DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

"Combining visualization and sonification techniques for monitoring and analyzing business process execution data"

verfasst von / submitted by

Tobias Hildebrandt

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Wirtschaftswissenschaften (Dr.rer.oec.)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 305 175
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Wirtschaftsinformatik
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Math. Dr. Stefanie Rinderle-Ma

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Vienna, ____________________ ____________________

Date Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am ____________________ ____________________

Datum Unterschrift

Abstract

Due to increased computerization and automation of business processes, the amount of available process data increases, which can make it more difficult to maintain a real-time overview of running processes and detect trends and developments in lager volumes of historical process execution data. Data visualization techniques are well-established in real-time monitoring and retrospective analysis of business process data, and these visualization techniques are typically combined with automated data processing, machine learning, and statistics. However, visualization poses certain limitations and challenges, such as the difficulty to focus on other work while performing visual process monitoring, or the limited number of visual dimensions onto which data can be mapped.

To alleviate some of these challenges, we propose to combine visualization and visual analytics with sonification (the representation of data using non-speech sound) to multi-modal systems. Sonification has several characteristics that make it suitable for monitoring and analyzing business process data, such as the possibility to monitor passively or human ability to detect even slight deviations in rhythm and pitch over time. However, as its application to the domain of business processes has hardly been researched, many open research questions remain in this area. This PhD thesis answers several of those open questions, e.g. how to effectively combine visualization and sonification to multi-modal systems for business process monitoring and analysis, and which sonification techniques and mappings are best suited for the tasks and requirements at hand.

To this end, the SoProMon system, a modular evaluation framework for multi-modal sonifications for process monitoring as a secondary task, has been developed. A quantitative and qualitative experiment based on this system has been devised. It aims to test the SoProMon framework and determine how different types of sonification can help in guiding attention.

One of the main results is that a variant combining visualization, auditory alerts and continuous sonification increases monitoring performance significantly over visualization alone. Another developed prototype combines both modalities to support users in the post-hoc analysis of business process execution data. While it seems that a tight integration of both modalities can support users in post-hoc data analysis as well, this will still need to be proven through qualitative and quantitative experiments. Overall, this thesis proves that sonification can be a valuable supplement to visualization in many areas, especially those involving real-time monitoring.

Kurzfassung

Durch die zunehmende Digitalisierung und Automatisierung von Geschäftsprozessen nimmt die Menge der verfügbaren Prozessdaten zu, was vor neue Herausforderungen für die Echtzeitüberwachung von laufenden Prozessen sowie die nachträgliche Auswertung von Prozessausführungsdaten stellen kann. In beiden Bereichen haben sich Techniken der Datenvisualisierung etabliert, welche meist mit automatisierter Datenverarbeitung, maschinellem Lernen und Statistik kombiniert werden. Die Visualisierung weist jedoch gewisse Schwächen und Herausforderungen auf, wie beispielsweise die Schwierigkeit, sich während der visuellen Prozessüberwachung auf andere Tätigkeiten zu konzentrieren, oder die begrenzte Anzahl von visuellen Dimensionen, auf welche Daten abgebildet werden können.

Um einige dieser Herausforderungen abzuschwächen, schlagen wir vor, Visualisierung und Visual Analytics mit Sonifikation (die Repräsentation von Daten mittels nicht-sprachlichem Klang) zu multi-modalen Systemen zu kombinieren. Die Sonifikation weist mehrere Eigenschaften auf, welche sie als sehr geeignet für die Überwachung und Auswertung von Prozessdaten erscheinen lässt, wie die Möglichkeit zur beiläufigen Überwachung, oder unsere Fähigkeit, kleinste Veränderungen in Rhythmus und Tonhöhe im Zeitverlauf wahrzunehmen. Da die Anwendung von Sonifikation auf den Bereich der Geschäftsprozesse bisher allerdings kaum untersucht wurde, gibt es in diesem Bereich noch einige offene Forschungsfragen zu beantworten.

Diese Dissertation beantwortet mehrere dieser offenen Forschungsfragen, beispielsweise wie Visualisierung und Sonifikation effektiv zu multi-modalen Systemen für die Überwachung und Auswertung von Geschäftsprozessdaten kombiniert werden können, oder welche Sonifikationstechniken und -abbildungen für die vorliegenden Arbeitsaufgaben und Anforderungen ideal sind.

Zu diesem Zweck wurde das SoProMon System, eine modulare Evaluationsumgebung für multi-modale Sonifikationen zur beiläufigen Prozessüberwachung, entwickelt. Eine damit durchgeführte Studie kommt zu dem Ergebnis, dass eine Variante welche Visualisierung, auditive Benachrichtigungen und kontinuierliche Sonifikation kombiniert, die Überwachungseffektivität im Vergleich zur alleinigen Visualisierung signifikant verbessern kann. Ein weiterer entwickelter Prototyp kombiniert beide Modalitäten, um Nutzer in der nachträglichen Auswertung von Geschäftsprozessausführungsdaten zu unterstützen. Es gibt Hinweise darauf, dass eine enge Integration von beiden Modalitäten dabei helfen kann, ein solcher Effekt muss aber noch durch qualitative und quantitative Experimente bewiesen werden. Alles in allem beweist diese Dissertation, dass die Sonifikation in vielen Bereichen eine sinnvolle Ergänzung zur Visualisierung darstellt, insbesondere in der Echtzeitüberwachung.

Acknowledgements

I would like to thank several individuals who played an essential role in the creation of this thesis:

- My supervisor Prof. Dr. Stefanie Rinderle-Ma, for her continuous support, trust, patience and guidance
- Dr. Thomas Hermann, for inspiration, feedback, cooperation and guidance
- Felix Amerbauer, for his instrumental contribution to the development of the ProSon prototype
- Dr. Paul Vickers, for providing helpful feedback and agreeing to be my second reviewer
- Current and former colleagues, for feedback and support (especially, but not exclusively, Jürgen Mangler, Simone Kriglstein and Conrad Indiono)
- Several members of the staff of CITEC at Bielefeld University, especially Eckhard Riedenklau, Jan Hammerschmidt, Matthias Sterz und Jens Hellriegel; for helping in executing user studies, writing project proposals, discussing research and in general being welcoming
- All participants of the focus group discussion held in Vaasa, Finnland, and of the SoProMon experiment conducted in Bielefeld, Germany
- Prof. Dr. Frieder Nake and Petra Klusmeyer, for seeing the potential of sonification and for their encouragement
- The digital media master's program of the University of Bremen, for preparing me for research
- The ICAD community, for providing a platform for exchange and inspiration
- All anonymous reviewers, for their feedback
- My wife, for support, feedback, direction, encouragement, and purpose. None of this would have happened without her.

Parts of this thesis have been inspired and funded by the EU FP7 project ADVENTURE (http://www.fp7-adventure.eu/) and SBA research (https://www.sba-research.org/).

Contents

List of Figures

1. Introduction

> 'You see, any aspect of a piece of music can be expressed as a sequence or pattern of numbers,' enthused Richard. 'Numbers can express the pitch of notes, the length of notes, patterns of pitches (...)'
>
> Douglas Adams, Dirk Gently's Holistic Detective Agency, 1988

Processes, such as those used in manufacturing, logistics or robotics, are become increasingly complex, while simultaneously becoming more automated and computerized (Handfield et al., 2014; Horton, 2015). Business process, which represent a specific form of processes, can be defined as follows:

A business process is a structured set of tasks or activities that are designed towards reaching a business goal (Weske, 2007).

Business processes can take a variety of forms, ranging from processes whose individual tasks are performed mainly by humans (such as an insurance claim process), to highly automated tasks with no or little human involvement at all (such as a production process in a modern factory). They have a beginning and an end with clearly defined input and output parameters, while the tasks or activities constituting the processes have users or roles assigned to them. This type of description, which includes activities involved in a business process, combined with a definition of control and data flows, is called a process model, while each execution of such a process model is called a process instance (Weske, 2007). In a scenario where a process model describes the production of a specific type of car, each individual car produced would represent an instance of that process (Hildebrandt, Hermann, and Rinderle-Ma, 2014). Figure 1.1 shows a simplified part of a business process of a manufacturing company in the standardized notation Business Process Model and Notation (BPMN)[1]. The diagram presents the different steps necessary to produce a product, and the sequence in which these steps need to be performed.

[1] http://www.bpmn.org/

Business process performance is often critical to a company's success - a delayed delivery of raw materials can, for example, lead to delayed production and ultimately to customer dissatisfaction. In many cases, therefore, the ability to monitor business processes execution in real time is crucial. The increasing amount of data, on the one hand, puts increasing pressure on users who need to observe processes. On the other hand, this increase in data that are recorded and processed electronically offers enormous potential to better monitor and control processes. However not only is it crucial to be able to monitor processes in real-time, it is also important to be able to effectively analyze historical process execution data, for example, to detect trends over time, or to find bottlenecks or other anomalies (e.g. fraud).

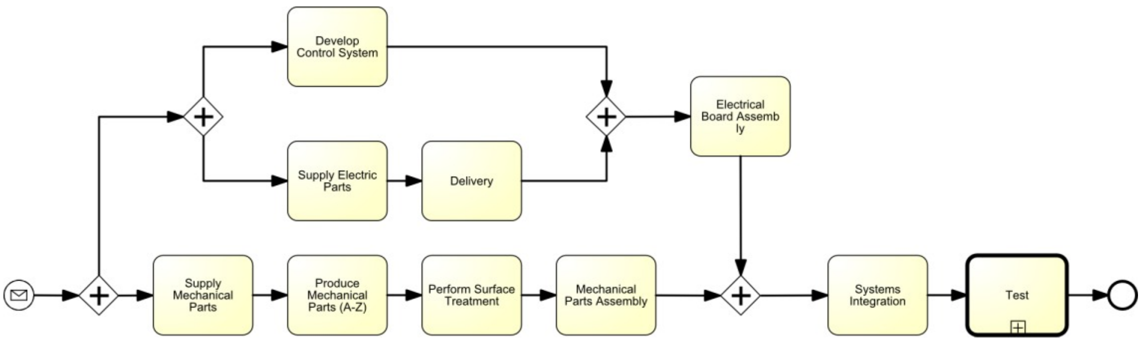


Figure 1.1.: Simplified example of a business process model.

Business Process Management (BPM) is a management approach that covers the design of processes, their operation (often also called enactment or execution) and the analysis of historical process execution data (sometimes also referred to as evaluation phase). This analysis frequently leads to an improvement or redesign of the process model, often referred to as the change phase. Therefore, the different steps are often represented as phases of a continuous life cycle. Although there are currently no standardized naming or numbering protocols for different phases, this thesis uses the phases depicted in Fig.1.2. This figure depicts the main tasks that users want to perform in each respective life cycle phase, along with the main challenges that arise during each task. Other definitions of the BPM life cycle include, for example, the phases *Design and analysis*, *configuration*, *enactment* and *evaluation* (Weske, 2007).

A term that is closely related, and often used synonymously with the term business process is workflow. The Workflow Management Coalition (WfMC) defines workflow as: "The automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules" (Weske, Aalst, and Verbeek, 2004). Even though the concepts of BPM and workflow management are often used synonymously, BPM is a more encompassing term, while workflow management concentrates mostly on the technical implementation and execution of business processes.

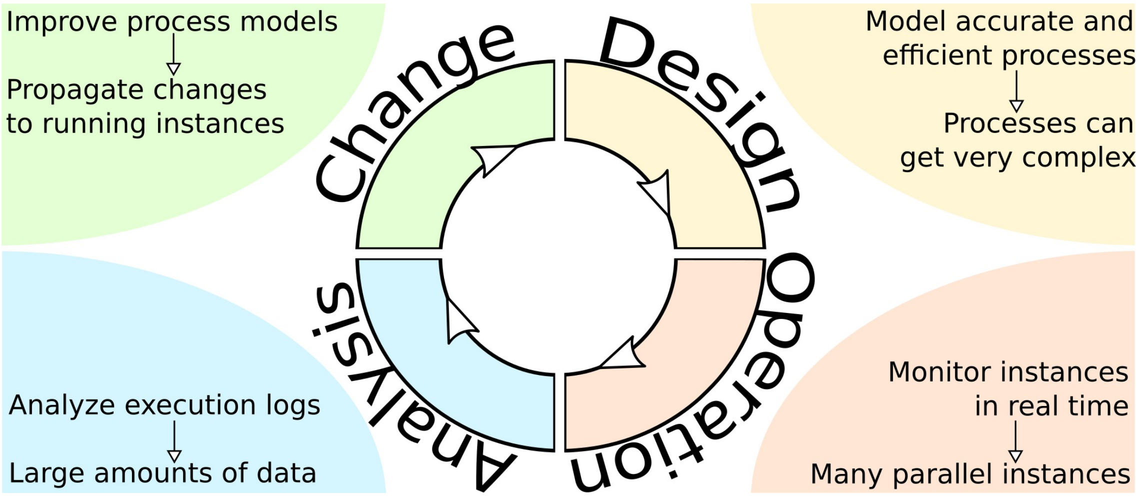


Figure 1.2.: Business process life cycle with main tasks and challenges.

There is an array of software that supports business processes in all phases of the life cycle. All those systems can be subsumed under the term PAIS (Process Aware Information Systems). The most comprehensive types of PAIS are BPM Systems, which usually support users in all phases of the life cycle. While workflow engines typically concentrate on the enactment/execution of processes (operation phase), BPM systems also support other phases such as process analysis (Weske, Aalst, and Verbeek, 2004).

Different user groups are involved in the different phases of the life cycle. Gadatsch (2009) defines the following roles for organizations that have already adopted BPM:

- CPO (Chief Process Officer): responsible for adoption and integration of the BPM system and for process analysis, simulation and optimization on a strategic level
- Process owners: responsible for reaching targets and process optimization on a day-to-day basis
- Process employees: responsible for modeling, execution and monitoring of individual tasks and components of the processes

There is research that aims to support users in their assigned tasks during the various phases of the life cycle. However, considering the importance of generating business insights, process monitoring and analysis in particular seems to be an area neglected by research thus far. Van der Aalst (Aalst, 2012) defined 20 typical BPM use cases by analyzing papers from past BPM conferences. He pointed out that there has been remarkably little research in this area so far, especially given the needs of industry in terms of business process monitoring:

"In this context it is remarkable that the use cases monitor (mon) and analyze performance using event data (perfed) have a much lower frequency (...) Given the practical needs of BPM one would expect more papers presenting techniques to diagnose and improve the performance of business processes." (Aalst, 2012)

Both, real-time monitoring and retrospective performance analysis, are sometimes encompassed by the term Business Process Intelligence (BPI). Similar to Business Intelligence (BI), BPI deals with generating insights into business-related data, but is tailored specifically to business process execution data. This data is typically available in the form of process execution logs, which contain entries for all the events that occur during execution. As, depending on the process, hundreds of thousands of these logs entries might accumulate every day (or even per hour), of course for process owners, controllers, or control center operators it would, depending on the number of process events, be either tedious or outright impossible to read all log file entries. Therefore, BPM systems typically present aggregated data, and different means of statistics and machine learning are applied to derive trends or forecast developments.

However, as we are well equipped to see patters in unstructured data, the use of visual analytics is recommended to convey business process-related data to the user (Aalst et al., 2012). Most existing BPM systems already include at least rudimentary visualizations for process monitoring and analysis. However, systems that rely only on the visual sense to present data to the user face certain challenges. Process monitoring challenges include the difficulty of passive monitoring, which in turn prevents users from simultaneously completing other tasks. Users must focus their attention a screen at all times in order to remain aware of potentially critical developments. Another challenge is the fact that a large number of process instances in different execution states can exist in parallel. Depending on the scenario, frequently updated data streams, such as sensor data, need to be observed as well. In process analysis, challenges can include the detection of trends and developments over time, or the detection and analysis of anomalies in large amounts of data. The visualization of high-dimensional process events is especially challenging, as there are a limited number of visual attributes onto which data can be mapped. This limitation can be alleviated by supplementing visualization with another means of data representation - sonification.

Sonification is the use of non-speech audio to convey information or perceptualize data (Kramer, 1994). Kramer et al. (1999) found that auditory perception is particularly sensitive to temporal change. Process executions are time-based by nature; therefore, sonification seems to lend itself to this purpose. As sound can be processed more passively than visual elements, and does not require a specific head orientation, sonification can further alleviate the challenge of constant visual attention during process monitoring. Therefore, it should be investigated whether combining existing techniques of visual analytics with sonification to multi-modal systems can better support users during their tasks in process monitoring and analysis.

1.1. Sonification for Monitoring and Analysis of Data

Examples of the usage of sonification in practice are the Geiger counter that displays radioactive radiation, or the auditory parking aid that conveys the distance to obstacles by beep sounds. Apart from these rather simple forms of sonification, research that uses sonification to represent more complex high-dimensional data exists. Spehr (2008) concludes that sonification is more suitable than visualization for complex, irregular or even chaotic data. This promises advances when trying to convey process exceptions and changes to users. Moreover, sonifications can very well be recognized, remembered and recalled later on (Hermann, 2002). We as humans are equipped with highly developed listening capabilities that are equally as powerful as our visual processing capabilities, but differ in terms of characteristics and bandwidth. Listening offers a number of benefits that enable sonification to enhance visualization in many applications, and in some scenarios, surpass the capabilities of visual interfaces.

Aside from auditory alerts and warnings, more complex sonifications can represent continuous values by mapping data to auditory features of sound, such as pitch, level, duration, timbral characteristics, brightness, roughness or spatial location of sound events in a stereo or multichannel display. As listening to sound does not require users to focus their attention on a specific visual display, it is already commonly applied in cars, cockpits or operation theaters, usually combined with visual displays. However, as sonification can be accessed and interpreted without any visual cue, it is also very suitable as sensory substitution for visually-impaired or blind users.

An example of how we utilize complex listening skills in everyday situations is car drivers, who can detect even subtle changes in the sound of their car engine and recognize if there is a problem, even if they do not have any technical knowledge about car engines. This mode of listening has been defined as *monitory listening* (listening, whether something is wrong; (Supper and Bijsterveld, 2015)). If a problem has been detected, the technique of *diagnostic listening* (Supper and Bijsterveld, 2015) can then be applied to determine what exactly is wrong. Other modes of listening are *explorative listening* and *synthetic listening*, both used to detect and decipher patters in sound (and thus in the underlying data; (Supper and Bijsterveld, 2015)).

To summarize, sonification (especially in combination with visualization) offers several benefits compared to visualization that make it suitable to alleviate some of the aforementioned challenges of visual-only process monitoring and analysis:

- No visual attention required
- Rapid detection & processing
- Habituation and sensitivity to changes
- Ability to guide the users' attention
- In combination with visualization: enhanced range of data dimensions

Although all phases of the BPM life cycle may benefit from this additional means of data perceptualization, the use cases (a) monitoring and (b) analysis are expected to benefit the most. This can be attributed to several factors, including the fact that sonification, due to its temporal nature, is best suited for temporal data, such as process execution data (which needs to be monitored and analyzed during the monitoring and analysis phases, respectively). For the use case monitoring, the expected benefit for the users is perhaps more obvious than for the use case analysis, as during monitoring it is typically more critical to react to process developments in time, while users often need to work on other tasks at the same time. Therefore, the bigger focus of this thesis is put on process monitoring compared to analysis. In this thesis, the term monitoring is being defined as observing process execution data in real-time, while analysis refers to the task of analyzing historical process execution data.

In the process design phase, on the other hand, users are mainly concerned with creating process models, which consist mostly of spatial and structural data, a data type for which visualization may in many cases be better suited - even though first approaches for enhancing this task with audio exist (Gulden, 2014).

1.2. Research Gap

A range of sonification approaches for monitoring different types of processes already exist, as summarized in (Vickers, 2011), as well as approaches for sonification-based retrospective data analysis. However, none of those approaches are targeted specifically at business process execution data. As business process data possesses certain characteristics that do not prevail in data of other domains (such as the hierarchy between processes and instances), the feasibility of transferring such research results to the BPM domain still needs to be investigated. Furthermore, it is still unclear if sonifications and multi-modal systems that support the tasks of monitoring and analysis of business process data require different approaches than such that are designed for monitoring and analysis in other fields. Open questions in this area concern, for example, how sonification can help to guide the operator's attention during monitoring as a secondary task - a typical scenario especially in smaller and medium-sized enterprises.

A significant research gap exists concerning multi-modal systems that combine the two modalities visualization and sonification, as a majority of related sonification research seems to focus on single-modal sonification. Finally, the main share of sonification approaches has not yet been evaluated in qualitative or quantitative user studies. Thus, an open question remains as to whether the selected design decisions and mappings are the most effective and suitable in their respective cases. The research gaps for the two tasks are analyzed in more detail after the related work in Chapter 3.

1.3. Methodology

Major portions of the research gap concern guidelines for the application of sonification techniques and mappings for business process execution (or similar) data, while for visualization, one can draw on existing research in many cases. Therefore, although this thesis explicitly investigates multi-modal systems combining both modalities, more emphasis is placed on sonification compared to visualization.

Based on the research gap, this thesis answers the following main research questions:

- What are the peculiarities of business process execution data and what implications do these have with regard to sonification?
- Which user requirements exist concerning multi-modal solutions during the process operation and analysis phases?
- What implications do user requirements have on possible sonifications in process monitoring and analysis?
- Which sonification techniques and mappings are best suited to the requirements defined for real-time data analysis and the analysis of historical instance execution data?
- What are the strengths and weaknesses of visualization and sonification techniques during the operation and analysis phases?
- How can visualization and sonification techniques best be combined to support users during the real-time monitoring and post-hoc analysis of business process execution data?
- Do multi-modal solutions better match the requirements than visualization or sonification alone during the phases operation and analysis?

More detailed sub-questions are presented later in different sections of the thesis. For single-modal systems based on visualization and visual analytics, many of those questions have already been answered.

The process of answering most of the stated research questions has been guided by a design science approach, which is tailored to research in information systems (Wieringa, 2014). It is based on the development and evaluation of artifacts that are built to address unsolved problems. Such artifacts include, among others, algorithms (e.g. for information retrieval), human/computer interfaces and system design methodologies (Hevner et al., 2004).

In contrast to other research approaches such as explanatory research, academic research objectives are often pragmatic (Vaishnavi and Kuechler, 2015). Most of the previously defined research questions can be answered through the design and evaluation of artifacts (see Sec. 1.5 for a detailed list) such as multi-modal prototypes, that have been developed and evaluated, and thus lend themselves to a design-science based research approach. As the topic of this thesis is rooted in information systems research, yet interdisciplinary in nature (including areas such as Human-Computer Interaction (HCI)), not all of the stated research questions have been guided by design science research alone. For example, questions regarding the effectiveness of the different modalities and how they can direct attention might better be tackled using explanatory approaches, and their results should be transferable to research areas other than information systems. In many cases, HCI and sonification research primarily intends to describe and understand a specific phenomenon, while "Design science research by definition changes the state of the world through the introduction of novel artifacts" (Vaishnavi and Kuechler, 2015, page 31). Thus, while the development of prototypes in this thesis is in fact a constructive creation of artificial artifacts (something inherent to design science research) many of the questions that these prototypes try to answer are empirical in nature, as they try to describe reality (i.e. human perception and behavior). Furthermore, some of the research questions in the area of sonification (and of this thesis) are not directly based on real-world business problems (which is one of the main goals of design science research), but instead are driven by research interests. Therefore, this thesis pursues a mix of design science and explanatory research approaches. The concrete research methods that have applied to generate requirements and to evaluate prototypes have been a mix of qualitative and quantitative methods and include a focus group study, experiments and questionnaires.

1.4. Structure of this Thesis

This thesis is structured by first presenting an overview of existing research and a discussion of its foundations, before presenting requirements for multi-modal sonification in monitoring and analysis. The main results of the thesis are then presented along the BPM life cycle phases *monitoring* and *analysis* in subsequent chapters.

Chapter 2 investigates the tasks of monitoring and analyzing business process execution data, starting with an analysis of its data structure, and followed by an analysis of the peculiarities of the two tasks in practice, as well as of their challenges. Chapter 3 outlines the foundations of data visualization, sonification and multi-modal displays in general as well as their benefits and challenges, and presents the related work relevant to this thesis along the lines of sonification for process monitoring and sonification for data analysis. The chapter concludes with a detailed analysis of the research gap in the stated areas. Chapter 4 presents an analysis of the suitability of the different existing sonification techniques for business process data, as well as an analysis of the requirements for the tasks of process monitoring and analysis, and their implications on multi-modal systems in these areas. Chapter 5 presents the SoProMon system, a modular evaluation framework for multi-modal sonifications for process monitoring as a second task. In the same section, a quantitative and qualitative experiment based on this system is presented. The experiment aims to test the SoProMon framework, in addition to determining how different types of sonification can help in guiding attention in a dual-task scenario for process monitoring compared to a visual-only system. The results of this experiment are presented and discussed. Furthermore, this section contains two further, different sonification concepts for process monitoring. Chapter 6 presents an approach for multi-modal process analysis. This thesis concludes with a summary and discussion in Chapter 7.

As sonifications, unlike visualizations, are not easily conveyed in a static PDF document, external sound and video files are referenced throughout the thesis. This is indicated in the document by the following speaker symbol:

In the PDF version, a click on such a speaker symbol will lead to a URL that contains an audio file.

1.5. Contribution

The main outcomes of this thesis are several artifacts, the majority of which have already been presented in conference and journal papers. The following list presents the main outcomes, the chapter in which they are presented in this thesis, and where these results have been published. A substantial portion of the contents in these chapters is based on the respective publications.

1. A systematic literature review concerning sonification in computer security (Sec. 3.3.2)
   - Tobias Hildebrandt and Stefanie Rinderle-Ma (2015). "Server sounds and network noises." In: *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 45–50. DOI: 10.1109/CogInfoCom.2015.7390562
2. An analysis of research concerning sonification of data of similar structure as business process data (Sec. 3.3.3), leading to the development of best practices for business process monitoring (Sec. 4.4)
   - Tobias Hildebrandt and Stefanie Rinderle-Ma (2013). "Toward a Sonification Concept for Business Process Monitoring." In: *19th International Conference on Auditory Display (ICAD 2013)*. Lodz, pp. 323–330
   Figures from this paper are used in Fig. 5.35 and Fig. 5.37.
3. An analysis of existing sonification techniques and their suitability for the phases of the BPM life cycle (Sec. 4.3)
   - Tobias Hildebrandt, Simone Kriglstein, and Stefanie Rinderle-Ma (2012b). "On Applying Sonification Methods to Convey Business Process Data." In: *CaISE 2012 Forum*. CaISE Forum. CEUR. URL: http://eprints.cs.

univie.ac.at/3449/
Figures from this paper are used in Fig. 4.3.
- Tobias Hildebrandt, Simone Kriglstein, and Stefanie Rinderle-Ma (2012a). "Beyond Visualization: On Using Sonification Methods to Make Business Processes More Accessible to Users." In: *18th International Conference on Auditory Display (ICAD 2012)*. Georgia Institute of Technology, pp. 248–249

4. The outcomes of a focus group study concerning the requirements of users in industrial production monitoring (Sec. 4.1)

   - Tobias Hildebrandt, Jürgen Mangler, and Stefanie Rinderle-Ma (2014). "Something doesn't sound right: Sonification for monitoring business processes in manufacturing." In: *LABEM 2014 : 1st International Workshop: Lowering the Adoption Barrier of Enterprise Modelling*. Geneva, Switzerland: IEEE. URL: http://eprints.cs.univie.ac.at/4066/
   Figures from this paper are used in Fig. 4.1 and Fig. 5.36.

5. A system for evaluating sonifications for process monitoring as second task (the SoProMon system) (Sec. 5.1)

   - Tobias Hildebrandt, Thomas Hermann, and Stefanie Rinderle-Ma (2014). "A sonification system for process monitoring as secondary task." In: *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 191–196. DOI: 10.1109/CogInfoCom.2014.7020444
   Figures from this paper are used in Fig. 5.7.

6. An evaluation scenario for the SoProMon system, including questionnaires, study procedures and data measurement and -analysis methods (Sec. 5.2)

   - Tobias Hildebrandt, Thomas Hermann, and Stefanie Rinderle-Ma (2016). "Continuous sonification enhances adequacy of interactions in peripheral process monitoring." In: *International Journal of Human-Computer Studies* 95, pp. 54–65. ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2016.06.002. URL: http://www.sciencedirect.com/science/article/pii/S107158191630074X
   Figures from this paper are used in Fig. 5.12-5.33.

7. Statistical analysis of the evaluation of the SoProMon system and performance comparisons of process monitoring as a second task between the modes visual-only, visual+auditory alarms and visual+continuous soundscape sonification (Sec. 5.3, 5.4, 5.5).

   - Thomas Hermann, Tobias Hildebrandt, et al. (2015). "Optimizing Aesthetics and Precision in Sonification for Peripheral Process-Monitoring." In: *International Conference on Auditory Display (ICAD) 2015*. Proceedings of the 21st International Conference on Auditory Display. Graz, Austria, pp. 317–318. ISBN: 978-3-902949-01-1. URL: http://eprints.cs.univie.ac.at/4395/

- see (Hildebrandt, Hermann, and Rinderle-Ma, 2016)

8. A prototype for analyzing historical process data in the form of a ProM plugin[2] (Sec. 6)

   - Tobias Hildebrandt, Felix Amerbauer, and Stefanie Rinderle-Ma (2016). "Combining Sonification and Visualization for the Analysis of Process Execution Data." In: *2016 IEEE 18th Conference on Business Informatics (CBI)*. Vol. 02, pp. 32–37. DOI: 10.1109/CBI.2016.47
     Figures from this paper are used in Fig. 4.2 and Fig. 6.3-6.8.

   The presented prototype was developed in the context of the following master's thesis:

   - Felix Amerbauer (2017). "Multi-modal Sonification of Business Process Event Logs." MA thesis. Austria: University of Vienna

9. A Concept for motif-based monitoring (Sec. 5.6).
   - see (Hildebrandt, 2013).
10. A concept study for soundscape-loop based monitoring (Sec. 5.7).
    - see (Hermann, Hildebrandt, et al., 2015).

The main contributions of this thesis in the operation and analysis phases of the business process life cycle are summarized in Fig. 1.3.

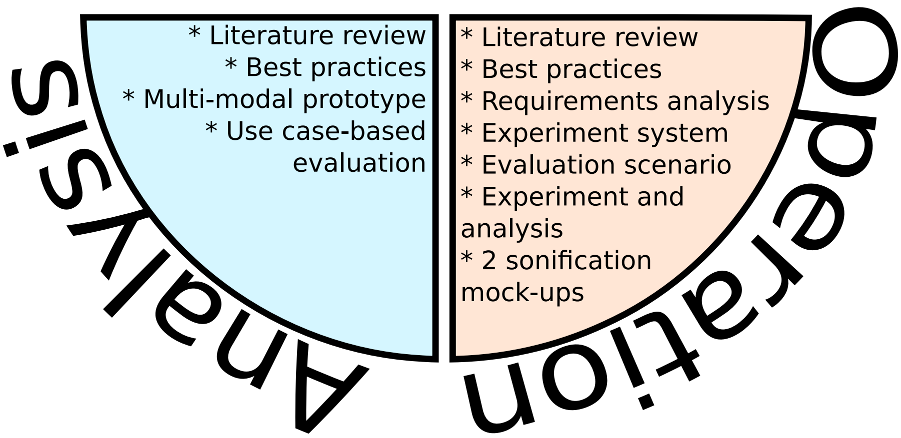


Figure 1.3.: Main contributions in business process operation and analysis.

Together, these contributions build the first step towards integrating (multi-modal) sonification into BPM- and workflow systems.

[2]http://www.promtools.org

2. Challenges of Monitoring and Analyzing Business Process Execution Data

This chapter discusses the structure of business process execution data, as well as how it is being monitored and analyzed in practice, and the challenges that arise thereof. Various process-aware information systems (PAIS) exist, ranging from systems that are mainly reasearch-oriented (like the Cloud Process Execution Engine CPEE[1]), over open-source systems like Apache ODE[2], the Activity BPM Platform[3] or the jBPM suite[4], to commercial applications such as Apian BPM[5] or IBM Business Process Manager[6].

Ideally, PAISs create event logs during process execution, for example in the XES (eXtensible Event Stream) standard[7], that can then be imported and analyzed using process mining software, for instance. Process mining refers to a bundle of techniques operating on business process execution data for example, to discover process models, check the conformance of execution data with a given process model, and conduct process performance analysis (Aalst et al., 2012).

There exists a variety of such systems, both open source (e.g. ProM[8]) and commercial (e.g. Fluxicon Disco[9], QPR Process Analyzer[10] or Celonis Process Mining[11]). However, in practice not all PAISs create event logs in a standardized format.

The tasks of monitoring and analyzing business process-related data are often subsumed under the term *Business Process Intelligence*, which is the application of business intelligence techniques (technologies and methods for data collection, integration, analysis and presentation of corporate information) to business processes. This happens through the pre-processing and aggregation of process

[1] http://cpee.org
[2] http://ode.apache.org
[3] http://activiti.org/
[4] http://www.jbpm.org/
[5] http://www.appian.com/business-process-management-software/)
[6] http://www-03.ibm.com/software/products/en/business-process-manager-family
[7] http://www.xes-standard.org/
[8] http://www.promtools.org/
[9] https://fluxicon.com/disco/
[10] http://www.qpr.com/products/qpr-processanalyzer
[11] http://www.celonis.com

logs and the analysis of this aggregated data. The following features of business process intelligence can be distinguished (Grigori et al., 2004):

- The analysis of completed process executions
- The derivation and application of prediction models
- The monitoring and analyzing of running process instances
- The intervention with business process management systems to avoid undesired situations

2.1. Business Process Execution Data

For each enactment of a business process a process instance is created, and for the execution of each of the different tasks that make up the process (sometimes also called *activities*) one or more events are logged. The sequence of events within such a log is often called *trace*. The execution of a process instance (sometimes also called a *case*) usually entails several or (for large process models or processes with many cyclic parts) sometimes even hundreds or thousands of such individual events (see Fig. 2.1). (Aalst et al., 2012)

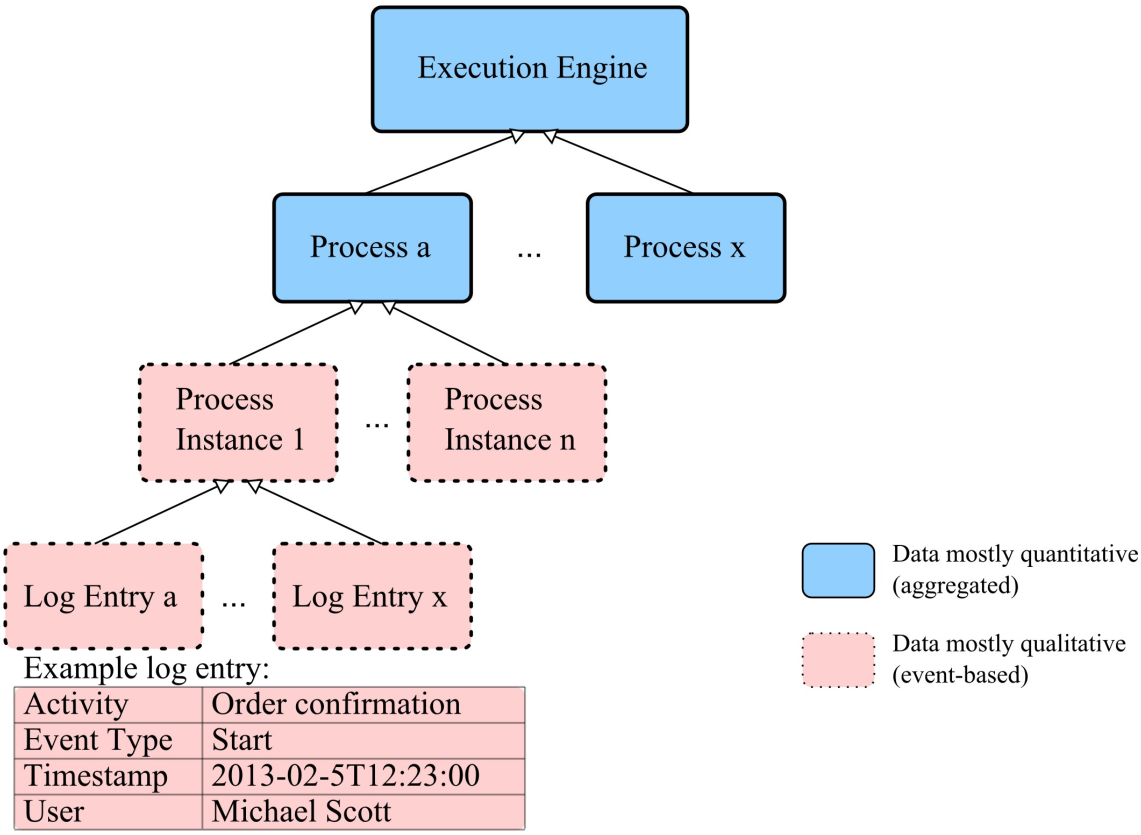


Figure 2.1.: Structure of business process execution data.

A log entry often consists of the name of an activity that has been performed or is about to start, a time stamp and an assigned user. Depending on the PAIS, log entries can also be created for other events, like the modification of variables or the occurrence of errors during execution. Aside from this minimum set of information, depending on the organization and the PAIS in place, an arbitrary amount of additional semantic and quantitative information can also be included in the log entries. Individual event logs thus typically contain the following data types:

- Time-based data: the timestamp of occurrence of an event (e.g. the start of an activity).
- Nominal data: e.g. the name of the concerned activity or variable, the name of a responsible user and/or department.
- Relational data: the relation of an event log to a process instance and a process.
- Quantitative data: data that is related to an event, like the price of an ordered item for an order event, or the duration of an activity.

Such event logs are created by PAISs in different formats such as in the aforementioned XES standard, in the previous de-facto standard MXML (Mining eXtensible Markup Language)[12], or in vendor-specific, non-standardized formats.

In the following, an individual log entry from an example trace in the XES format is presented:

```
<event>
<string key="concept:name" value="Send second reminder"/>
<string key="concept:instance" value="shutdown-process/256"/>
<string key="id:id" value="a11"/>
<string key="lifecycle:transition" value="start"/>
<date key="time:timestamp" value="2016-12-09T07:22:31+01:00"/>
</event>
```

In general, process events can be classified into three categories:

- Control flow perspective (regarding the ordering of activities)
- Data flow perspective (changes in variables)
- Organizational perspective (involved users and their roles)

The individual events are typically aggregated to quantitative data on the process and engine levels. Typically such aggregations occur in order to calculate so-called Key Performance Indicators (KPIs). Such KPIs can be calculated for single process instances (like the current execution time in relation to the average execution time for instances of the respective process), but are typically calculated either over all instances of a specific process (e.g. the share of instances in faulty states) or over all instances of all processes that an organization runs (e.g. the average process instance overdue time).

[12] http://www.processmining.org/logs/mxml

A special challenge of business process execution data, therefore, is that it differs in many aspects from data of many other application domains, which is often primarily of quantitative nature. The essence of the data at hand, however, is based on discrete, temporal events that contain mostly qualitative, semantic data (like the name of a user or a department). In contrast, the individual log entries are often linked to related quantitative data variables, like sensor values, and are also linked to a specific process instance and a process model. On these higher levels, event- and instance-spanning KPIs are quantitative in nature. Due to this specific data structure, *classical* systems for data analysis are, without adaptations, only partly suitable for analyzing business process execution data.

Another challenge is that the amount of data that accumulates over a certain period of time varies strongly, depending on the types of business processes an organization runs. In labor-intensive, long running processes it may happen that there is only one running instance at a time (if, for example, a small workshop produces only one unit at a time), with only a few events occurring per day. On the other hand, with highly automated manufacturing processes in which, for example, events are triggered based on sensor data, thousands of instances can exist that may possibly lead to thousands of occurring events per second. This puts special requirements on high-performance data processing and monitoring.

2.2. Monitoring Business Processes

As stated in the introduction, in many domains it is crucial for organizations to keep track of process executions in real-time, for example, in order to avoid standstills or delays and thus loss of income, unsatisfied customers or even contractual penalties. Organizations can have hundreds or more business process models (Yan and Grefen, 2011), while there are usually many instances of those process models running at any given moment. Those instance executions can be forced to a stop due to, for example, technical problems, and a high-level overview of all instances of a process can give an accurate overview of a process' or even a company's performance. Therefore, enterprises want to keep track of the executions of their processes, often in real-time. Depending on the domain and the user, there are different types of process-related data that users want to monitor. Furthermore, the users would like to be informed about exceptional situations, such as delays or certain changes during process execution. These can be caused by activities that deviate from their usual order, or, for example, production-related sensor values exceeding a certain threshold. Additionally, users are interested in obtaining a broader overview of the performance not only of individual resources, such as specific machines or departments, but also of entire processes.

2.2.1. Practice

Different user groups (e.g. maintenance staff, engineers, supervisors and managers) rely to different extents on monitoring systems that present the data mainly by visual information displays. In large-scale process monitoring, for example in large manufacturing settings, monitoring is typically performed in control centers where users observe production on multiple screens, using video features as well as schematic overviews of processes and machines/facilities, charts/graphs, textual descriptions and alerts (Sauer, 2004).

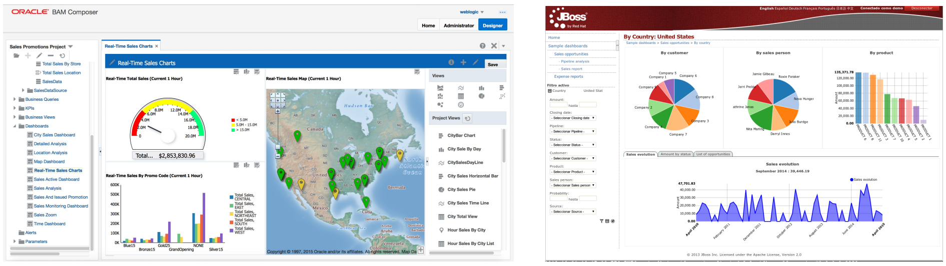
For scenarios without dedicated control centers, process monitoring systems also offer dashboard and cockpit views that aggregate individual process execution events to KPIs and present them in real time, using visualization techniques such as speedometers, as shown in Fig. 2.2 and Fig. 2.3. Such systems offer graphical representations of process performance-related indicators, while notifications about the individual events, on which the higher-level representations are based, are available in textual form. Some monitoring applications provide graphically annotated views of process models, in which the transitions of currently executed instances between the processes' activities of a particular process model can be observed in real-time.

Figure 2.2.: Dashboard-based process monitoring in manufacturing (ProVis.Agent by Fraunhofer IOSB). Photo from press release (`https://idw-online.de/de/news368643`).

Examples of such applications are ARIS Process Performance Manager[13] or IBM Business Monitor[14]. A variety of commercial systems that support users during business process monitoring exists. SAP, for example, offers *Business process monitoring*[15], a solution with different dashboards that present mainly KPIs and high-level process-relevant data, and optional alert notifications that can also be

[13] www.softwareag.com/corporate/products/aris_platform/aris_controlling/aris_process_performance/
[14] http://www-03.ibm.com/software/products/en/business-monitor
[15] https://wiki.scn.sap.com/wiki/display/SM/Business+Process+Monitoring

sent by email or SMS. In Oracle BAM[16], users can create dashboards in which KPIs can be visualized, and the program also enables to drill down from the KPIs to the individual events on which those KPIs are based. JBoss jBPM supports



(a) Oracle BAM (http://www.oracle.com/technetwork/middleware/bam/overview/index.html).

(b) JBoss BPM business activity monitoring (http://www.oracle.com/technetwork/middleware/bam/overview/index.html).

Figure 2.3.: Example views of business monitoring applications.

the ability to build a dashboard for visual real-time monitoring of KPIs as well. The aforementioned ARIS Process Performance Manager (PPM) enables both real-time and post-hoc process analysis. Users pay attention to these dashboard overviews periodically, while at the same time interrupting other activities they are currently working on, unless they are full-time monitoring employees.

In contrast to visuals, sound is typically only used as a means to convey warnings and alerts in environments such as modern production settings. In these cases, sounds usually indicate an alarm signaling when a machine has broken down or a predefined threshold has been exceeded (Siemens AG, 2007). In a production scenario, for instance, this could for instance be the case when the stock level of a resource has dropped below a critical level, or when a temperature sensor of a machine has measured a critical temperature, indicating imminent machine failure (SAP SE, 2015).

Supervisors are often more interested in KPIs, such as current average throughput times, energy consumption, or other performance-related indicators. Typically, the higher a person is in the hierarchy of a company, the less he or she is interested in monitoring individual execution events and more in obtaining aggregated data (usually in the form of KPIs; (Petzmann et al., 2007)).

To conclude, the various existing PAISs share many similarities in terms of functionality in real-time process monitoring. Most approaches offer customizable visualizations, usually in the form of a dashboard view that incorporates elements such as speedometers and bar charts. Furthermore, in most cases alert rules can be created, which, for example, can be triggered once a predefined threshold value has been exceeded. These alerts are usually shown in the dashboard, but are often also sent via SMS or email.

[16] http://www.oracle.com/technetwork/middleware/bam/overview/index.html

2.2.2. Challenges

Particularly in smaller organizations, process monitoring is typically a passive activity, which is usually performed while the user concentrates primarily on another task, in contrast to, e.g., process analysis, a task to which users typically dedicate their full attention. Thus, the current practice in process monitoring poses the following drawbacks:

- In typical monitoring scenarios without dedicated full-time monitoring personnel, users such as engineers or supervisors periodically glance at their monitoring application. As a result, they may see time-critical events or alerts too late to take action, while simultaneously becoming distracted from their main task.
- Some environments, such as in large production facilities, have full- time personnel dedicated to monitoring processes. In these cases, these users cannot effectively perform other tasks concurrently.
- Visual attention and a specific head orientation are required.
- Information is often conveyed in textual form, but users can only read a certain amount of text in a given time range.
- For situations in which a high number of events occur, e.g., with automated production, there can be several KPIs that need to be monitored. Screen space however is limited.

As previously mentioned, domains such as industrial production, involve systems that already apply auditory alarms and alerts. These auditory signals, however, typically only convey events requiring immediate action, by signaling, for instance, that a predefined threshold value has been exceeded. They often do not convey the exact nature of the alert or problem, forcing users to check a machine or screen to determine if user action is required, and which specific action is required.

Furthermore, many systems do not consider information leading up to an alert prior to reaching a predefined threshold, or information without existing defined alert rules. Such information could be of interest to users from a preventative standpoint. Moreover, as alerts and alarms only convey the occurrence of a (supposedly) exceptional situation that requires immediate action, they are not designed to be aesthetically pleasing but to grab immediate attention. However, especially if thresholds are defined too low and therefore *unnecessary* alerts are raised, over the course of a work day they can be annoying and distracting, in particular to those people who are not targeted by the alarm. Furthermore, critical situations are difficult to define: on the one hand, if rules that define alert-triggering thresholds are defined too *conservatively*, i.e., requiring strong evidence before issuing positive classifications, potentially critical situations such as machine failures might occur without issuing an alert.

On the other hand, if the values are defined too *liberally*, i.e. risking high false positive rates, the resulting flood of often unnecessary alerts and alarms might lead to information overload for the user, or a situation in which the user stops to take the alerts seriously. Furthermore, in many scenarios engineers are not able to define all states and values that might lead to a critical situation beforehand. Levels and values that might constitute a critical state are often complex to establish, as, for example, the question of whether a specific parameter value constitutes a critical situation often depends on the context, given by various other parameters.

But even if all possibly critical situations are covered by alerts and alarms, in most cases operators might prefer to be informed even *before* a situation might become critical, thus enabling them to *anticipate*, *intervene* and *prevent* the problem. A constant awareness of states and values through an *auditory ambient information system* might enable such an anticipation of critical situations.

2.3. Analyzing Business Process Execution Data

Business process execution data is a valuable source for answering process-related questions, and many of these questions can be answered with process mining (Aalst et al., 2012). Process mining has three main goals (Aalst, 2011):

- (a) Process discovery
- (b) Conformance checking
- (c) Process enhancement.

The goal of process discovery is to automatically discover a process model by analyzing logs of recorded process events. It is typically done by enterprises that have either not explicitly defined all their business processes yet, or that have done so, but have no software system in place that enacts and observes them. To this end, event logs from different systems (e.g. BPM systems, ERP (Enterprise Resource Planning) systems, CRM (Customer Relationship Management) systems or other types of information systems) are analyzed in order to find out what processes exist in a company, and how they are defined. The processes that are discovered this way can then serve as a basis for later process improvement, or for enactment and observation by PAISs.

In conformance checking, executed process instances are compared to the previously defined process models (or to *typical* process executions) with the aim of finding deviations. On one hand, one can rate a mined process by asking how far it deviates from the process log, i.e. how many cases are valid instances of the mined process. Some PAISs are also flexible enough for (authorized) staff to deviate from the predefined process model by applying ad-hoc changes to a running process instance, like adding a new task or changing the sequence of preexisting tasks. PAISs that support such flexiblity include the CPEE and the AristaFlow BPM suite[17]. In such a case, conformance checking can be used to find deviations from the *official* process models.

[17]http://www.aristaflow.com/de/bpm-suite/ueberblick.html

Process mining techniques are also used for process enhancement, for example to identify bottlenecks and weaknesses of processes. For the analysis phase of the BPM life cycle, goals (b) and (c) are particularly important. Other goals that users have during the analysis of process execution data might not necessarily need process mining techniques to achieve, but can instead by achieved by *traditional* means of data analysis, such as statistical analysis or data visualization. Therefore, not all user questions related to their process data base represent one of the three main goals of process mining - in many cases, users may want to analyze process performance, without having the primary intention of improving the process itself. Some optimization results that can arise out of process analysis may not directly affect a process itself, but rather the resources assigned to it (human or otherwise), or business decisions independent of the process. Another example is anomaly detection. Even though there is no unified definition, an anomaly can be defined as (i) a rare or infrequent event; (ii) a deviation from a normal form or rule; (iii) an unexpected result; or (iv) a state outside the usual range of variations (Bezerra, Wainer, and Aalst, 2009). Anomaly detection is not necessarily performed to reach one of the three stated goals of process mining, but can also be applied, for example, to detect intrusions.

In summary, user questions that may arise during the process analysis phase can include:

- How busy was the system/were the processes over time?
- How *popular* have the different processes been?
- What was the average lead time?
- Which paths of a process have been taken the most? Is it, for example, necessary to assign further employees to certain tasks?
- Have there been errors/problems, such as technical problems, deadlocks, etc.?
- Have there been anomalies or deviations during process execution?
- Have there been intrusions?

Those aspects of process analysis that are not covered by the previously stated definition of process mining are covered by the earlier defined term *business process intelligence*.

2.3.1. Practice

As stated in the introduction, different groups of users perform business process data analysis for different purposes, depending on the organization. Typical users are CPOs, process owners, and managers or supervisors in general, who are interested in the performance of specific processes. Business analysts and data scientists use process execution data data to perform process discovery, or to improve existing processes.

Different tools that support users in the previously defined tasks exist. BPM systems and other PAISs offer, as previously mentioned, dashboards and other options for monitoring process execution. The same dashboards that are used to monitor in real-time can usually also serve for post-hoc data analysis, for example to conduct process improvement or performance measurement. Users can often define the time range for which KPIs are calculated (e.g. the number of customer returns in relation to the number of sold products within the last week or the last year). Process mining techniques such as process discovery or conformance checking are not supported by most PAISs, but as already mentioned, tools specifically designed for this purposes exist, such as the ProM framework.

2.3.2. Challenges

One of the grand challenges in process mining is to provide understandable representations of the analysis results (Aalst et al., 2012). The reason is that discovered process models and process execution data might hold a high complexity (Aalst, 2011), hampering understanding and interpretation by non-experts. The Process Mining Manifesto (Aalst et al., 2012) advocates visual analytics, visualization, and interactive process mining for harnessing the *"the amazing capabilities of humans to see patterns in unstructured data"*. Visualizations along different process perspectives have been developed and several interaction techniques such as *Abstract/Elaborate* or *Explore* are offered by process mining tools such as ProM (Kriglstein, Pohl, et al., 2016). However, visualization of process execution data faces certain challenges, such as the number of data dimensions that can be conveyed visually, or the ability to convey patterns or details on individual events in large process execution logs.

Table 2.1.: Challenges of business process execution data and affected tasks.

Challenge	Monitoring	Ex-post analysis
Potential to miss time-critical events	x	
Difficulty to concentrate on other work	x	(x)
High amount and complexity of data but limited screen size	(x)	x
Perception of patterns or details in large logs		x

2.4. Conclusion

As business process execution data differs structurally from that of many other domains (for example, by combining time-based, quantitative and qualitative data), real-time monitoring and ex-post analysis of process executions face certain challenges. Some of those challenges are tackled by supporting users with methods from machine learning, visualization, and visual analytics, but some challenges still remain. Table 2.1 summarizes the main challenges that cannot be alleviated completely by state-of-the-art visualization and visual analytics.

3. Visualization and Sonification for Process Monitoring and Data Analysis - Background and Related Work

This chapter focuses on the benefits and challenges of visualization and sonification for to the representation of data in general, and of process execution data in particular. Furthermore, it addresses the question of which use cases the different modalities are suited to, with regard to data structure and tasks, as well as how they can be best combined into systems for multi-modal data analysis. It starts by giving a general introduction to visualization and sonification, and evaluating their benefits and challenges in a domain-independent way, before analyzing the body of related work that exists for the two modalities concerning the domain at hand. It describes and discusses, therefore, both background information concerning the two modalities, and the body of related work concerning their application for the monitoring and analysis of process execution data.

However, as visualization in general has a longer tradition and the suitability for different types of data and tasks is better researched than concerning sonification, more emphasis is put on the body of sonification research, compared to visualization. One reason for this is that unlike visualization, the comparably young discipline of sonification has few generalized guidelines and best practices to date (Kramer et al., 1999). However, as there exist hardly any research approaches specifically concerning sonification for monitoring and analyzing business process data yet, the related work section is extended to cover the research concerning sonification in monitoring and analysis of data from other domains in which the data is of similar structure (i.e., in principle based on discrete events).

This chapter contains ideas, results, and text from my previous publications. The systematic literature review concerning sonification in computer security contains text and results that were published in (Hildebrandt and Rinderle-Ma, 2015). The analysis of strategies for sonification mappings in process monitoring contains text and results that were published in (Hildebrandt and Rinderle-Ma, 2013).

3.1. Data Visualization

Data visualization is a technique that is widely accepted in society. Graphs of economical statistics are frequently printed in newspapers, and their interpretation is common knowledge. Many web pages feature interactive information visualizations. Due to the fact that several different visualization techniques exist that are established and can be interpreted and understood by many people, almost all software applications for data analysis feature visualization in some form, from software primarily aimed at business users, over tools to analyze scientific data, to software that is often used by end users as well (like Microsoft Excel or LibreOffice Calc).

There are many ways to visualize data, but some are broadly accepted across disciplines:

- Scatter-plot diagrams: A scatter plot is a data visualization where dots are positioned on a two-dimensional chart based on two or more attributes of a given data set. Even though scatter plots are typically two-dimensional, three-dimensional scatter plots exist as well.
- Bar charts: Bar charts are typically two-dimensional diagrams with rectangular bars, where the length of the bars is mapped in proportion to the data.
- Time-series graphs: Time series graphs are typically two-dimensional diagrams, where one line displays the development of one data set over time. The time is typically being mapped to the horizontal axis while the data property is being mapped to the vertical axis.
- Pie charts: Data is mapped to the size of a „slice" of a circular graph. Pie charts are used to illustrate proportions.
- Graph diagrams: A graph diagram is a usually two-dimensional representation of a mathematical graph (a set objects in which several pairs are connected by links). Therefore, graph diagrams are typically used to display relations between objects.
- Map-based visualizations: Map-based visualizations are visualizations that combine geographical information (such as a world map or a country map) with other data (often economical, political, sociological or medical). Publicly available map tools (such as Google Maps) are often used to create map-based visualizations.

Visualizations can be interactive or non-interactive. Non-interactive visualizations, once created, cannot be changed or adjusted in any way (like a graph in a newspaper or an image file). Visualizations might be animated (for example in the form a video file), but are still considered to be non-interactive if they do not react to user action. A visualization that builds on parameters and settings adjustable beforehand by the user offers a higher degree of interactivity than the types previously described. One could, as an example, think of a scatter-plot visualization for which the user can select which parameters are mapped to the horizontal and vertical axes.

In general, the most frequently used visual dimensions on which data are mapped are the following:

- Position (2D/3D)
- Color
- Shape
- Size.

Additional information is often displayed as text or numbers.

3.1.1. Benefits and Challenges

Due to its nature, visualization offers several benefits over other types of presenting data, like textual information or sonification (Frauenberger, 2006):

- Easy integration with text/numbers
- Suitability for comparison/reference with different information
- Good for fast overview at a glance
- Very suitable for spatial information (like geodata, process models)
- Possibility to take a discrete *snapshot*, for example of an animation
- Knowledge to read and interpret already widespread; western societies more visually than aurally oriented (Biesheuvel, 1949).
- Creation and consumption without technical means possible (e.g. using pen and paper or printout)
- Usually easier to focus on single aspect in visualization compared to sonification - suitable for cases in which singular data have more importance than their relations and general structure (Frauenberger, 2006)
- Possibility to ignore (look away, close eyes)
- Visualizations often more suitable for precise representations of data than sonifications as, in some areas, sound has lower resolution.

Of course, visualization faces also certain challenges. In general, as with other modalities, the benefits and challenges depend heavily on the context, like the amount and type of data to be presented, the working condition and environment of the user, and his/her preferences and training:

- Limited number of properties onto which data can be mapped
- Difficult to process multiple different data streams in parallel
- Limited number of space, particularly on mobile displays
- Hard to perceive in background to other task
- Certain head orientation and eye focus necessary, therefore not suitable for emergency situations
- Difficult to detect small variations in development over time
- Not suitable for blind/visually impaired

- Often more difficult to remember than highly salient musical patterns (Kramer et al., 1999).

3.1.2. Visualization for Business Process Monitoring

Visualization of process data often refers to the visualization of process models (Grossmann and Rinderle-Ma, 2015). There is e.g. a large body of work that deals with questions of layout and presentation of process models. However, such approaches typically aim to support users in the process design phase (which is not a focus of this thesis), but usually not during real-time monitoring and post-hoc analysis. Exceptions are works like (Bobrik and Bauer, 2007) or (Rinderle et al., 2006), that allow for process model visualizations that can be updated in real-time to reflect the status of process executions.

In contrast, only few approaches have been proposed for visualizing process event data. In general, although the inclusion of visualization techniques is very common in commercial and open-source solutions for monitoring and analyzing process execution data, e.g. in the form of dashboards, there seems to exist no significant body of research in this regard. Existing research in the area of e.g. business process intelligence focuses mainly on aspects of gathering, integrating and processing data, and less on the presentation of such data to the user (e.g. Sayal et al., 2002). Most of the few existing approaches that do investigate such aspects seem to be based on completed process instances, and thus on supporting users in process analysis rather than in real-time monitoring.

3.1.3. Visualization for Analyzing Business Process Data

A few research projects try to tackle the challenges of visualizing historical business process execution data. Schönhage, Ballegooij, and Elliëns (2000) propose the usage of three-dimensional visualizations. The authors created a three-dimensional visualization of business process simulation data. While the three-dimensional version was not as accessible as a two-dimensional version, the authors tackled the challenge of the limited screen space, as the amount of data that can be simultaneously displayed increases with an additional dimension. Hao et al. (2006) apply visualization methods for the analysis of business process models as well as instance execution data. These methods aim to identify factors that influence business metrics and resource parameters. To do so, the authors first apply data mining techniques to identify important factors and then use interactive three-dimensional circular graphs to visualize them. Suntinger et al. (2008) apply visualization methods for the historical analysis of business process event data. They do so by using a 3D-tunnel metaphor, where newer events are visually represented in the front and older ones in the back, and related events are grouped together. Thus, the event tunnel enables users to identify irregularities and patterns during process analysis. In (Kriglstein, Wallner, and Rinderle-Ma, 2013), the visualization of instance traffic refers to an aggregation of the number of executed process instances along process paths. In the ProM framework, dotted charts are the main visualization for the occurrence of process event data over time (Aalst, 2011). Kriglstein, Pohl, et al. (2016) analyze the different visualization techniques that exist within the ProM framework. The authors show that different data types and interaction techniques are supported.

3.1.4. Research Gap

Although visualization is included in the majority of commercial and non-commercial applications for monitoring and analyzing business process execution data, there seems to be hardly any research concerning how such visualizations need to be designed to best support users in their tasks. What is needed therefore is research that investigates such aspects with potential users, especially for the challenging task of real-time monitoring.

3.2. Sonification and Multi-Modal Displays

Sonification can be defined as the "presentation of data using sound" (Hermann, 2002). This presentation is usually intended to support the listener or user to gain new insights into the presented data. Beyond the usage of sonification to enhance visual means (multi-modal sonification), it is also applied as an alternative to visualization in situations where visual focus and attention are needed elsewhere (e.g., in cockpits or operating rooms) or to support blind or visually impaired people. This section begins with a brief, general overview of sonification in research and practice. Subsequently, a short history of sonification is given, followed by an overview of how data can be mapped to sound, and which established sonification techniques already exist. This section ends with an analysis of the benefits and challenges of sonification.

There is a growing amount of sonification research concerning applications, methods, and perception. There exists research that applies sonification in many different areas, such as astronomy, volcano activity, ice glaciers, RNA structures, brain activities and weather data. Furthermore, there are also a few examples in the fields of social sciences, such as sonifications for population developments and election outcomes, sport sciences and economics (e.g., sonification of stock market data (Ciardi, 2004)).

For instances in which sonification is already being applied in a real-world context, the applied sonification techniques are typically very basic, and are applied in areas where a constant visual focus cannot be guaranteed, such as airplane cockpits, medical operation theaters, cars (e.g. parking aids), or navigational aids for blind or visually impaired people. Those sonifications usually convey the occurrence of one of several pre-defined events (e.g. alarms), or map one parameter to sound (like the parking aid or the Geiger counter). Sonification for data analysis or exploration, for example as a part of an analysis software, is generally still a research topic, although, to our best knowledge, there are no studies concerning the adoption of sonification techniques and software in science or business. However, the usage of sonification to present and analyze scientific data seems to be increasing over time, as, for example, various sonifications of experiment data at the CERN (European Organization for Nuclear Research) demonstrate (e.g.[1]). There are a few examples of sonifications used in the area of data journalism that convey data to the general reader.

3.2.1. History of Sonification

One of the oldest examples of a sonification is probably the Morse code (used to transfer letters by means of a sound code), invented around 1837. Another early sonification is the Geiger counter (used to sonify rates of radiation), invented in the early 20th century. Also from the early days of sonification comes the pulse oximeter, which is used to sonify the fraction of oxygenized hemoglobin in blood during medical surgeries. Apart from these sonification examples, sirens have been used for a long time to draw attention to certain events. According to Hermann (2002) the two pioneers of research in data sonification were Pollack and Ficks (1954), who investigated the usage of abstract auditory variables for the presentation of quantitative information. Speth, in 1961, proposed the usage of audification of seismic data (Speeth, 1961). Neurophysiologists have been listening to neurons firing for at least 40 years. Chambers, Mathews, and Moore (1974) researched in the field of three-dimensional auditory displays as enhancements of scatter plots. Bly (1982) developed parameter mapping sonifications based on non-ordered, multi-dimensional data sets. An important milestone for sonification was the inaugural International Conference on Auditory Display (ICAD) in 1992. Since then, a growing amount of research concerning perception, the development of sonification models and the creation of methods has accumulated in various fields of application. Apart from selected areas such as medicine, sonification is still mainly a research subject; however, there is a growing amount of sonification used as public outreach of scientific results, or to present data to general users. As an example, sonification techniques are used on the web presences of renowned newspapers[2] and are regularly broadcasted on the radio[3].

[1] https://lhcsound.wordpress.com/
[2] http://www.nytimes.com/interactive/2010/02/26/sports/olympics/20100226-olysymphony.html/
[3] http://tweetscapes.de/about/

3.2.2. Mapping of Data to Sound

As with visuals, audio data can be mapped onto a number of dimensions, some of which are more suitable than others, depending on the use case. Acoustic properties that suggest themselves for sonification in particular (Hermann, Hunt, and Neuhoff, 2011, chapters 2 and 4) are frequency and spectrum; tempo and gain are usually less suitable.

However, one should consider that some dimensions are not independent and therefore interact with one another. For example, mapping data to frequency and gain in parallel may influence the perception of pitch. Furthermore, the different dimensions can be grouped into different layers (Hermann, Hunt, and Neuhoff, 2011, chapter 15).

The basic auditory properties that data can be mapped onto are the following:

- Frequency
- Spectrum (distribution of the different frequencies that are present in certain sound)
- Tempo
- Gain (how much a sound is amplified, which influences loudness of sound)
- Duration
- Envelope (the way in which the level of a sound wave changes over time)

These properties influence the perceptual grouping on the sound object layer:

- Pitch (subjective perception of frequency)
- Loudness (subjective; influenced by gain but also many other factors)
- Timbre (in music perceived as instrument, mainly influenced by spectrum and envelope. In non-musical sounds perceived as brightness and roughness)

The following factors should be considered when designing a sonification, alongside the grouping on the auditory scene level:

- Masking (certain sounds/sound layers are masked by others)
- Location (sounds can appear to come from a certain direction, for example left/right (panning), or from any direction assuming a corresponding speaker/headphone setup).
- Fusion (stream fusion: if two or more sounds are played at the same time, depending on certain attributes in the perception of the listener they can be perceived as belonging to the same coherent sound stream)
- Segregation (the opposite of fusion, i.e. although several sound streams are played at the same time, they can keep individual "identities" that the listener perceives as separate streams)
- Harmony (depending on different attributes and subjective perception of listener, a sound or parallel combination of sounds (like a musical chord) can be perceived as harmonious)

- Rhythm (pattern of sounds, meaning several sounds played in a row; perception very subjective)

For a detailed overview on how these different properties are grouped and influence each other, please refer to (Hermann, Hunt, and Neuhoff, 2011). Aesthetically, auditory displays can range from being very *musical*, for example by conveying data by notes on a classical instrument such as a piano, to consisting of abstract sounds (such as basic wave forms) or pre-recorded audio samples (Walker and Nees, 2011).

3.2.3. Sonification Techniques

The four most common established sonification techniques (audification, auditory icons, earcons and parameter mapping) are presented in this subsection.

Audification

An audification is the direct conversion of data points into samples. One sample is the smallest discrete unit of digital sound. This means that for a typical sample rate of 40kHz, 40,000 samples are needed for one second of sonification. In order to be audible, the sounds should fall within a frequency range from around 50 hertz to around 20,000 hertz. Therefore, for one second of audification, around 40,000 values between 50 and 20,000 are needed, thus limiting the field of possible modes of operation. The problem of the audible frequency range can, however, be solved by scaling the existing data (Hermann, 2002).

Auditory Icons

Auditory icons are everyday sounds that directly represent the events that are being sonified (Hermann, 2002). One example is the sound of a paper basket being emptied that is played back upon emptying the metaphorical paper basket in the Windows operating system.

Earcons

Earcons are non-verbal audio messages consisting of motives, which are short rhythmic sequences of pitched tones with variable timbre, pitch and amplitude. Timbre describes the *basic properties* of sounds and is a subjective characteristic that enables the differentiation of two sounds, even though they might have the same loudness and pitch (Hermann, 2002). The concept of earcons is similar to that of auditory icons with the difference that auditory icons are everyday sounds that directly represent the event that is being sonified, whereas earcons can be *abstract* symbols that are not similar to the real-world sound of the represented event or object.

Parameter Mapping

Parameter mapping is the mapping (either direct or by scaling) of data values to specific attributes of sound. These attributes typically include volume, pitch, panning (the position of a sound in the stereo field) or timbre. Other possibilities to map parameters of sound are repetitions and pauses between distinct sound events in loops. Other approaches include filtering out specific ranges of frequencies according to data. Due to these characteristics parameter mapping is often being said to be the sonic pendant to a scatter plot diagram (Hermann, 2002).

Summary

Table 3.1 summarizes the strengths and weaknesses of the presented sonification methods.

Table 3.1.: Strengths and weaknesses of different sonification methods.

Representation	Advantages	Disadvantages
Audification	able to compress large amounts of data into short sonifications	restrictions on the data to be sonified, limited means of sound design
Auditory Icons	very intuitive and easy to learn	limited means of sound design, not able to convey quantitative data
Earcons	extensive means of sound design	not able to convey quantitative data, often not intuitive and easy to learn
Parameter Mappings	extensive means of sound design, ability to convey quantitative data	often not intuitive and easy to learn

3.2.4. Benefits and Challenges

Auditory perception has several properties that seem to make sonification ideal as a supplement to visualization in many cases and for many data types, at least as an optional addition. However, there are certain data types and use cases for which sonification is particularly suitable as an enhancement, or even replacement, for visualization. Time-based data (like process execution data) suggests itself for sonification in particular, as do most use cases where data has to be monitored in real-time.

With state-of-the-art monitoring solutions based on visualizations, on the other hand, users currently have to switch their visual focus (and thus their attention) between the process monitoring application and the main task. Particularly in peripheral or serendipitous-peripheral monitoring situations, where the attention is focused on a primary task and other information is monitored indirectly at the same time, visual means are not well suited, as pointed out by Vickers (2011, ch. 18, p. 455). Sonifications, on the other hand, might better be able to keep users informed about background activities without being disruptive and in peripheral monitoring, auditory cues might be more useful than visuals and less intrusive. Not only is sonification useful as sensory substitution of visual displays, but also as a supplement during direct monitoring tasks, for instance when the visual displays are very complex or if many parameters need to be conveyed simultaneously. For the use case of process monitoring in particular, sonification offers the following benefits:

No visual attention required: Listening is not limited to a specific head orientation, therefore we cannot miss information only because the screen is out of view. Thus, monitoring personnel can move freely in the room, and can also focus their visual attention on other tasks while they are listening to a sonification.

Rapid detection & processing: Sound is processed faster than visual signals, thereby reducing reaction times. This is crucial when time-critical notifications during process execution need to be conveyed.

Habituation and sensitivity to changes: We can notice even small changes in rhythms and sequences of sounds over time. This effect can be leveraged by sonification, for example by enabling users to hear if the activities of a process have been executed in a different sequence than usual, or to hear if there has been an especially long break between two subsequent activities. As our auditory perception is able to habituate to regular soundscapes (acoustic environments), such a sonification can remain unobtrusive during 'normal' operation, while even small changes in sound over time are able to immediately grab our attention in case of process deviations or undesired process behavior.

Guiding Attention: Sounds attract our attention and let ourselves search for the cause – our ability to estimate the sound source location can thus be used to reorient us to relevant situations.

To state an example concerning our ability to perform auditory process monitoring, machine maintenance experts in factories and plants have been listening to the acoustic patterns machines produce for decades. They are often able to evaluate if a machine is about to break down, or a specific part needs to be replaced soon, by listening to the frequencies and patterns of the sounds a machine produces, a technique referred to as vibration analysis (Renwick and Babson, 1985). Crucial vibration properties are amplitude, frequency, phase and modulation (Renwick and Babson, 1985). However, by using sonification methods, these inherent properties of audio can be leveraged and made accessible to a wider range of people. On the one hand, sonification can decrease the need for experience that is necessary for vibration analysis significantly, as the data can be aggregated and filtered according to the individual users information needs, while at the same time optimizing the resulting audio for our cognitive abilities. On the other hand, for most processes other than in a production scenario, vibration analysis has never been an available means for monitoring.

Challenges

In general, sound is primarily a temporal medium, in that it can only exist over time. In contrast to animation, one cannot freeze a discrete state of sound in time, meaning that sound often shows most potential for time-based data (like process execution data). Furthermore, sonification is usually not as suitable as visualization for conveying concrete text or numbers. An activity name or an exact number, for example, can only be conveyed aurally using speech, which can be very distracting. Therefore, sonification is often preferable for conveying trends and developments, while a corresponding visual display can then be used to convey detailed, concrete information (such as exact numbers), if necessary. As pointed out by Walker and Nees (2011), complex auditory displays are relatively new and therefore, unlike for visualizations, the skills to interpret such are not widespread in our society yet. Thus, potential users of auditory displays generally require more training than those of visual displays. This challenge can be alleviated by designing sonifications to be as intuitive as possible, for example by using fitting auditory icons to convey the occurrence of certain events or by applying intuitive mapping analogies. In order to do so, specific attention has to be paid to what data dimension is mapped to which acoustic property (e.g pitch or tempo) and how (e.g. linear or exponential). Furthermore, there are certain challenges that specifically have to be overcome when applying sonification for real-time monitoring, as defined by Vickers (2011):

- Potential intrusion and distraction
- Fatigue and annoyance
- Aesthetic issues
- Comprehensibility and audibility

The first two challenges are based on the fact that users of potential auditory process monitoring systems may have to listen to them for several hours a day. If no headphones are used, sonifications can disturb coworkers in shared work spaces (unlike visualizations). Furthermore, visualization is in general more adequate in conveying spatial information (e.g. geographic data or static process models), as well as data that is not time-based (such as detailed information on specific aspects). In general, while sonifications usually require technical devices for creation - you cannot *draw* a sonification on a piece of paper like you can with a visualization - sonification always requires technical means to play back, and, for example, cannot be printed on a piece of paper.

3.3. (Multi-Modal) Sonification for Process Monitoring and Analysis

This section presents the related work in terms of sonification for process monitoring and -analysis. A substantial amount of research concerning applications for monitoring different types of processes using sonification has accumulated, spanning various areas such as industrial production processes, program execution and web server behavior. Rauterberg and Styger (1994) developed sonifications for the monitoring of industrial production environments. Tran and Mynatt (2000) researched into sonifications to monitor home environments. Another popular application area of auditory process monitoring is computer program execution. As early as 1949, a computer had its circuits wired to an audio channel in order to support audio-enabled debugging (Knuth, 2011). There were several other research projects that investigated how to use sonification for program debugging purposes, such as using music in HCI, as proposed by Alty (1995) or Vickers and Alty (2005). Furthermore, examples for sonifications for monitoring web servers and computer networks exist, such as *Peep* (Gilfix and Couch, 2000) or *WebMelody* (Ballora, Panulla, et al., 2010). A survey of existing approaches to the usage of sonification for the task of process monitoring summarizes these different areas (Vickers, 2011).

There has, however, been little research to date that specifically addresses sonification for business process monitoring. In the ARKOLA simulation, Gaver, Smith, and O'Shea (1991) describe a real-time multi-modal sonification of a bottling plant. In this simulation, users manually control the settings and adjustments of several interconnected machines, trying to avoid stops and bottlenecks. Events such as liquid spills are communicated to the user by appropriate sounds as they occur. After analyzing the findings of user studies, the authors came to the conclusion that the audio feedback played an essential role in the participants' interaction with the system. Hermann, Niehus, and Ritter (2003) investigate sonification for process monitoring; however, their research concerns the field of robotics. This leads to the assumption that there is still a substantial amount of untapped research potential in this area.

3.3.1. Auditory Mechanisms for Attention Allocation

A very important aspect in developing solutions for process monitoring is attention allocation, as solutions should be able to attract users' attention when necessary, while avoiding mental overload. Sonifications in particular should ideally be unobtrusive. However, generally, the more information a sonification conveys, the greater the risk of being disturbing. This trade-off has been researched by Gaver, Smith, and O'Shea (1991), among others. There is a wide selection of research that investigates how sonifications can guide users' attention, such as (Seagull, Wickens, and Loeb, 2001).

Attention allocation is important in every monitoring application, but especially so in peripheral monitoring, where the user is primary engaged in another task, while performing monitoring as a second task. McClimens and Brock (2010) investigated the effectiveness of auditory displays to improve dual-task performance. Bakker, Hoven, and Eggen (2010) designed several sonifications that were intended to run during a whole workday in the background. The authors concluded that users found a sonification using recordings of rain to be useful yet unobtrusive. The authors further found out that after a few weeks, the sounds shifted away from the center of attention, grabbing the users' attention only during unexpected occurrences. The same effect was discovered by Kilander and Loennqvist (2002), who suggest that natural sounds are better accepted as part of the environment, especially if they constitute a *constant murmur* instead of a stream of individual sounds. Caldwell and Viraldo (2014) suggest to investigate the usage of complex sonification to convey state-based information in control-room scenarios in order to tackle the problem of information overload and alarm flooding. Cohen (1994) also researched in the area of unobtrusive ambient information systems, for example by using steady sounds, with sharper sounds used to attract the users' attention.

Studies testing effectiveness have only been conducted on a few of the existing approaches. To state an example of such studies, Watson, Sanderson, et al. (2003) conducted research on the general effectiveness and utility of audio in process monitoring by measuring participants' performances on different distractor tasks (e.g. solving simple arithmetic problems) and patient monitoring in three conditions (within-subjects-design): visualization, sonification and multi-modal. The sonification and multi-modal modes were based on a continuous sonification in which several parameters were mapped onto two sound streams based on pure tones. The highest results of the arithmetic task were observed, when the peripheral patient monitoring was conveyed using the auditory-only condition, and the lowest results in the visual-only condition. The participants in most cases achieved better patient monitoring results during the multi-modal condition. The performances under the conditions that included sounds were even better, when the respective part of the experiment was not the subjects' first part of the experiment, but the second or third.

Overall, it seems that studies comparing a visual-only condition to a multi-modal condition with sporadic auditory alerts or alarms conclude that auditory signals do not significantly affect the performances in both tasks (McClimens, Brock, et al., 2011; Brock, McClimens, and McCurry, 2010). In a few cases, both tasks are negatively affected (e.g. McClimens, Stroup, et al., 2004). Typically there are fewer head movements and attention switches measured in the multi-modal condition (Brock, Ballas, and Stroup, 2002), something that has also been observed for continuous sonifications (Sanderson et al., 2004). When comparing performances in the main task – often simulated with arithmetic problems – in experiments that include continuous sonifications, the results are mixed: in some experiments, fewer mistakes were made under the multi-modal condition compared to the visual-only condition (Watson, Sanderson, et al., 2003; Poguntke and Ellis, 2008), while in other experiments, the best main task performance was observed under the visual condition (Crawford et al., 2002; Watson and Sanderson, 2004).

In tendency, the main task performance seems to be slightly negatively affected by sound (more so in multi-modal conditions than in auditory-only conditions), although these differences between the conditions are not statistically significant in the majority of studies. Even so, results under conditions that include sound in particular are typically better when the respective condition is not the subject's first, but second or third, condition of the experiment (e.g. Watson, Sanderson, et al., 2003 or Poguntke and Ellis, 2008). Thus, the observed distraction by sound seems to be smaller when the participants are already used to and familiar with the two tasks themselves. Performance in the secondary task (monitoring) is typically significantly higher in multi-modal conditions that feature continuous sonifications, compared to visual-only conditions (Watson, Sanderson, et al., 2003; Crawford et al., 2002), although a few studies report the opposite result (e.g. Sanderson et al., 2004). As with the main task, there seems to be a strong familiarization effect that particularly benefits multi-modal conditions (Watson, Sanderson, et al., 2003), which may be an explanation for why the advantage of auditory conditions over the visual-only condition seems to be greater for domain experts than for amateurs (Crawford et al., 2002).

Participants, especially domain experts, when asked for their opinions, generally state a preference for multi-modal conditions including continuous sonifications, attributing the preference to a feeling of being in control, among other reasons (Poguntke and Ellis, 2008; Crawford et al., 2002).

3.3.2. Sonification in Computer and Network Security Monitoring

As computer and network security monitoring is one of the most prominent domains in sonification research, while sharing certain commonalities with business process monitoring, a literature survey and analysis of this area was conducted.

While most approaches in this area aim to support operators during network surveillance in general, a few systems are designed specifically for web server monitoring. The developed significations are based on different types of data, including unfiltered event data (such as individual web server requests), filtered event data and aggregated state-based data (such as e.g. the traffic load of a server) and alert notifications. Of course, the existing research is in different stages of maturity, ranging from vague system proposals, over prototypes that can be used to conduct user studies, to systems that are already publicly available for download and in use.

Method

The following literature survey is based on a Google Scholar search performed on 4 March 2015. The keywords "sonification" and "security" were used in all searches, and were combined alternately with one of the keywords "monitoring" and "analysis". Papers that did not fit the topic were excluded. Examples are papers in which the term "sonification" is only mentioned in the related works section and that do not feature an approach or system that actually incorporates sonification techniques, or papers that only mention "security" as one of several application domains in which a generic framework or sonification technique can be applied, without specifying a computer security-related use case. In cases where several papers by the same authors describe the same approach (e.g. in more detail), only the most recent paper has been included.

Findings

Table 3.2 summarizes the different existing approaches. The field "Area" states the intended area of use, such as network monitoring. "Granularity" describes which kind of data is sonified (e.g. alerts) and "Modality" categorizes the publications into approaches that are "standalone" sonifications and such that are combined with visualization into multi-modal systems. "Status" describes the maturity of the respective research approach, like "concept". Finally, the field "Study" states if a user study has been conducted or not.

Ballora et al. (Ballora, Panulla, et al., 2010; Ballora and Hall, 2010) describe systems that sonify entries of web server log files, stating that their systems are designed to work in a real-time mode as well as for post-hoc data analysis. In a later publication (Ballora, Giacobe, and Hall, 2011) the authors suggest a new system based not only on raw events but also on aggregated state-based data (e.g. current throughput/traffic rates), reiterating its suitability for real-time monitoring as well as for retrospective analysis. Of course, many of the other systems technically support post-hoc data analysis as well, as they often use log files as a basis. However, as they are not designed specifically with the use case of retrospective analysis in mind, there is a certain likelihood that they are not very suitable for this use case.

In terms of the application area, 17 of the 20 presented papers deal with network monitoring in general, typically with the aim of supporting users in detecting intrusions. Three approaches more specifically aim at the goal of web server monitoring (Ballora, Panulla, et al., 2010; Ballora and Hall, 2010; Barra et al., 2002). The first two of those approaches (Ballora, Panulla, et al., 2010; Ballora and Hall, 2010) are the only ones that solely present unfiltered event-data.

Most other approaches are based on either filtered event data, for example by presenting only events that fit pre-defined criteria or auditory alerts, or on higher-level quantitative, state-based data. Vickers, Laing, and Fairfax (2014) describe

Table 3.2.: Auditory and multi-modal approaches in computer security.

Publication	Area	Granularity	Modality	Status	Study
Fairfax, Laing, and Vickers, 2015	Network	State-based data	son	con	No
Vickers, Laing, and Fairfax, 2014	Network	State-based data	son	prot	No
deButts, 2014	Network	Filtered events	multi	prot	No
Giot and Courbe, 2012	Network	State-based data	son	con	No
Ballora, Giacobe, and Hall, 2011	Network	Raw events, State-based data	son	prot	No
Ballora and Hall, 2010	Web server	Raw events	multi	prot	No
Ballora, Panulla, et al., 2010	Web server	Raw events	son	prot	No
Muñoz-Arteaga et al., 2009	Network	Alerts	multi	con	No
Brown et al., 2009	Network	State-based data	son	prot	No
Qi et al., 2007	Network	State-based data	son	prot	No
Miguel A Garcia-Ruiz, 2006	Network	Not specified	multi	con	No
Varner and Knight, 2004	Network	State-based data	multi	con	No
Gopinath, 2004	Network	Filtered events	son	prot	Yes
Papadopoulos et al., 2004	Network	State-based data	multi	prot	No
Malandrino et al., 2003	Network	Filtered events, State-based data	son	prot	No
Varner and Knight, 2002	Network	Not specified	multi	con	No
Kimoto and Ohno, 2002	Network	Filtered events	son	prot	(Yes)
Barra et al., 2002	Web server	Filtered events, State-based data	son	Tool	No
Gilfix and Couch, 2000	Network	Raw events, State-based data	son	Tool	No

Legend: son=sonification, multi=multi-modal system, prot=prototype, con=concept

a system that sonifies the SOC (Self-organized Criticality) of a network. Brown et al. (2009) sonify the bit-rates and packet-rates of a delay queue, an approach shared with Qi et al. (2007). Ballora et al. sonify, in addition to raw events, the amount of network traffic (Ballora, Giacobe, and Hall, 2011). The *Peep*-system sonifies state-based data, such as the average server load or the number of users on a specific machine, as well as event data (Gilfix and Couch, 2000). Giot and Courbe (2012) propose to sonify various network-state information, such as packet sizes or port usage statistics. Papadopoulos et al. (2004) propose a system that sonifies quantitative network traffic data, without mentioning the specifics of this proposed software.

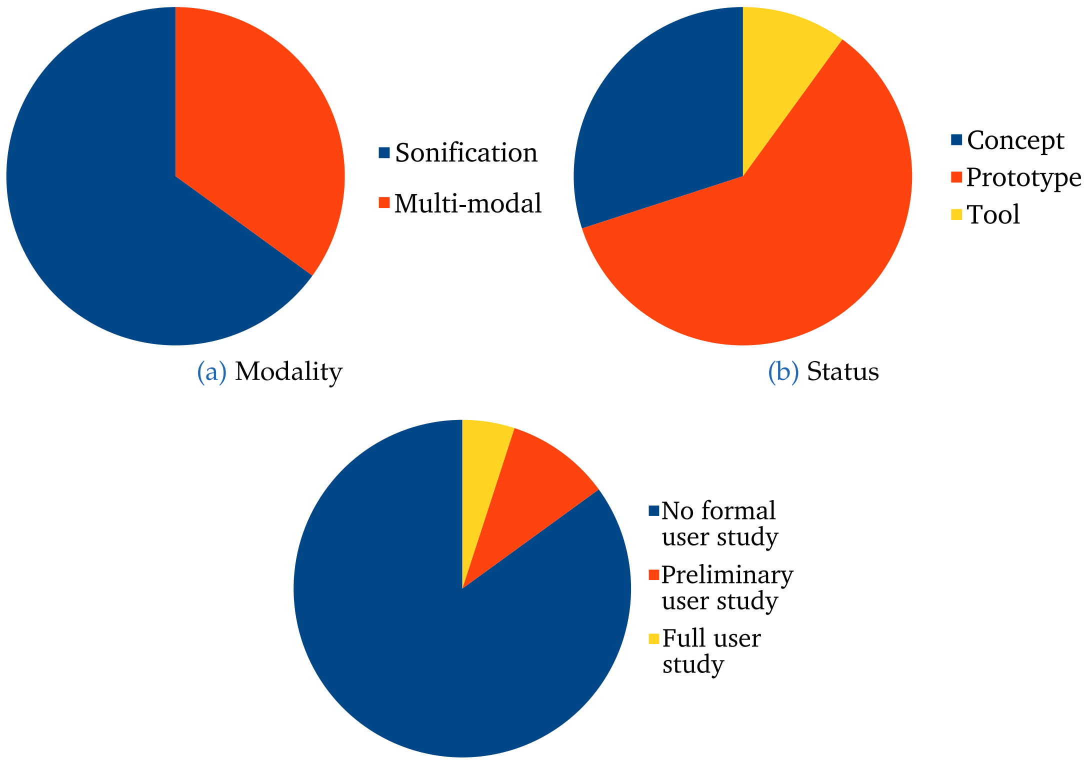
In terms of modality, 6 of the 20 publications feature multi-modal solutions that combine visual and auditory means. Looking at the integration between acoustic and visual conveyance, the existing literature covers the whole range of possibilities: Muñoz-Arteaga et al. (2009) suggest applying audio purely as a means to convey intrusions that have already been discovered by intrusion detection software, while visual means are then used to find out additional information.

The opposite of this interplay of modalities is proposed by Varner and Knight (2004). In this proposal, visualization is used to convey the status of network nodes, while the user can then select a specific node to hear a sonification that conveys additional details. A more integrated approach is described by Ballora and Hall (2010), where network information is simultaneously conveyed aurally using a multi-channel sonification and visually using a 3D visualization. Papadopoulos et al. (2004) propose a system that combines 3D audio and 3D visuals in parallel to convey different network-related data. Miguel A Garcia-Ruiz (2006) propose a system that combines 3D audio in conjunction with visualization to convey network attacks. The details of this proposed system are not specified, however. The same holds true for a position paper (Varner and Knight, 2002) that proposes the use of visual and auditory means without specifying details. Fig. 3.1a summarizes this aspect.

In terms of maturity, several of the presented research papers constitute research agendas or proposals for systems and prototypes that still need to be developed. Most of the publications, however, describe already-developed prototypes in several stages of maturity, while only two publications seem to focus on ready-to-use tools that are publicly available (Barra et al., 2002; Gilfix and Couch, 2000), as can be seen in Fig. 3.1b.

One striking characteristic is that even the mature systems have undergone little or no evaluation for usability (see Fig. 3.1c). In several papers, case studies describe functionality tests in a certain context; however, mostly without involving potential users. Most papers mention no kind of testing at all, while a few publications include informal preliminary testing. There seems to be only one publication in which a formal user evaluation is described (Gopinath, 2004).

Kimoto and Ohno (2002) mention a first tentative user study with 4 users. In this study, users were played a sequence of sounds indicating traffic loads at different points in time, which users more or less successfully estimated based on the sonifications. However, the number of subjects is too low to provide generalizable results. The mentioned user study did not include a comparison of the respective



(a) Modality

(b) Status

(c) Evaluation

Figure 3.1.: Summary of literature survey on sonification in computer security.

sonification with the status-quo in the respective area (typically visualization-based systems). Such a comparison has been performed in a study conducted by Gopinath (2004) with 20 subjects. The author conducted different user studies, comparing sonification-based approaches with control groups that performed the same tasks without auditory support. The sonification groups performed the tasks (detection and identification of intrusions) significantly faster than the control groups. However, the control group was instructed to perform the tasks by reading log files without access to visualization-based tools, thus possibly limiting the generalizability of the results.

After conducting the literature search, three new systems have been presented that apply audio for computer security purposes. In (Hauer and Vogt, 2015), a prototype that sonifies log files of streaming servers is presented. In its current state, it is designed to allow for the analysis of historical log data, but will be adapted for real-time monitoring in the future, according to the authors. It is based on a continuous sonification of aggregated server data, such as server load, as well as metadata on the content handled by the server. The authors conducted a small case study with five participants, who were able to perform most of the provided tasks effectively. The NetSon system sonifies and visualizes network traffic in real time, with a focus on larger-scale organizations (Worrall, 2015). The system presents server load and aggregated metadata from random samples that are taken at a pre-defined sampling rate. No formal user studies have been conducted with the system as yet; however, it is being used productively at Fraunhofer IIS, who provide a live web stream of their installation.[4] Furthermore, even though it has not been presented in a scientific publication yet, Microsoft's multi-modal system *Specimen Box* is worth mentioning. It allows for the detection and analysis of botnet activity in real-time and in retrospect.[5]

Conclusion

As stated earlier, the area of computer and network security monitoring shares certain commonalities with business process monitoring, therefore it can be assumed that certain findings and results of this area might be transferable to the domain at hand. Unfortunately, very few of the presented approaches applied formal or informal user testing, which makes it difficult to differentiate designs, strategies, and mappings that work well, from those that do not.

3.3.3. Sonification Mappings for Process Monitoring

In this chapter, all sonification projects that have been presented examplarily by Vickers (2011) are analyzed in terms of how similar the underlying data structure is compared to that of business process execution data. Afterwards, the resulting list of related publications is analyzed in terms of which mappings and techniques were applied, and whether they were successful or not, based on formal or informal evaluations by the papers' authors (if conducted). Based on these results, a list of recommendations and guidelines for sonifications of similar data structure is compiled, the results of which are subsequently transferred to the domain of business process monitoring. The criteria for inclusion in this analysis were firstly, that the data serving as the base for the respective sonifications had to be of primarily qualitative nature and convey occurring events in real time.

Results and Findings

The publications focus on a variety of application domains, such as

- Industrial monitoring (ARKola (Gaver, Smith, and O'Shea, 1991); Sharemon (Cohen, 1992) and (Cohen, 1994))
- Home and shared work environments (ListenIn (Schmandt and Vallejo, 2003); Music Monitor (Tran and Mynatt, 2000); WISP (Kilander and Loennqvist, 2002); RAVE (Gaver, Moran, et al., 1992); Workspace Zero (Eggen et al., 2008))

[4]http://www.iis.fraunhofer.de/en/muv/2015/netson.html
[5] http://o-c-r.org/2014/11/15/specimen_box/

- Systems for external auditory representations of programs (Caitlin (Vickers and Alty, 1996; Vickers and Alty, 1998; Vickers and Alty, 2002); Infosound (Sonnenwald et al., 1990); Logomedia (DiGiano, Baecker, and Owen, 1993); Sonnet (Jameson, 1994); ADSL (Bock, 1994); Listen/LSL & Jlisten (Boardman et al., 1995); Program Slices (Berman and Gallagher, 2006))
- Web server and internet sonification (Peep (Gilfix and Couch, 2000); Web-Melody (Barra et al., 2002))
- Interface tasks (SonicFinder (Gaver, 1989)).

The majority of the publications analyzed describe sonifications that convey occurrences of discrete events, such as the aforementioned ARKola Simulation by Gaver, Smith, and O'Shea (1991) in which occurring events, such as spills of liquid, are being sonified using auditory icons with a predefined sample length. Process execution data, on the other hand, is also based on discrete events, but contains additional activities with certain durations. These are marked by the respective "activity started" and "activity finished" events. There are a few sonifications, like the CAITLIN project by Vickers and Alty (1996), where constructs that have a certain duration (like a for-loop) also convey these durations aurally by using continuous sounds throughout the duration of an ongoing activity. In the Sonicfinder project (Gaver, 1989) even sonic metaphors that can give a clue about the remaining duration of an ongoing activity are applied. The authors applied sounds of a jug being filled with water to sonify an operation system's copying procedure.

In terms of the selected mapping techniques, two almost equally large fractions can be identified: sonifications that apply mappings of events to auditory icons, and sonifications that apply earcons, which in most cases are based on harmonic principles of western music, or musical motifs (short melodic patterns) like (Vickers and Alty, 1996). A few of the analyzed sonifications let the user define the sounds used, as well as the events that trigger their playback. Some approaches (like Vickers and Alty, 1996) aurally convey not just event occurrences, but also structural or hierarchical information, mostly by using hierarchically structured classes of motifs or timbres.

In the following, based on formal and informal experimentation conducted by the authors of the respective publications, sonification techniques and mappings that have successfully been applied to convey certain concepts are presented and summarized. In OutToLunch (Cohen, 1994), a sonification based on motifs was deemed more pleasant than an earlier version with auditory icons, although this assumption has not been formally proven. On the other hand, Berman (2011) concluded (based on user evaluations) that the associations of concrete sounds are easier to recall than those of abstract sounds, and that they therefore should be preferred over musical sounds if they are available, as they entail less cognitive overhead. Vickers and Alty (1998), applying musical earcons, further suggested (based on user testing) not only to use melodic constructs, but to include percussion and rhythm in sonifications as well. Furthermore, the authors state that tonal music should be preferred over direct mappings to frequencies (Vickers and Alty, 2002).

Different approaches have been taken to differentiate aurally between different constructs. Vickers stated that by using different motifs (short melodic patterns), users can often distinguish different constructs. He also stressed the importance of using different timbres/instruments for the different constructs (Vickers and Alty, 1996; Vickers and Alty, 2002). He suggests that motifs should also differ in rhythm and tempo and that they should convey durations of activities by using sounds that are played back for the whole duration of an activity (e.g. by using drones - continuous humming sounds; (Vickers and Alty, 1996)). This concept has also been applied by Francioni, Albright, and Jackson (1991). Vickers and Alty (1996) furthermore suggest specifically investigating the usage of musical contour (the direction and shape musical notes move in, like a short melody that is rising in pitch) as it might yield better results than motifs that do not take such considerations into account (Vickers and Alty, 2002). Francioni, Albright, and Jackson (1991), on the other hand, distinguished different concepts (in this case different processors) by assigning each concept a different timbre. Thus, while the different events that occurred during execution were each assigned a different note, those notes were played in the respective instrument for the processor in which the events occurred.

Another concept that has been conveyed aurally (e.g. by Vickers and Alty (1996)) is hierarchical information. Vickers and Alty mapped hierarchical information to hierarchical groups of leitmotifs and derived motifs, stating that most users were able to distinguish between the different top-level constructs sonified using leitmotifs, while they had more problems in distinguishing the derived motifs assigned to the respective sub constructs. Further, they concluded that in general, the developed sonification for debugging purposes proved beneficial for very complex programs in particular, while it was less helpful for debugging simpler constructs.

Vickers and Alty also suggested allowing users to not only decide which instruments the different constructs are being mapped to, but to let users influence the melody creation as well. (Vickers and Alty, 2002; Vickers and Alty, 2005)

3.3.4. Sonification for Process Analysis

Due to several factors, many of the earlier presented results for sonification in process monitoring are not directly transferable to process analysis: process monitoring is often performed as a background task in parallel to other work, and executed for a long period of time (often over a complete work day). Process analysis, on the other hand, is typically a task performed for a confined period of time, but with the users' full attention. This has implications both for sonification aesthetics, and the amount of data that can be (and needs to be) conveyed aurally. For monitoring purposes, the sonification design has to be more aesthetically pleasing and less intrusive, which typically also means that less data can be conveyed at the same time. On the other hand, in process monitoring typically less data needs to be conveyed than during process analysis. Thus, for process analysis, sonification techniques that can condense large amounts of data to a short sonification are particularly adequate.

To our best knowledge, there have previously been no approaches that apply sonification for the specific purpose of analyzing historical business process execution data. There is, however, a growing body of research concerning the usage of sonification for analyzing historical data outside the business process domain. Unfortunately, the majority of this research seems to focus on quantitative, continuous data, and less with qualitative data, like discrete events. Sonification for debugging purposes can be seen as a related area, as a process model is very similar to a program code, while one process instance can be compared to one program execution. Both consist of a control flow, a data flow and errors and warnings. Therefore, results from this area may also be beneficial for the domain at hand.

In (Ballora, Cole, et al., 2012), sonification is applied to detect anomalies in twitter data. However, the authors sonify condensed meta data, and not individual events. The same approach was taken by the tool *E-Rhythms Data Sonifier*[6], in which the frequency of different types of event occurrences can be mapped onto volume or pitch. One of the few approaches that sonify discrete events was presented in (Cullen and Coyle, 2005). The authors present a sonification of nominal event data using melodic earcons. Ballora, Panulla, et al. (2010) present a tool that sonifies HTTP requests from web server log files. It sonifies individual events that contain primarily nominal data (such as return codes). There are a few other approaches based on logs of individual events, but sonify only aggregated data (e.g. deButts, 2014).

3.3.5. Research Gap

Although there is very little existing research specifically targeting sonification for monitoring business processes, approaches exist for areas such as monitoring industrial production processes, which are specific forms of business processes. Existing approaches in this area focus mainly on scenarios where monitoring a specific system is the user's main, or even only task. However, particularly in small and medium-sized organizations, potential users, such as supervisors, typically have to dedicate their main attention to tasks other than process monitoring. Even so, they need timely information on process changes, updates, deviations, alerts, and problems. Auditory process monitoring systems that support such use cases need to be designed differently than ones that require permanent active attention. Among other factors, they need to be less obtrusive, in order to facilitate listening over long periods of time.

[6]http://www.jackjamieson.net/blog/e-rhythms-data-sonifier/

In contrast to research in sonification for business process monitoring, a variety of studies have investigated sonification for peripheral monitoring in general (see 3.3.1). However, it is interesting to note that although research often suggests natural soundscapes for dual-task and peripheral monitoring, very few quantitative experiments have been conducted in this regard. In the majority of studies, more basic sounds are used, such as (parameterized) simple wave forms, instead of natural sounds that belong to the same acoustic ecology. However, such experiments would be crucial to find out if natural soundscapes are perceived as more pleasing and thus can enhance acceptance for long-term use, as is often suggested.

Furthermore most studies for peripheral monitoring compare three conditions: visual-only, auditory-only and multi-modal. The sonifications are mostly designed in a continuous, stream-based fashion that aims to provide an acoustic overview of the running processes at all times, while some approaches are based on sporadic auditory alerts or cues, like they are used in many real-life monitoring scenarios. There are, to our best knowledge, no experiments in this area that compare not only the three mentioned conditions, but also continuous sonifications with auditory alerts or cues.

Furthermore, performance in the secondary task (typically process monitoring) is usually measured with user reaction times to stimuli, as well as attention switches, which are often measured by head movements. Often, additional factors are accuracy or error rates. However, what is typically not measured, are *false positives*, meaning situations in which the user assumes a need to interact when in fact no interaction was required. By measuring those interactions, one would also be able to identify interactions that were *too early* or *unnecessary* as well, in addition to such that were *too late*.

Finally, in a majority of quantitative studies on peripheral monitoring, the systems' effectiveness is measured for both tasks, but typically the users themselves are not questioned in terms of their preferences and opinions on the different conditions of monitoring, or their understanding of those. Very few approaches ask the user for visual/acoustic impairments, experiences with sound and/or domain knowledge concerning the second task. While many experiments seem to include no questionnaire at all, in some cases the participants are asked for their opinion on the presented modes. Seldom explored is the degree to which users find presented sonifications pleasing, exhausting or obtrusive, and whether they understood the individual mappings. This would be beneficial, however, as peripheral monitoring is typically done over a complete work day; thus, users' opinions on willingness and ability to use the sonification for an extended period of time should be considered.

In summary, several open questions remain concerning the support of users in monitoring as a secondary task with regard to sonification design, as well as how different types of sound- enhanced process monitoring affect attention and concentration in main and secondary tasks. These considerations are summarized in the following:

- To our best knowledge, no quantitative experiments using soundscapes in dual-task settings have been conducted so far.
- Sonification designs in previously conducted experiments for peripheral process monitoring were based on either auditory cues or continuous sonifications, but not both in combination.
- Furthermore, dual-task experiments that have been conducted in this area measure the performance in both tasks, typically using either binary correctness measures and/or response times but not more fine-grained measures.
- In most quantitative experiments, the user's opinion on, for example, the different conditions and his/her understanding of those is not gathered, or if it is, not in a very fine-grained way.

Process Analysis

It seems that no previous research has specifically applied sonification to the analysis of business process execution data. This is remarkable, given the fact that other domains boast a number of approaches that use sonification for retrospective data analysis. Although most of those approaches use data with different structures, approaches exist for domains with similar data structures as process execution data. In general. however, of the few approaches that exist in the area of sonification for the analysis of qualitative and nominal data (outside of the domain of business process data), the majority are *pure* sonifications, not multi-modal approaches. Therefore, especially questions concerning integrating visualization and sonification could not be answered by previous works.

3.4. Conclusion

Visualization and sonification both have their benefits and challenges, many of which depend on the data on which they are based and the problems they are intended to solve. However, many challenges of the two modalities can be addressed by combining the modalities into multi-modal systems. For both modalities, relatively little research has applied them to the monitoring and analysis of business process execution data, leaving many open questions, especially regarding the combination of both modalities. This research gap is especially extensive concerning sonification, as, unlike for visualization, few guidelines and mapping strategies exist thus far.

4. Requirements and Design Decisions

This chapter analyzes the requirements that multi-modal sonifications for monitoring and analyzing business process execution data should fulfill, as well as the suitability of existing sonification techniques and mappings for those tasks. Most of those requirements fall into two areas: a) the structure and peculiarities of the data that needs to be represented (see Sec. 4.3) and b) the different tasks that the users need to perform during real time monitoring and post-hoc data analysis. Those requirements have implications regarding sonification and interaction design.

This chapter contains ideas, results, figures, and text from my previous publications. The section with requirements for auditory process monitoring contains text, figures and results that were published in (Hildebrandt, Mangler, and Rinderle-Ma, 2014). The suitability analysis of existing sonification techniques contains text, figures and results that were published in (Hildebrandt, Kriglstein, and Rinderle-Ma, 2012b) and (Hildebrandt, Kriglstein, and Rinderle-Ma, 2012a). The findings on sonification designs for data of similar structure contains text, figures and results that were published in (Hildebrandt and Rinderle-Ma, 2013).

4.1. Requirements for Auditory Process Monitoring

This section presents the results of a focus group study intended to generate user requirements for auditory process monitoring. The domain of production monitoring has been chosen for the focus group as this is an area in which real-time monitoring has traditionally always received high attention. In order to support the discussion, two sonification mock-ups have been created before to the focus group. One of them demonstrates a sonification of different process events using melodic motif-based earcons. The second mock-up is based on KPIs, and demonstrates the mapping of different parameters to acoustic properties of sound loops. Example recordings of these mappings can be found at:

🔊 http://soundcloud.com/tobias_hildebrandt/

Three mappings have been deemed especially promising by users in informal tests:

1. Playback speed: One parameter can slow down or accelerate the playback speed of the background sample (in this case a recording of rain sounds), which leads to a lower or higher pitch.[1]
2. Multiplication with sine wave: One parameter is mapped to the frequency of a sine wave, which is multiplied with the background sample. This leads to a modified envelope of the background sample.[2]
3. Multiplication with formant: One parameter is mapped to the frequency of a formant synthesizer (a synthesizer that contains fixed frequency peaks) that is multiplied with the background sample. This also leads to different envelope modifications of the background sample.[3]

In first informal tests, participants were usually able to identify the points in time when a parameter changed its value (and thus influenced the sound file). For both wave-multiplication mappings (2 and 3) the participants seemed to be able to recognize intuitively if the value was increasing or decreasing. This direction could not be recognized, however, without instructions for the mapping on playback speed (mapping 1). A rising value in mapping 2 has been associated by participants with something becoming increasingly urgent, making this mapping suitable for parameters related to warnings or alerts. Mapping 3 conjured up associations with "static noise" or "something that is broken" and it was deemed to be artificial-sounding and unpleasant. Therefore, this mapping might suggest itself for parameters that are related to critical events or states. For mapping 1, no salient association has been observed. The most promising approach, mapping 2, has been selected to be presented at the focus group for demonstration purposes.

Our next step was to find out which sonification types (sonification of individual events vs. aggregated parameters) are suitable for different user groups (e.g. technicians, supervisors) in different settings (e.g. factory floors, offices) and in different scenarios (e.g. manual versus automated production). Furthermore, we were interested in general constraints on auditory process monitoring in such environments. A focus group method seemed to best fit such a semi-structured discussion.

4.1.1. Method

The focus group consisted of two moderators and six participants from different European countries. The main goal was to discuss auditory process monitoring in manufacturing from different angles, which is why the focus group consisted of practitioners as well as researchers. In order to be able to differentiate the monitoring needs and requirements for different manufacturing industries and

[1] direct link: https://soundcloud.com/tobias_hildebrandt/kpi-sonification-playback
[2] direct link: https://soundcloud.com/tobias_hildebrandt/kpi-sonification-sine-wave
[3] direct link: https://soundcloud.com/tobias_hildebrandt/kpi-sonification-formant

company sizes, we approached members of a consulting and systems development company who deal with different types of companies and thus different production monitoring scenarios. This group of practitioners was complemented by a participant who has in-depth knowledge of process monitoring in his specific domain. The following individuals participated in the focus group discussion:

- Participant A is the operations and improvement manager at a medium-sized engineering and manufacturing company. He has been able to gain broad experience in monitoring and improving business processes, specifically concerning quality standards.
- Participants B and C are the directors of a small company that provides training, consulting, and systems in the area of production monitoring. Participant B, the managing director, has over 20 years of experience in control systems within the manufacturing industry. Participant C, the technical director, has broad experience as a control systems consultant as well as in information communications technology. Participant D works in the same company as a development manager.
- Participant E is an associate professor and research project manager in the area of business process engineering and industrial management at a university; participant F is a researcher at the same department.

The focus group interview was structured by a couple of open questions while the moderators asked further questions if deemed necessary. The discussion was started by giving a short introduction on the general topic of sonification and our approach for real-time monitoring. Afterwards, short audio recordings of the event sonification mock-up and the parameter-based sonification mock-up were presented. This was done to enable the participants to familiarize themselves with the concept and possibilities of sonification. Following the discussion on the presented audio recordings, the participants were asked if they could imagine (a) the usage of sonification for production monitoring in general and (b) the concrete application of the two presented concepts. This was followed by a more detailed discussion concerning the suitability of different types of sonification for various user groups and constraints, concluding with feedback and suggestions. The following subsections summarize the results of these discussions and quote noteworthy statements.

4.1.2. Benefits and Challenges of Sonification in Production Monitoring

Several potential challenges for the usage of sonification in production monitoring were mentioned:

- Deaf people are not able to hear sonifications. (B)

- In many production environments people have to wear noise protection. (A) - Response: Noise protection headphones can be used to transmit sonification. (B)
- Visual displays are very contained (spatially), audio travels and can therefore create noise pollution. On the other hand, employees could use headphones. (C)
- Previously, many machines conveyed verbal warnings that have been deemed too irritating and have been deactivated. (B) - Response: Non-speech sonifications are different than speech, as speech needs to be processed actively, while tones can be processed more passively. (D)

In general, it seems that auditory process monitoring is already an accepted technique in the manufacturing sector:

- "On the other hand, quite often some processes already DO have sound. So if you walk past it, it will go 'BIP BIP' and you know everything is fine." (B)
- "Sometimes when our machines broke down, when it's missing cork stoppers, for example, it spreads a red light. Sometimes our customers ask us to introduce a buzzer. Alarm." (A)
- "Perceptually, what it is, is: previously, you would have people working in a factory for 25 years, and they would know things by sound, by smell, these sorts of things. What we're doing is, removing the need for the 30 years, we are taking the knowledge. But we are still using the same almost holistic approach, where you are listening instead of actively doing things, and you could still properly hear. But you're taking the knowledge, you are already doing the knowledge, but instead of (,) you are still using the same method that the experts used, it's just (,) in a different world." (D)

To summarize: a few challenges concerning the usage of sonification in production monitoring have been mentioned; however most arguments have subsequently been weakened by other participants. Furthermore, it seems that auditory monitoring already has a place in production monitoring where multi-modal monitoring combining different senses has even been described as *holistic*. Furthermore, customers have specifically asked to equip machines with auditory warnings.

During the presentation of the KPI-based recording, the participants were asked to identify changes in a fictitious KPI by noticing changes in the resulting sonification. Most of the participants were able to do so without further instructions. As this mock-up is more subliminal than the event-based one, concerns regarding the attention that is necessary to perceive changes were raised:

- It should be tested if people could still hear changes in sound while they are talking to other people. (D)
- The monitoring personnel would have to actively pay attention to derive the desired information from the sonification. (B) - Response: "There is a part of your brain which kind of tunes in into that background, which is

like: danger! You know, you are kind of like, something's not quite right." (C)

- Instead of mapping to natural background sounds, the KPIs should be mapped to music in a playlist. (A)
- The KPIs can be mapped to several acoustic parameters of played music (such as low-pass-filters). (D)

To conclude: at least several of the participants seemed to be able to notice sonification-induced changes during audio playback. Participants raised suggestions regarding further testing and new sonification methods for future prototypes.

4.1.3. Auditory Process Monitoring in Different Scenarios and Settings

In general, all participants claimed that they believe sonification can be useful in production monitoring, but only under certain conditions and in specific scenarios. One participant suggested that a possible sonification in production monitoring would have to go beyond already existing auditory alarms and offer something new (B). During the course of the discussion it became clear that the potential user group (factory workers, maintenance staff, engineers or supervisors and managers) as well as the work environment (on the factory floor or in a separated office) both heavily influence the circumstances under which auditory process monitoring is beneficial. Concerning the usage of sonifications on the factory floor level, one participant recounted his personal experiences on working as an operator for a big manufacturer of consumer electronics several years ago.

> "We constantly had a sound testing booth. The whole floor could hear it. The noise was not much different to these (remark: the presented event-based mock-up). As an operator: first you kind of notice it. But in factories, unlike what we are doing, which is sort of creative work, you are doing a mechanical job. You got into a different mindset. As long as the frequency of it isn't too infrequent or changing, its OK. If its quite constant, as a factory worker, it's OK. So I don't think, being annoying I don't think you should worry too much about. From a maintenance or a operator perspective, you have got so many annoying noises around anyway, doesn't matter." (D)

At a later point, however, the participant remarked that such a noisy environment also poses challenges to sonification in terms of perception. Later he remarked that the sonification of the statuses of individual machines could prove difficult in case the sound source would be located at those machines, unless the machines are far away. He therefore concludes that sonification should concentrate on aggregated KPIs, such as line-wide KPIs or even factory-wide KPIs. Another participant suggested conveying an auditory alarm to maintenance workers in case there is a

problem with machines (C). However, even though most participants seemed to support the idea of sonification on a shop floor level, one participant stated that such a system would be too disturbing (B).
In contrast to the application of auditory process monitoring on the shop floor level, the participants seemed unambiguously positive towards its usage in office scenarios (such as e.g. in maintenance or supervisor offices). One participant stated that auditory process monitoring might be beneficial to supervisors to free them up from looking at a screen, if designed in a subliminal fashion (C). This statement was complemented by another participant:

> "If it's one guy in an office, and he's monitoring something, then maybe yes. Someone on the shop floor, no." (B)

The same participant continued to suggest that a sonification of the current status might make sense in engineering offices, as they are typically separated from the factory floor. Independent of the workplace (shop floor vs. office), the participants also suggested that the requirements for and potentials of sonification differ depending on the type of data to be conveyed. Several participants are positive towards the sonification of individual events in case they are conveyed in quite regular intervals (B, D). One participant suggested to keep in mind which action is desired from the user when monitoring KPIs. He exemplified this by stating that the sonified KPIs should convey data that goes beyond what could be conveyed by mere alarms. He further stated that a performance analysis with the aim to improve production processes is usually done in a weekly basis on historical data, not in real-time. Another participant suggested to use sonification to convey the statuses of Kanban systems (E).

4.1.4. Summary and Discussion of Results

To conclude, during the focus group participants stated several points that need to be considered when developing sonifications for this domain, specifically that they have to be subliminal, and that they have to offer more than plain auditory alarms. Particularly, the proposed system has to take into account which actions have to be taken by the users when they hear changes or events conveyed aurally. A critical remark has been made concerning the differentiation of different instruments. This is a valid remark, as even though recognition in sonifications typically increases with training time, the number of different activities that can be identified in such a way are limited. For small processes with only a handful of different activities, the presented approach might, however, provide a valid method. The remark that the presented approach could possibly be annoying, depending on the frequency of occurring notifications, cannot be easily dismissed. Therefore, as already mentioned, the presented approach is probably only suitable for processes that consist mostly of manual tasks and that therefore exhibit a relatively low frequency of occurring events. Concerning the KPI-based mock-up, it needs to be evaluated if the subtle changes can still be perceived (and

interpreted) during operation in a factory. Other mapping techniques, such as mapping onto music files, need to be investigated.
Concerning potential scenarios, it seems apparent that a majority of participants does not see the main potential of sonification for production monitoring on the factory floor, but instead mainly in offices of engineers and production supervisors, although there seem to be some conflicting opinions regarding this. All participants, however, seem to agree on the usefulness of sonification in engineering and supervisor offices, both for event-based (as long as it offers more insight than auditory alarms, and its frequency does not vary too much) as well as for sonification that is based on quantitative parameters (as long as it is subliminal). Concrete scenarios that have been mentioned include the sonification of machine statuses for maintenance personnel, the sonification of factory- or line-wide KPIs for supervisors/engineers and Kanban sonifications.
Fig 4.1 summarizes the suitability of different sonification approaches (event-based vs. indicator-based) for different user groups and different data densities.

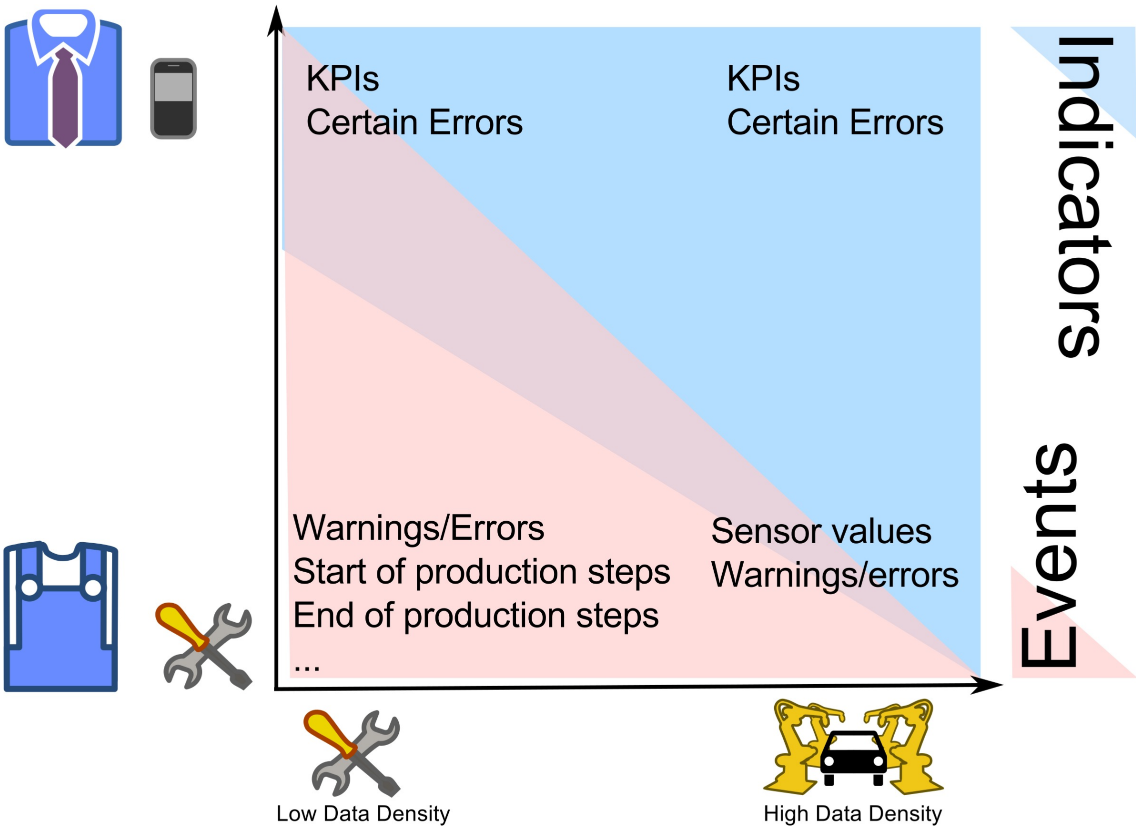


Figure 4.1.: Suitability of sonification approaches for different scenarios.

Furthermore, it was suggested that the system should be designed in such a way that it is both unobtrusive and not distracting, but alarming and attention-grabbing when necessary. This might put limits on the presented event-based sonification mock-up based on melodic motifs and suggests that the also presented natural-sound-based approach might be more promising, as such sounds are generally considered to be more unobtrusive. This links with the fact that several participants suggested that an auditory monitoring system has to go beyond simple conveying of event occurrences - something that existing auditory alarms already offer, albeit in a less detailed and fine-grained manner. Furthermore, as suggested, typical users of such monitoring systems are likely to be engineers and supervisors instead of line workers, and thus are more likely to be interested in aggregated quantitative parameters anyway.

Therefore, it seems sensible to build on the presented concept of using natural sounds to convey KPIs and sensor data, and try to incorporate important events such as warning or errors into this system where necessary.

A sonification concept similar to that developed by Gilfix and Couch (2000) seems best suited to fulfill these requirements. In their system, different types of natural sounds are integrated into a homogeneous *soundscape*, as found in nature (e.g. a forest environment containing sounds of wind, rain and birds).

After the presentation of the event-based mock-up, the participants were asked if they would be able to identify the different event types in the presented recording, and possibly even the related activities. Several, but not all, participants stated that they would be able to, while a few comments on the presented mock-up were made:

- There is a limit to how many different instruments can be remembered and distinguished. (C)
- "Just to play devil's advocate - what if somebody said: I already have something that beeps in a negative way, when something is wrong? If somebody said that, what would be the counterargument?" (C) - Response: "It is different - suppose a machine is working, it seems that everything is OK. But you can say, there is something in the movement not normal (Participant imitates machine noise). Something strange. In a situation like that, you can imagine the normal pattern. You need to correlate this, if you correlate it, you can identify the difference." (E)
- A sonification of this type can get annoying, depending on the frequency of sonified events. (C)

To summarize: most participants mentioned that they would be able to at least identify the different event types from the played melodies, while this seems more difficult for the related activities.

4.2. Requirements for Auditory Process Data Analysis

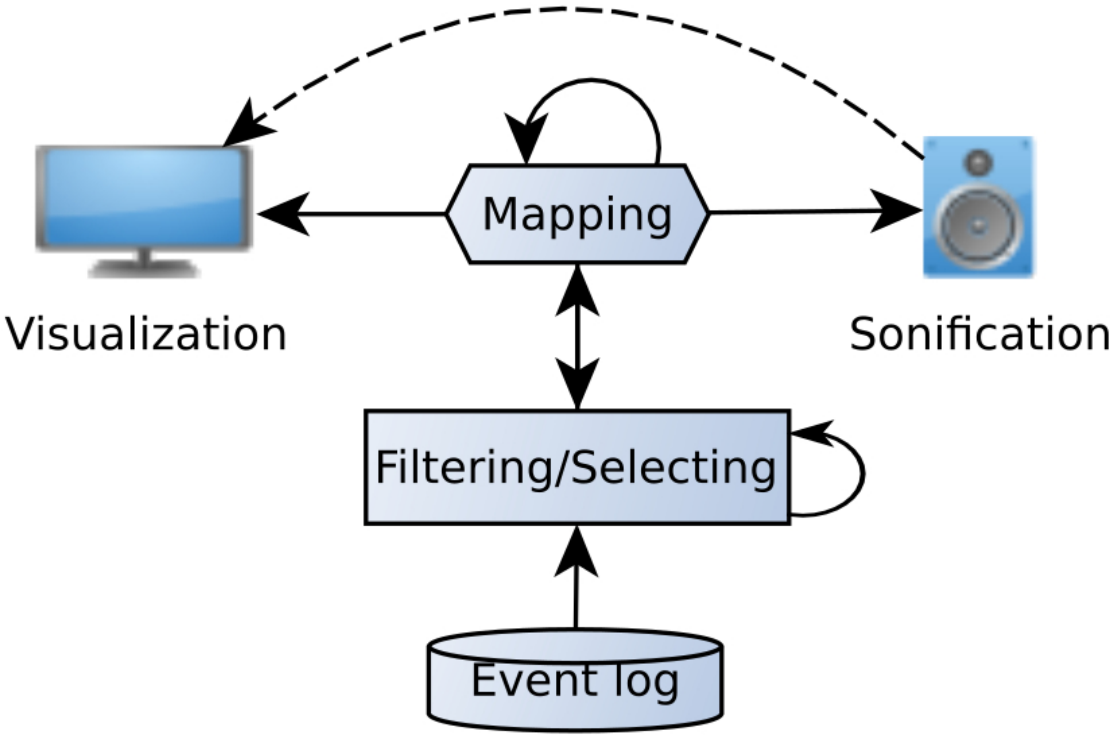
Several hypotheses can be derived from the related work in terms of interaction design and technical aspects.

4.2.1. Interaction Design

It is obvious that sonifications for the active analysis of process execution data can be designed more intrusive than those for (passive) process monitoring, as (a) process analysis is typically performed as the main and only task and (b) it is usually not performed for complete working days. However, the literature shows that sonifications that are perceived as "aesthetically pleasing" are typically better accepted by users. As preferences are subjective, users should be able to customize the mapping from data to sound, and adjust it according to their preferences.

The interaction design should support the user in his/her process analysis workflow as much as possible. One major aspect in reaching this goal is to simplify the creation and testing of a sonification mapping and help the user to memorize how information aspects available in the event logs are sonified. As the user may not be trained in defining a sonification, an iterative creation of the mapping might be preferable. In such an iterative process, as demonstrated in Fig. 4.2, users should be able to quickly change mappings in order to try out different options.

Figure 4.2.: Interaction flow of analysis concept.



Both visualization as well as sonification should follow the information-seeking mantra "Overview first, zoom and filter, then details-on-demand" (Shneiderman, 1996), which should be reflected in the user interface. This includes the incorporation of the ability to zoom in visually, as well as speed up and slow down the sonification.

4.2.2. Technical Aspects

There are several requirements that a tool for multi-modal process data analysis should fulfill on a technical and functional level. In general, as the aforementioned XES file format is the IEEE standard for event logs[4], a multi-modal analysis tool should be able to import files of said format and handle at least the basic concepts of the format.

The field of visual analytics typically builds on a strong interconnection between machine learning and visualization, where the user is able to derive hypotheses from a visualization, which at a subsequent step can be verified by machine learning, and vice versa. It can be assumed that this approach can also be applied to multi-modal, audiovisual analytics. Therefore, it would be beneficial to integrate the multi-modal presentation of process data with other functionality to perform statistical computations, machine learning and data processing operations as tightly as possible. As sophisticated machine learning algorithms already exist, it would not make sense to reinvent the wheel and implement those again in an audiovisual analytics tool. Instead, it is probably more advisable to integrate the approach into an existing process mining suite that already possesses these functionalities. As the already mentioned ProM suite is probably the most extensive and versatile process mining suite, it is recommended to integrate an audiovisual approach as a plugin into ProM.

The ProM framework itself is developed in Java, and therefore plugins also have to be developed in this programming language, or at least in a language that runs in the Java Virtual Machine. This puts technical restrictions on the sound synthesis, as audio programming languages such as SuperCollider[5] typically enable sophisticated synthesis methods better than Java libraries. It would be possible to develop the sound synthesis in another programming language and integrate it into the plugin using open protocols such as OSC (Open Sound Control[6]), but this would make the installation of the plugin by the intended user, and thus the usage, more difficult.

4.3. Suitability of Sonification Techniques

The following subsection presents the different existing sonification techniques and their theoretical suitability for business process execution data for all phases

[4]http://standards.ieee.org/findstds/standard/1849-2016.html
[5]http://supercollider.github.io/
[6]http://opensoundcontrol.org/spec-1_1

of the business process life cycle.

4.3.1. Basic Sonification Methods

The four sonification methods presented in Section 3.2.3 are not equally suitable for the representation of business process data. The method of *audification* is not as flexible and versatile as the others that have been presented. Due to its necessity to incorporate a large number of quantitative data, it seems less suitable for sonifying cases involving little or no quantitative data, as opposed to events. Additionally, due to the fact that audification converts data directly into digital sound signals instead of relying on *high level* elements such as emulations of musical instruments, it might be very difficult or even impossible to distinguish between several *streams* of sounds. This makes audification unsuitable for most sonifications of process data, as different process instances typically run at the same time and in a different execution state. A process sonification would therefore have to fulfill the requirement to enable the users to distinguish between different execution events and states, thus likely proving difficult when relying solely on the limited means of audification.

Auditory icons

Auditory icons, however, seem suitable for the sonification of business process data: in the analysis phases, but also in the operation phase, the sonic pendants of the involved activities and events could be played back upon their incidences. As an example, the sound of a shopkeeper's bell could signify the reception of a new order. Analogously, the process event "customer has payed his invoice" could be conveyed by playing the sound of a cash register being opened, while the activity "delivery" could be sonified by applying motor sounds. Fig. 4.3 shows a schematic overview of how such a sonification of a process instance could be realized. The x-axis is the time axis, whereas each row on the y-axis contains a sound file that represents one activity or event. These sound files are played back sequentially from left to right.

In this example, the sounds that convey events have been assigned a fixed length of 1.5 seconds (as the time axis in the lower part of the figure shows). The lengths of the audio signals that convey activities, on the other hand, represent the actual duration of the represented activities. Thus, a motor sound with a duration of three seconds could, depending on the scaling, signify, for example, that the transport took three days. Analogously, a *silence* between activities and events could signify a waiting period. If, for example, there is a gap of two seconds between the playback of the sounds that represent the activities "production" and "packaging," one can conclude that there has been a waiting period of two days between the production of and the transport of said goods. This could be a hint that there are inefficiencies in the process.

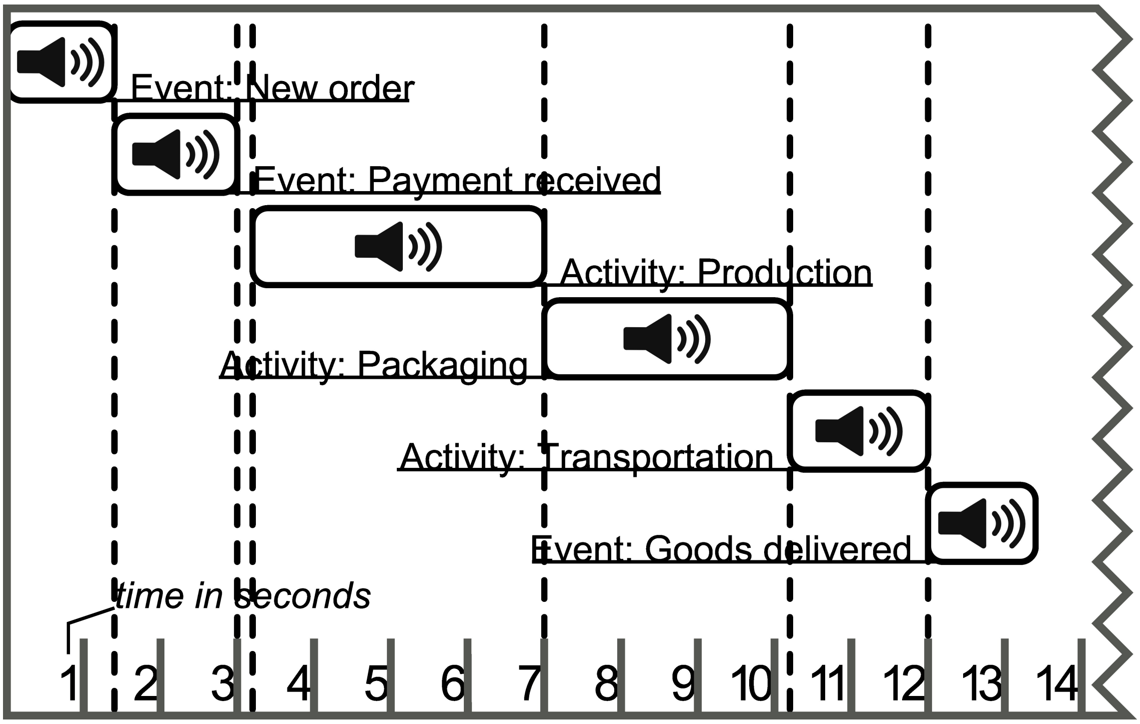


Figure 4.3.: Schematic overview of an auditory-based sonification of an exemplary process instance.

Audio files of three different instance sonifications of this example process are available online[7]. The audio file "Example one - average process instance"[8] shows a sonification of an average process instance. The audio file "Example two - no payment"[9] is a sonification of a process instance in which no incoming payment has been registered. The audio file "Example three - Production and transport delayed"[10] is a sonification of a process instance, in which the activities production and transport have been delayed (which can be recognized by the pauses before the respective audio signals). This simple example tries to show that auditory icons are able to point out deviations in process instances.

Earcons

Earcons are, in a similar fashion, suitable for process sonifications because their concept is similar to that of auditory icons. Their advantage over auditory icons is that the sonification designer is more flexible in choosing appropriate sounds for process events and execution states. For some states, it could prove difficult to find real-world-sonic analogies. For example, it could be a challenge to find sounds that are sonic analogies to the states "customer is already registered"

[7] http://soundcloud.com/tobias_hildebrandt/
[8] direct link: http://soundcloud.com/tobias_hildebrandt/business-process-sonification
[9] direct link: http://soundcloud.com/tobias_hildebrandt/business-process
[10] direct link: http://soundcloud.com/tobias_hildebrandt/business-process-1

and "new customer". This differentiation would, therefore, be hard to convey using auditory icons. As earcons are not based on real-world sonic analogies, a sound designer could easily assign almost arbitrary sounds to sonify these states. However, potential users may need more training with earcon-based sonifications than with those that are based on auditory icons, as earcons are mostly not as self-explanatory as auditory icons. It would, however, be reasonable to use several distinctive, memorable sounds or short rhythmic/melodic patterns to sonify the various events and activities of business processes.

Parameter Mapping

At first sight, parameter mapping might not be the most obvious choice - parameter mapping relies on quantitative data that varies over time, rather than on information on events and their sequences of occurrence, as is typically the case in business processes.

One process can have hundreds of instances running in parallel, whose events and sequences could be sonified using auditory icons or earcons. However, one would probably not be able to distinguish individual process instances from the resulting sonification, as the individual sound streams would be overlaying each other. Thus, parameter mapping-based sonifications might be more suitable for this task, as they could convey various aggregated information on running process instances (instead of directly sonifying all process events). Parameter mapping also lends itself to the task of analyzing quantitative data. KPIs might be mapped to one or several sound streams. These sound streams might then, for example in the process operation phase, be played back continuously, making it feasible for the user to recognize patterns and modifications. Some of the data sets that might be mapped to attributes of such sound streams could be:

- The current number of running instances of a specific process. If several processes should be monitored simultaneously, one could assign a distinguishable sound stream to each of these processes and map the number of processes instances to the different sound streams accordingly.
- The current average execution time of a process or of certain activities.
- Current average capacity utilizations depending on the type of processes (e.g., machines or staff).
- The current number of processes in faulty execution states.

These quantitative parameters could either be mapped onto various attributes of a sound stream (e.g., volume, pitch, panning or timbre) or, if several sound streams are being played back simultaneously, to all sound streams respectively. The utilization of several parallel sound streams could, for example, be realized by the usage of emulations of different musical instruments like piano, guitar or percussion, while each sound stream uses a different instrument.

Table 4.1.: Suitability of basic sonification methods for typical process scenarios.

Process scenario	Audification	(Parameterized) Auditory Icons/Earcons	Parameter Mapping
Process Models/Process Patterns		X	
Individual process events		X	X
Aggregated Quantitative Data		(X)	X

During the process phases "operation" and "analysis", a parameter mapping would probably be well suited to convey a general overview of the running process instances and cases that are possibly critical.

In general, processes often generate quantitative information (which would make parameter mapping seem to be a suitable sonification method), but overall probably contain more information about events and sequences (which are traditionally domains of auditory icon- and earcon-based sonification).

4.3.2. Advanced Sonification Methods

There are, however, sonification methods that are combinations of parameter mapping and auditory icons and earcons, and combine the *simple* event-occurrence method of auditory icons and earcons with parameter mapping: parametrized auditory icons and parametrized earcons. In these, sounds convey the occurrence of events, but at the same time, quantitative data can be mapped to these sounds. This mapping is analogous to the parameter mapping method of mapping data to sound attributes. Parameterized earcons usually provide more extensive means to map data to sound attributes in comparison to parameterized auditory icons. Earcons are usually generated in real-time and therefore are more subjected to manipulation. Auditory icons, on the other hand, are often based on prerecorded audio samples, which can easily be adjusted in terms of volume, pitch or panning, but are less flexible when it comes to timbre manipulation. However, approaches towards the generation of auditory icons in real-time exist as well. Thus, the use of parameterized auditory icons or earcons can convey not only that a certain event has occurred, but also one or several quantitative data attributes that are connected to that event. One could, for example, imagine an auditory icon conveying the occurrence of an event "incoming payment," while the sum of the payment is mapped to the volume or the pitch of that auditory icon.

Table 4.1 lists the sonification methods that are potentially best suited to convey the data of the monitoring and analysis phase (indicated with an 'X') and these that seem suitable with restrictions (indicated with an '(X)').

Summary

In conclusion, of the presented methods, parameterized earcons and parameterized auditory icons seem to be best suited for presentation of individual events as they occur in all phases of the process life cycle, while parameter mapping seems best suited for aggregated KPIs and other indicators, as they mainly occur in process monitoring and analysis. Such data could be mapped to parameterized earcons or auditory icons, but parameter mapping offers more possibilities to map quantitative data to sonic attributes.

4.4. Sonification Design

This subsection presents findings from related research concerning the design of sonifications for the representation of data of similar structure as that of business process execution data (see Sec. 3.3.3). Due to the fact that, as previously mentioned, only a small amount of research on the retrospective analysis of similar data using sonification exists, the survey focuses on sonifications for real-time monitoring. In both direct and indirect monitoring, attention allocation is an important aspect that needs to be considered when designing auditory process monitoring.

4.4.1. Attention Allocation

In general, the more information a sonification tries to convey, the more active attention is required (Hudson and Smith, 1996). This could negate one of the greatest advantages of auditory process monitoring: its possibility of passive listening that enables concentration on parallel activities. Thus, it has to be kept in mind that humans can only process a certain number of audio streams and acoustic mappings in parallel, especially passively. On the other hand, it is desirable to aurally convey at least a certain number of parameters and events in order to free users from observing screens as much as possible. Therefore, it is crucial to maximize effectiveness by concentrating on sonifying only those parameters and events that are most relevant to the user, e.g. by offering customizable filtering mechanisms, and to convey the other parameters by means of visualization. The number of data dimensions and events that are conveyed in parallel can be maximized by applying principles from perception, such as the segregation of different audio streams by applying different timbres, and the avoidance of similar frequency ranges and positions in the stereo field (Walker and Nees, 2011).

Sonifications that are designed with these factors in mind should be able to attract users attention when deviations or specific situations occur. Otherwise, such sonifications should be unobtrusive enough to enable the user to concentrate on his/her main task and perceive information in the background. An example of this principle could be driving a car. The sound a car motor makes during *normal* operation is usually still subtle enough to let the user concentrate on his/her main task, in this case driving, while sudden changes in motor noise attract users attention and may even give an experienced driver an idea about the nature of possible technical problems.

It seems beneficial to offer the user possibilities to select what data is sonified in what level of detail (e.g., by offering drop-down menus with various filters and selections) in real time. Nevertheless, in the following, guidelines on how to design an initial prototype for auditory business process monitoring are presented. These guidelines take into account event occurrences as well as continuous KPI conveyance.

4.4.2. Individual Events

As one outcome of the literature analysis was that different constructs can be conveyed by melodic earcons (such as motifs, e.g. following the guidelines of Vickers and Alty (1998)) with different timbres, it can be expected that those concepts can also be applied to the domain at hand. These motifs could perhaps take into account the concept of melodic contours, as suggested by Cullen and Coyle (2006). In principle, process execution data has two orthogonal types of constructs that can be conveyed:

- The hierarchical relations between processes, instances and activities
- The hierarchy of event base types (control flow, data flow, alerts) and the concrete event types (e.g. activity started)

In most cases, users are probably less interested in the event type hierarchy, and more in distinguishing the different individual events. Thus, it is possible to reduce the complexity of the conveyed information by only distinguishing between the different event types without conveying their type hierarchy. This would leave only the hierarchy of processes, instances and activities to be sonified, which could be done by either using hierarchically structured motifs or hierarchically grouped timbres.

As companies typically have several processes, for each of which might exist a high number of instances and activities, users could easily become overwhelmed when trying to convey all of those concepts over hierarchically structured timbres. Thus, it seems sensible to let users choose the construct that is most important for them to distinguish, and then convey only this concept (either processes, instances or activities) using timbres that are as distinct as possible. Thus, when using motifs and timbres to convey the different constructs, two options remain:

1. Convey the different processes, instances or activities over different timbres and the different event-types over distinct motifs

2. Convey hierarchical relations between processes, instances and activities over hierarchical structured motifs, and the different event-types over different timbres

Which of the two options yields better results will need to be evaluated in experiments. As this will also likely depend on the type of processes a company runs, along with the information needs of the user company, users should be able to choose between the available options. It seems, however, that the motif creation is easier for (1), as it is probably easier for users to associate a specific earcon or contour motif with an event type (e.g. "activity started") than with an abstract concept (like e.g. "process instance 4"). As an example, a contour motif consisting of a few notes that rise in pitch could signify that an activity has started, while a falling pitch could mean that an activity has finished.

There may also be cases where users are only interested in distinguishing one of the two concepts, and may choose to apply different timbres as well as motifs to distinguish this concept. For example, for some users it may be particularly important to distinguish between a variety of events (e.g. between different warning/error types) while the specific process or instance in which they occur may be less important. In this case, the user may decide to map the event type hierarchy to hierarchical motifs that also are played in different timbres.

If the users would choose (2), all the individual events that occur could then be conveyed by playing the specific motif of the instance or the activity. For both, (1) and (2), a possibility to convey the fact that an activity is ongoing (i.e that the "activity started" event has occurred), could be to loop the motif that is assigned to this event until the respective "activity finished" has occurred. Of course, this method could potentially be annoying in cases of long-running activities and is only beneficial for a limited number of parallel activities.

4.4.3. Quantitative KPIs

As already discussed, aggregated quantitative parameters are an essential means to monitor the performance of business processes. Techniques from parameter mapping are an established means to convey such continuous, quantitative data, which is why it can be assumed that such techniques will be suitable for the conveyance of KPIs as well.

For KPIs on the process level, one solution could be to create a different continuous drone sound for each process model and map different process level KPIs to acoustic properties of these drones. If possible, the same concepts that are used to distinguish between the different processes for the conveyance of event occurrences could also be used for those drones. If, for example, events of different processes were represented by different grouped timbres (e.g. stringed instruments), the respective KPI drones for each process could also correspond to those timbres (e.g. by using a drone that sounds like a stringed instrument).

If a company runs too many different processes to enable a discrimination over different drones, a possible solution would be to use only one or a low number of drones and map selected process-spanning KPIs to acoustic properties of these drones. Instance level parameters (such as e.g. variable values that are of particular interest to the user) could, for example, be mapped to acoustic properties of all motifs that are being played for a particular instance.

4.4.4. Design Decisions

One could imagine, depending on the scenario, either a constant real-time sonification of all running process instances, or a *sonic summary* of a certain time period (for example a shortened sonification of the last 24 hours). After a learning phase, it should be possible to detect deviances or critical situations in such a sonification during the execution of process instances. A multi-modal solution could combine sonification with the possibility to visually explore root causes or other details, once such a situation has been recognized in the sonification. Similar approaches could be applied in the process analysis phase.

Users should have the possibility to adjust the systems to their preferred granularity level. Some users might only want to be informed about errors or certain alerts, while others nay also want to hear constant sonifications of various KPIs. There might also be users who want to hear a sonification of all events that occur for certain, or even of all processes or instances.

It can be expected that in companies that run processes in which thousands of events per minute occur, a sonification of all individual events would not be very helpful and interesting to the user, while for processes that only have a handful of events per day (such in processes whose tasks are mainly executed manually) such a sonification might be beneficial.

Thus, it is recommended to design the system in such a way that it is flexible enough to offer different modes of conveying event occurrences in order to adjust to different data densities that exist for different processes. For example, the Sonification Design Space Map (Campo, 2007) suggests that a sonification that should be based on individual notes/motifs for high event frequencies, as they might exist for highly automated processes. For such processes, grain clouds based on granular synthesis, as suggested by the Design Space Map, might be a better option than the usage of notes or motifs. A possible solution could be a system that automatically switches between different modes of aurally conveying event occurrences, depending on how many events currently occur per second or minute. Furthermore, even for cases that are suitable to convey individual event occurrences using, for example, motifs, it probably makes sense in terms of perception to only play a very limited number of notes simultaneously (if any), and instead queue occurring events in order to play them sequentially (perhaps starting with urgent events such as alarms). Another possibility would be to only play sounds in certain intervals, for example every second, in order to maintain a rhythm that sounds less chaotic.

Another possibility would be a multi-modal solution that combines visualization and sonification. During normal operation, the system would sonify notifications, alerts, and occurring events (if desired) as *sound events* as they occur. KPIs, on the other hand, are sonified as continuously updated *sound streams*. Another possibility would be to also play the sounds of events continuously, from the start event of an activity to its end activity. Of course, such a solution would only be feasible with (a) activities that have a short duration, or earcons that are very long or can be looped and (b) in systems with a low number of overlapping concurrent activities, as the result could otherwise be a large number of concurrent sound streams.

4.5. Conclusion

Process monitoring and retrospective analysis each have different requirements. For monitoring, one primary requirement is that the sonification should be unobtrusive, yet offer benefits over traditional auditory alarms and alerts. During process analysis, the user interaction should be designed as interactively as possible and integrate tightly with the user's analysis workflow.

To summarize the results from the literature analysis, based on the suggestions of the respective authors of the analyzed articles, the following considerations should be taken into account when designing sonifications for real-time monitoring of processes that are based on mainly qualitative, event-based data:

- Users should be able to customize the mapping from data to sound.
- If concrete auditory representations for the occurring events are available, the usage of auditory icons can yield positive results.
- When occurring events are mapped to earcons, complex timbres (possibly based on real-world instruments) should be preferred over simple timbres (like sine waves).
- Earcons should take concepts from the areas of motif design and melodic contours into consideration and adhere to *musical* concepts (such as the western tonal system).
- If motifs are being applied, they should differ not only in pitch, but also in rhythm and intensity.
- Different concepts can be conveyed by using different motifs (possibly hierarchically structured) and/or different timbres.
- In general, rhythm and percussion should be included in sonifications.
- Continuous sounds (such as drones) should be used to convey the duration of ongoing activities.

5. Multi-Modal Sonification for Business Process Monitoring

This chapter describes several approaches for multi-modal sonification during the operation phase of the business process life cycle (see Fig. 5.1). These include the development and evaluation of the SoProMon (Sonification for Process Monitoring) system for testing multi-modal sonifications as secondary task based on the previously defined requirements. This system is based on a prototype for an evaluation system with different components, as well as a tool chain for analyzing the experiment data. One concern frequently raised during the focus group discussion was that auditory process monitoring has to be subliminal, while offering more than plain auditory alarms. Furthermore, it was suggested that the frequency of event notifications should not change too much. The presented sonification design that is used to evaluate the SoProMon system was developed to fulfill both requirements. One requirement that the presented sound design does not yet fulfill is the focus on aggregated quantitative parameters, which is why this section also presents mock-ups for two further sonification designs, the latter of which is based on KPIs.

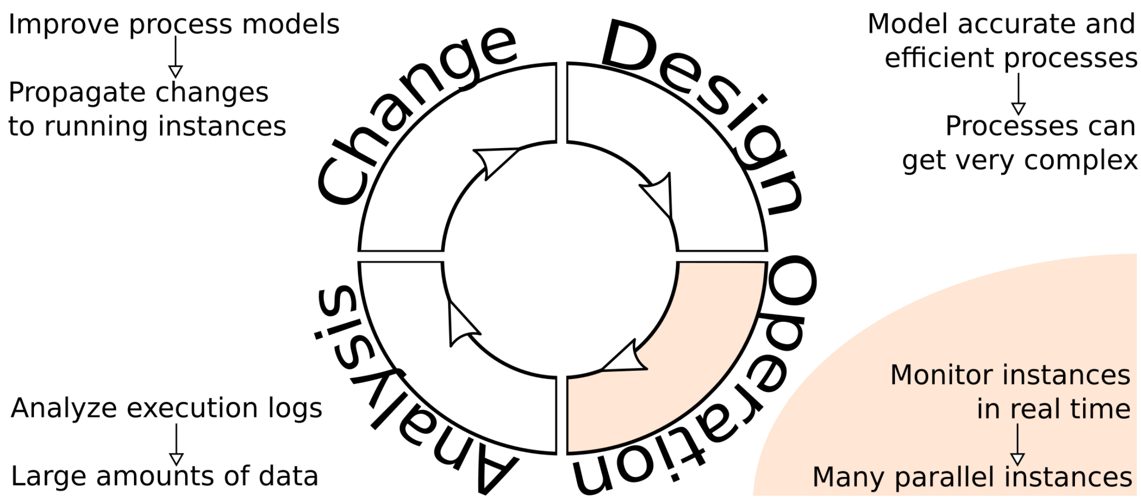


Figure 5.1.: Operation phase of the business process life cycle.

The SoProMon system and its evaluation are presented in Sec. 5.1 to Sec. 5.5, while Sec. 5.6 presents an approach that is more oriented towards direct, single-task monitoring. This chapter ends with the description of an additional sonification concept in Sec. 5.7, which builds on the feedback of the user study described in this chapter. As the results of previous chapters have shown that monitoring is typically performed as a secondary task, the main research focus has been put on the SoProMon system.

This chapter contains ideas, results, figures, and text from my previous publications. The description of the SoProMon system contains text and figures that were published in (Hildebrandt, Hermann, and Rinderle-Ma, 2014). The description of the SoProMon experiment and its outcomes contains text, results, and figures that were published in (Hildebrandt, Hermann, and Rinderle-Ma, 2016). Some of those results were also published in (Hermann, Hildebrandt, et al., 2015). The description of the motif-based concept contains text that was published in (Hildebrandt, 2013). The description of the soundscape-loop-based concept contains text that was published in (Hermann, Hildebrandt, et al., 2015).

5.1. The SoProMon System

Our system provides a generic testbed for evaluating different means of conveying process-related data with a focus on the aforementioned challenges of attention allocation in dual-task settings.

The core software components are (a) a process simulation, (b) a visual monitoring system including graphical user interface elements (buttons) to intervene, (c) a sonification system that allows different sonification types to be plugged in and (d) a main task console to draw the user's attention to a (different) focus. This is complemented by a set of service modules for logging all relevant data, including a video camera mounted atop the user to store head orientations, and calibration modules for user adjustment of the sound levels.

The system is highly modular and flexible, and individual modules can be replaced by other customized code if required. For our first practical implementation we decided to create a setting where a user is seated in front of two monitors, oriented perpendicular to each other, one for the main task and one for the monitoring console, in order to be able to stimulate and observe attention shifts. Furthermore, the keyboard and the mouse were fixed to the table, so that they could not be moved (see Fig. 5.2). Yet other implementations, such as letting users move freely in the room and solve practical problems, as opposed to computer tasks, are certainly also conceivable.

Concerning (a), the process simulation, we decided to base the sonification on a real-time simulation of a simple production process, in order to design the scenario to be as lifelike as possible. This also provided the users with a natural motivation to observe process sonifications, and grabbed their immediate attention in severe cases. This simulation consists of six production steps that partially run in parallel and require input of one or more previous production steps. We chose the arbitrary number of six, as it is high enough to provide

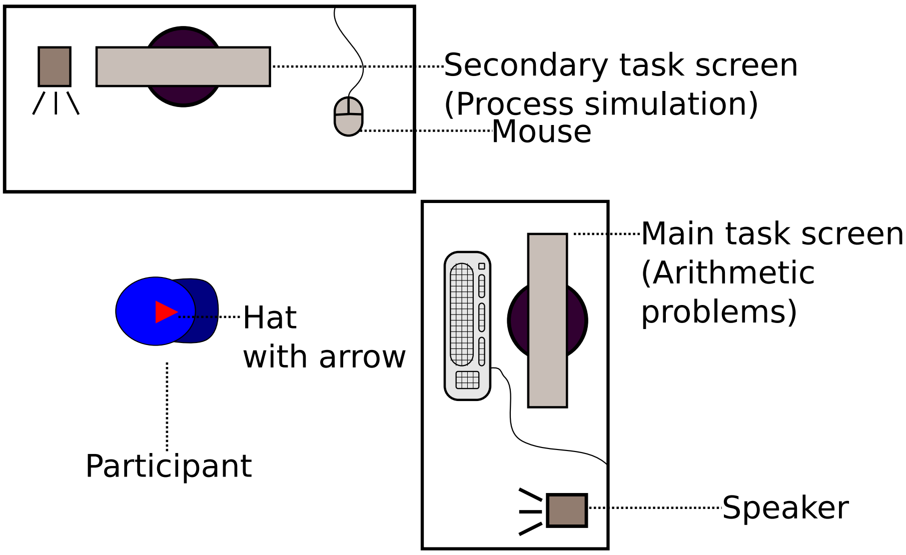


Figure 5.2.: Schematic overview of the SoProMon system setup.

adequate complexity, yet low enough to offer preliminary insights into processes of limited complexity. Even though, as already motivated before, the current implementation of SoProMon is based on the simulation of a production process, the system can be adapted to simulate different types of processes as well. The simulation has been designed in such a way as to require several actions by the users, in order to measure the performance of auditory monitoring in attention allocation and in interrupting the users during their main task. Furthermore, as a delayed or non-existing response of the users to certain situations can influence the performance of the simulated process, this provides an elegant way to measure the users' monitoring performance. The required interactions are:

Supply: One of the machines requires the user to refill the resource input of this machine in more or less regular intervals. In order to simulate a realistic environment, each machine contains a random factor that influences the time between when an input resource has been taken and the resulting material resource has been produced.

Empty: One of the machines requires the user to clear the output buffer (i.e. initiate a delivery/transport of goods) to free up space for assembled goods to be buffered.

Maintain: Two machines can encounter conditions of malfunction stops, which require active attention of the operator. In our case, these situations could theoretically be anticipated, as they occur regularly every pre-defined number of steps. However, this number is so large that it is rather surprising/interrupting for regular work on the primary task. For resolving, the user must click on the 'maintain' button, an action that could be modified to represent a more realistic scenario in which an operator would need to go elsewhere in the monitoring room to resolve a problem.

In order to perform an adequate number of user tests without unduly exhausting users, we planned a single evaluation time to be 10 minutes. Thus, we designed the simulation such that it provides an adequate number of interruptions within the evaluation time. In order to discourage users from performing unnecessary actions (e.g. prophylactically clicking "supply" every once in a while), we introduced a small waiting period after an action has been performed, in which the respective machine is paused. This is a realistic behavior, as real machines often require a short downtime when they are being refilled, and a longer one if they are being maintained or repaired.

As to (b) the visual monitoring, we depict a graph of the machine setup and flow of goods. The visualization can be assumed to be checked very quickly, leaving any time for the interpretation and reaction to be accounted for the interpretation of sounds. As our main concern is the assessment of sonifications that complement an existing visual monitoring console, the visual part remains invariant in all experimental conditions and thus does not require a dedicated motivation or testing.

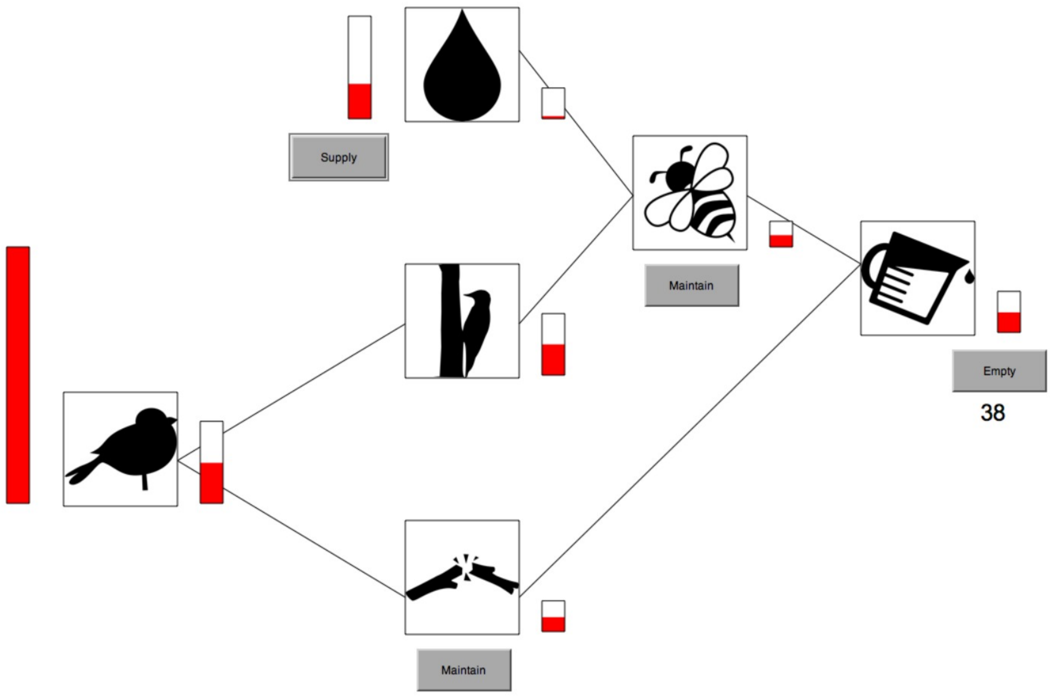


Figure 5.3.: SoProMon visualizations: filling levels are depicted in red, two machines include maintain buttons, the other buttons are for the management of buffers (supply/empty).

While Fig. 5.3 shows the *normal* state of simulation, in which all machines are working, the Fig. 5.4 shows a critical state in which several machines are out of order, due to either empty resource levels or maintenance needs.

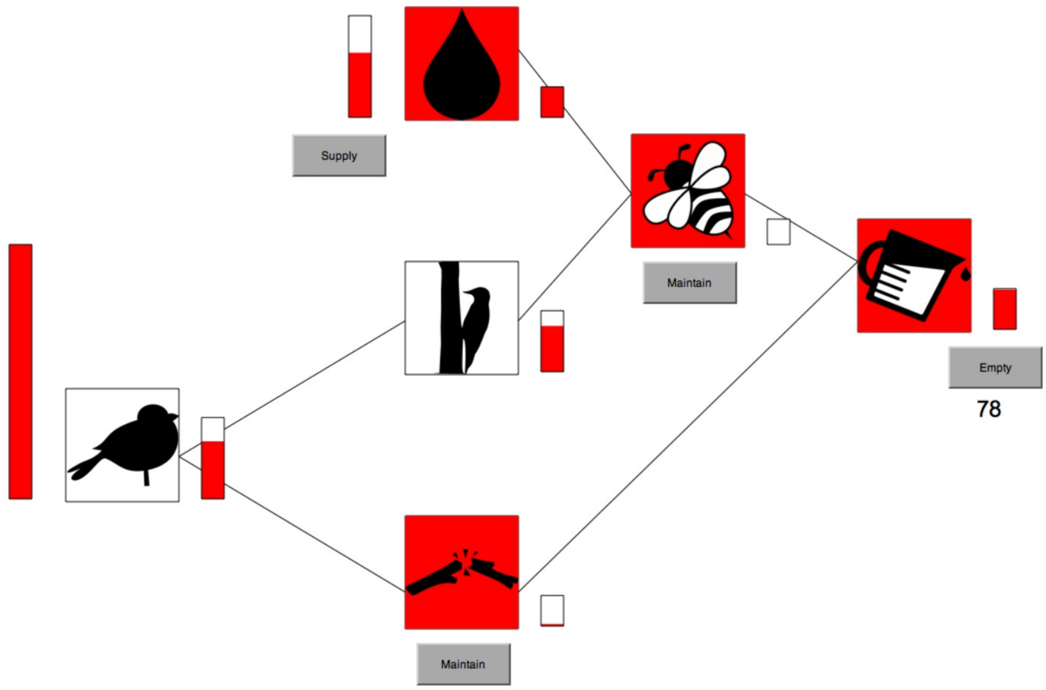


Figure 5.4.: Four out of the six machines have reached a critical state, respectively depicted by a red filling. The user can resolve these states by performing the respective actions.

5.1.1. Main Task

As the focus was to evaluate sonifications for process monitoring as a second task, it was first necessary to find suitable, yet realistic, primary (main) and secondary (monitoring) tasks. In a real-life setting, users interrupt their main tasks in more or less frequent intervals to observe the status of production. If necessary, these users then perform certain process-related actions such as ordering new raw material or repairing a broken machine.

We identified the following requirements for the main task: (a) it should be rather complex and cognitively demanding, (b) performance should be measurable, specifically task completion time and correctness of the result, ideally in an automated way, (c) the task should be repeated, rather frequently and with short duration, so that we can observe performance effects on small time scales, finally (d) it should be possible to interrupt the task in favor of attending to the processes to be monitored without consequences (i.e. the main task would then just wait). While requirements (a) and (d) aim to design the evaluation in a way that it resembles the real-life working conditions of supervisors as much as possible, (b) and (c) are needed in order to measure the performance and distraction of the test subjects in an effective way.

Typical tasks in real-world scenarios may be, depending on the user group, documenting, processing documents such as emails or postal inquiries, or planning/scheduling. These are rather domain-specific and heterogeneous in type. We argue that, for the sake of binding the cognitive resources, much simpler tasks are already effective to allocate and bind the users' full attention. For the sake of simplicity, we selected the *adding numbers* task, which is a simple mental arithmetic task of summing two numbers (each smaller than 50). The result is to be entered into a text field using the computer keyboard. On hitting the return key, the task, result and timestamp are logged and the next random numbers are drawn and presented (see Fig. 5.5). The task is displayed in the center of the screen, which is otherwise empty to reduce distractions.

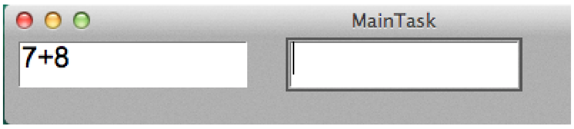


Figure 5.5.: Main task window of SoProMon system.

5.1.2. Sonification Requirements

Extending on the general requirements for auditory process monitoring defined in Chapter 4, the sonification design aims to meet the following requirements, specified for the tasks at hand:

(a) The sonification should provide an awareness of the ongoing process steps, allowing the listener to recognize significant changes in the overall process state.
(b) The sonification should represent information on the underlying process in a continuous manner, i.e. dimensions of the sound vary in tight connection with underlying data, so that the user can infer and anticipate states.
(c) The sonification should be compatible with verbal interaction, i.e. it should enable engagement in verbal conversations with limited/controlled disturbance.
(d) The sonification should be unobtrusive and support participants in letting it perceptually disappear in the periphery of attention, similar to how we can ignore car engine sounds while driving – yet surprising changes shall draw the attention to the sound.
(e) It should be possible to discern the sounds from different machines, allowing users to learn and associate distinct sounds with particular machines.
(f) The sound should be compatible with the acoustic environment in which the process monitoring situation is embedded.

The requirements for the sonification design result from combining our analysis of productive application opportunities with our understanding of the possibilities of sonification and our defined goals. We started with a first design, which we call *basic process identification sonification* and progress further to our latest design, *process data driven soundscapes*, which we explain in detail.

5.1.3. Basic Process Identification Sonification

A basic assumption is that each machine execution is an elementary (atomic) step that transforms an input into an output situation. The execution can be conceptualized as an event that takes place at a certain time. It seemed straightforward to represent these elementary events to corresponding acoustic events, leading to the perception of sound streams for each machine of rather repeated sound events. Thus, the sound aggregates have both the characteristics of identifiable events and a continuous stream. The rate at which the machines operate can be perceived by the rate of sound events, allowing an intuitive association: the higher the number of sounds per second, the higher the execution speed (in the same manner as a Geiger counter where radioactivity corresponds to the number of tick sounds).

To distinguish different machines, it is necessary to create machine-specific sounds that are easily discernible, yet coherent in their structure so that they can be perceived as belonging to the same auditory display. As the human listening system already performs a frequency analysis (by means of the tonotopic organization of the cochlea), a frequency analysis, we started with short percussive tones, tuned to a set of pitches so that tones can be easily discerned. Specifically, we use a complex timbre synthesized from a source filter model with 4 resonant filters at the fundamental f_0, second harmonic, $3.4 \cdot f_0$, and the seventh harmonic. There is no strong argument behind this timbre vector, yet it sounds pleasant and a bit like a glass/metal object. Still, in the selection of pitches lies a huge design space, as these tones can be set as either equidistant on a pitch perception scale, resulting in equal musical intervals between tones (e.g. a third, or fifth), or selected deliberately to form a coherent musical chord according to a harmonic system. As a first starting point, we chose 8 semi-tones between sound streams. However, as pitch is such a salient feature, it would be underused for merely identifying the machine. Instead we slightly increase the pitch depending on the degree to which the machine output buffer reaches the maximum. Perceptually, this compares to the increasing of pitch in sound when filling a bottle with water, indicating in an analogue and intuitively understood manner that 'something runs full'. Specifically, the pitch range spans 7 semi-tones for the range of semi-full to full output buffer.

A critical point is that a cacophony of parallel playing sounds makes it difficult to attend to the relevant sound streams. We thus decided to map the output buffer filling level to the sound level as well, so that empty output buffers (i.e. no problem) result in quiet sound events. Specifically, the sound level increases by 21 dB as the output level increases from 75% to 100%. To enable listeners to anticipate that input buffers run empty, we furthermore add a noise as the transient phase of the sound, mapping emptiness of input buffers to increasing noise levels. Thus, the more noisy the sounds are, the more critical the input situations become. The increase of noisiness becomes thus more and more discernible as buffers run empty. As the sound events have only a single sound level, we use this variable as an indicator of any kind of criticality: individual mappings for the cases 'filling output buffers', 'emptying of input buffers' and 'machine failure' have a mapping to amplitude, so that the more critical the situation becomes, the higher the amplitude becomes. Finally, the maximum of the three amplitudes is used as overall sound event amplitude. In consequence, sound events corresponding to events where no problem is apparent become quiet, and depending on settings almost inaudible. In case of machine failure, the machine sound is repeated at loud volume and low rate. This is immediately understood as an alarm condition, grabbing attention immediately. Our design has been iteratively refined by the authors while carefully balancing parameters to achieve subjectively acceptable sound streams for a sustained monitoring situation. However, the applied earcons are neither particularly well designed, nor do they meet the acceptability threshold for extended (e.g. full day) use. We concluded that more *natural* sounds, as frequently encountered on a daily basis when interacting with real-world objects, would be attractive and yield better compatibility with long-term use.

5.1.4. Process-data-driven Soundscapes

Since the previous approach appeared promising in terms of information richness and interpretability, yet suboptimal in terms of acceptability of the sonic texture, we mostly worked on this issue within the design. Our starting point was to question the way coherence between acoustic representations of machines is established within the display. Interpreting the process as a 'virtual world', a closed ecology which is distinct from our natural acoustic ecology, we see the opportunity to design the virtual acoustic environment based on the same guiding principles that structure natural soundscapes. For instance, in most ecosystems, animal sounds are optimized so that animal voices allocate disjunct spectra, thus reducing misclassifications. One design seed in this direction was to associate different bird motifs with machines, and thus have the running process yield 'the soundscape of birds', which is mostly regarded as relaxing and as a nice ambience. However, in the situation of our machine simulation, where events occur at a rather constant rate, the sounds are too frequent for complex bird motifs, and using bird chirps soon became exceedingly annoying. Looking at soundscapes in general, we see that most everyday sounds encountered while interacting with objects are transient, percussive yet stochastic. These sounds can include footsteps, closing a door, opening a bottle, etc.

We defined the soundscape with a *forest theme*, creating a soundscape with positive connotations and selected sound events from this area that work well, such as a cracking branch, a bee, a woodpecker, rustling leaves, water drops, and a snippet of a brook sound. While soundscape ecology would suggest an optimization process so that the bandwidth allocation reduces the risk of masking, we select only sounds based on subjective choices to fit into this theme. The mappings are largely identical to the ones described in the basic identification sonification above. This design, which emphasizes non-pitched sounds, is to our experience much less obtrusive, and much more capable to fulfill requirements (c), (d), (e) and (f). Since playback rate (and thus implicit pitch) is not as salient as pitch in the other design, (b) is slightly less clear, yet this sacrifice is probably acceptable given the improvements in the other categories. Of course, these observations represent our subjective impressions and require formal testing in empirical studies, as will follow in Sec 5.2. Based on the concept of *Cognitive Infocommunication Channels* (Csapo and Baranyi, 2012), this approach applies both Low-Level Direct Mappings (as all production steps directly result in respective parameterized auditory icons) and Structural Mappings, as for instance the mapping of filling buffer levels to pitch is an analogic representation.

The sonification is true to the event-like nature of individual machine executions, meaning that every process step in the production yields a tiny sonic counterpart so that their superimposition creates a soundscape that reflects the overall activity. Assigning different sounds/timbres to different machines results in 6 voices that play simultaneously.

When a machine has reached a critical level, its sound is repeated at a fast rate and high volume, until the problem is solved. Aside from the fact that a production step has been executed, the following additional data is conveyed aurally:

- We map the output buffer filling level to pitch, thus making use of the analogy of a filling jug.
- We map the input buffer criticality to increasingly louder noise to the initial/transient phase of the sound, thus enabling listeners to anticipate that input buffers run empty. The increase of noisiness becomes more and more discernible as buffers slowly run empty.
- We generally map the approaching of critical conditions to audio level, resulting in machine sounds gradually becoming louder and thus more salient over the 'normal' background soundscape, as the situation worsens. Specifically, the sound level increases by 21 dB, starting when the buffer level is at 25 % (machine with 'supply' button)/75 % (machine with 'empty' button), or when the condition of a machine has reached 25 % (machines requiring maintenance).

In general, a continuous soundscape sonification is unlikely to reach its full potential in a production process as simplified as the one we built, as here it would not be difficult to recognize or anticipate all critical situations, as well as to define respective thresholds for when a warning sound should be conveyed. Therefore, for the simplified process at hand, a continuous sonification possibly offers fewer advantages over an alert-based sonification than it would for more complex processes (although, obviously, our sonification concept is not suitable for a high number of machines and would have to be adjusted). However, the process used in this experiment should instead be seen as a very simplified representation of a much more complex real-life production process that includes more machines and variables to be monitored, and for which it would be significantly more difficult to model all potentially critical situations and define the respective thresholds. This is also the reason for the decision to sonify all machines, not only those that require interactions: real life processes are typically more complex than our scenario. In such a process, it could be beneficial to sonify even machines for which no interaction is necessary or possible. This could include, for example, the sonification of irregularities in the production of machines that precede other machines for which interactions are required, in order to help anticipate problems that might occur for those machines at a later stage.

However, it is obvious that the approach taken is not indefinitely scalable – in a scenario with a hundred machines, one cannot assign a unique sound to each machine. In such a scenario, one would have to either concentrate on sonifying KPIs (Key Performance Indicators) such as, for example, throughput times, or perhaps let the user dynamically select a handful of machines to be sonified. For the production process used in this experiment, however, a sonification of individual machines is feasible.

5.1.5. Technical Design

The SoProMon system is intended as a standalone component to conduct experiments. From a technical standpoint, it consists of the following components:

- The SoProMon experiment system, which logs experiment data.
- The analysis toolchain to parse and analyze experiment data.

The SoProMon system has been programmed in SuperCollider, a platform for audio synthesis and algorithmic composition. Although primarily intended to be used by computer musicians, SuperCollider has gained popularity among sonification researchers and practitioners. It combines the object-oriented structure of Smalltalk with aspects from functional programming with a C-family syntax. The reason for deciding on SuperCollider was mainly its extensive functionality of high performance sound synthesis as well as the possibility to create GUIs. Apart from SuperCollider, a few other programming languages and platforms specifically geared towards audio and sound synthesis exist; mainly Pure Data[1] and Chuck[2]. Furthermore, libraries for sound synthesis exist for almost all major programming languages.

The implementation has been developed in a loosely coupled manner, so that all the functionality connected to a specific machine can be passed. The sonification implementation is provided as a dynamic function that is executed for every process step, as is the code executed in case a machine fails.

[1] https://puredata.info/
[2] http://chuck.cs.princeton.edu/

The SoProMon system is intended to be run on two screens - one for the main task, and another one for the process simulation. A calibration and experiment control screen (See Fig. 5.6) opens when starting the system, enabling the experimenter to adjust the volumes of the different machines in cooperation with the participant. After the experiment has started, this window is closed automatically.

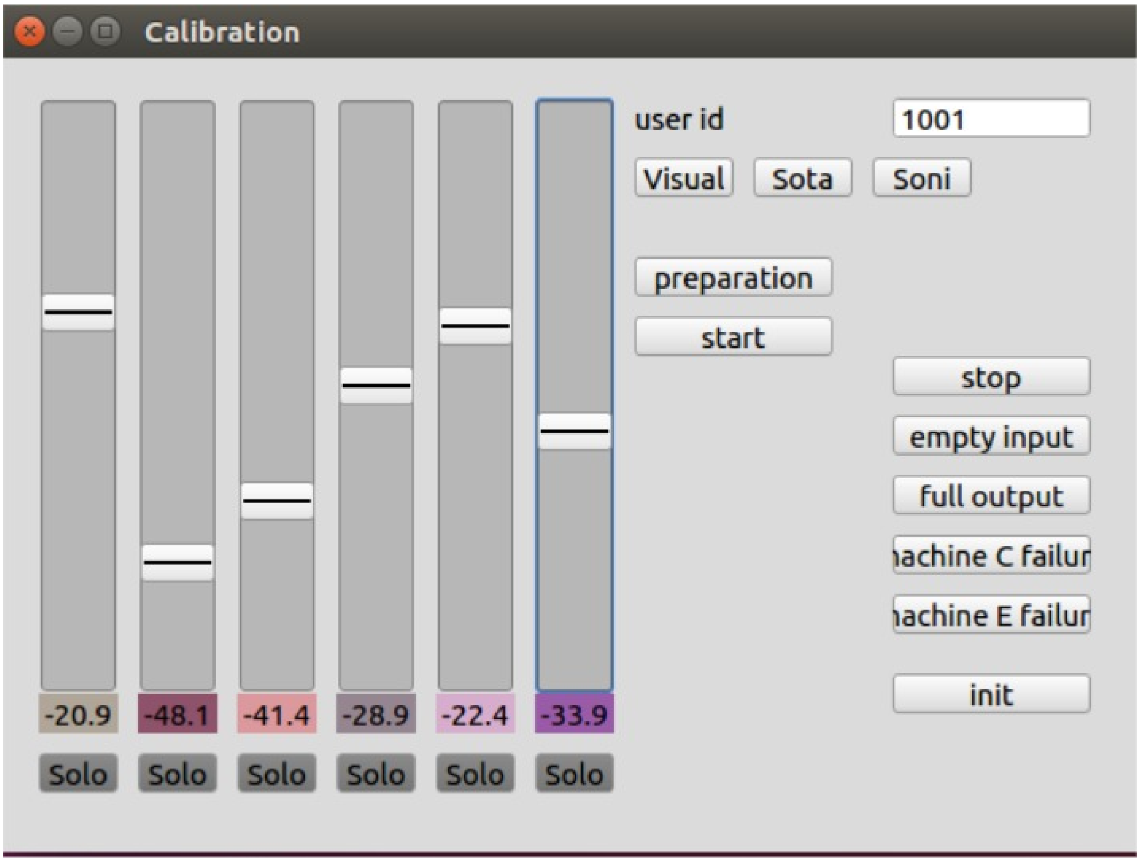


Figure 5.6.: Calibration window of the SoProMon system.

The second component is the data analysis pipeline, which consists of the steps ETL (Extract, Transform, Load), statistical analysis and creation of results/figures. The detailed description of the data analysis can be found in Sec. 5.2.9.

5.1.6. Implementation

To demonstrate the sonifications, video and audio recordings are available on the SoProMon Website[3], and as a data publication at the Bielefeld University website[4].

The file "Video of Son"
sopromon_son.mp4 presents a recording of the process simulation in the

[3] Project website: http://cs.univie.ac.at/wst/research/projects/project/infproj/1063/
[4] Bielefeld University data publication: http://doi.org/10.4119/unibi/2904377

Process-data-driven Soundscape mode. We suggest first to watch the video with muted audio so that the 'silent' visual-only monitoring can be understood. For this purpose the audio-only recording could also be used ("Audio of Son") SoProMon.mp3. As discussed earlier, all detailed information is rendered in the visual display, including the emptying of buffers (see time code: 0'50") or the anticipation of machine failures (see time code: 1'20"). Now, please unmute the sound and look and experience how sound augments the perception of the process. In general, the audio volume is adjusted by the user before starting the simulation, according to the hardware and the users' hearing capabilities. In the video example, the sound has been recorded with a high volume, so it is recommended to adjust it to a low volume which is above just audible. See also whether you can recognize which sound belongs to which machine. After 1-2 times watching the video, we recommend looking away from the screen and trying to identify from listening alone, when it is time to attend to the process. We learned from informal tests with a few subjects that it takes a short time to become familiar with the sonification and how it represents the information as sound.

For comparison, we provide the sonification example S2[5], which presents the first design *Basic Process Identification Sonification*, using the pitched sounds to identify machines. Again, from a handful of test listeners so far, we received the feedback that this is much more obtrusive than the process-data-driven soundscape design. In Appendix E, spectrograms of the sounds selected for the 6 different machines (in normal states, as well as in critical state) can be found.

The file "Video of Sota" SoProMon_sota.mp4 presents a recording of the process simulation in the Sota-mode (State-of-the-Art), in which auditory alerts are only conveyed in case a critical situation has already occurred.

Figure 5.7 shows a visualization of log data of selected buffer levels as well as the condition of a machine. The data have been recorded during a pre-test experiment. The three different types of user actions ('supply', 'deliver', 'restart') are clearly visible.

Figures (5.8, 5.9, 5.11 and 5.10) show photographs of an experimental setup of the SoProMon system at the laboratory of the Ambient Intelligence Group at CITEC/Bielefeld University in December 2014. Fig. 5.8 presents the setup with the top-camera visible. Fig. 5.9 demonstrates the positioning of the participants and the head-mounted device, used for head tracking, while Fig. 5.10 presents an additional side-camera. Fig. 5.11 shows a photograph of the view from the top-mounted camera.

[5] see http://doi.org/10.4119/unibi/2695709

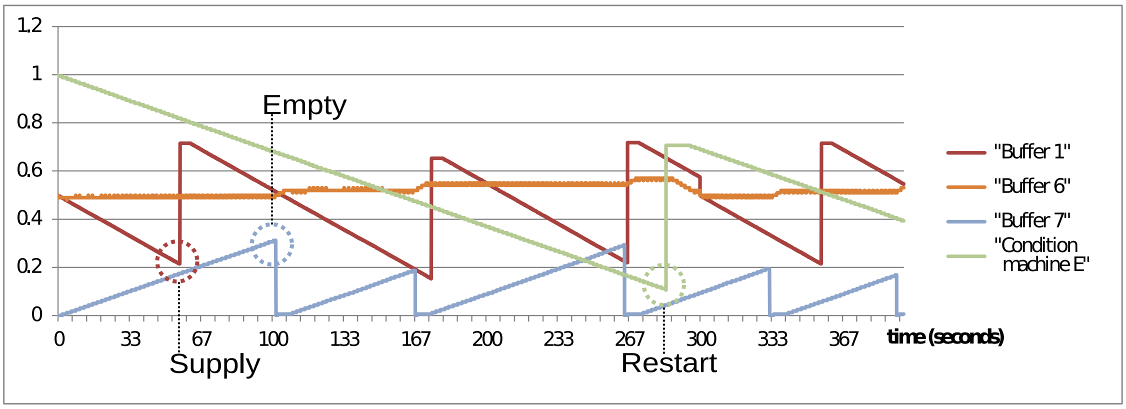


Figure 5.7.: Visualization of experimental log data.

Figure 5.8.: Photograph of SoProMon setup with top camera marked.

5.2. Design of the SoProMon Experiment

The performances of main and secondary task were compared under three conditions in order to find out if there are differences between the conditions in terms of distraction from the main task, and if a continuous sonification allows improvement of process monitoring performances over the two other conditions,

Figure 5.9.: Photograph of SoProMon system to visualize the positioning of the participants and the head marker.

Figure 5.10.: Photograph of SoProMon setup with side camera.

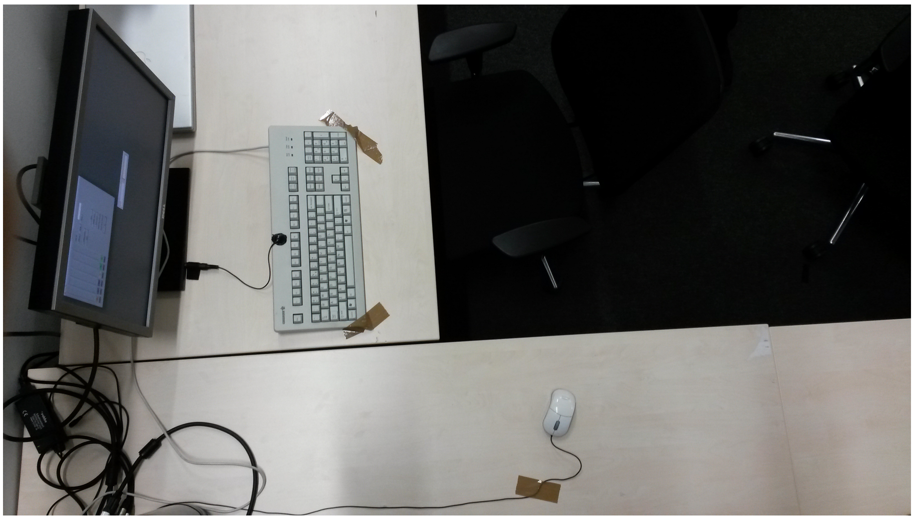


Figure 5.11.: Photograph of SoProMon setup with view from top camera.

for example by better avoiding critical situations. Furthermore, we wanted to find out, which condition potential users would find most helpful and pleasing.

5.2.1. Hypothesis

Based on the related research in Chapter 3, we define several hypotheses. As a baseline, we take the two most common modes of monitoring in current production scenarios (and of many other domains as well):

- $C_{\mathbf{vis}}$, in which the process status is conveyed using only visual means.

- $C_{\mathbf{sota}}$, that combines C_{vis} with auditory alarms.

These two are compared to

- $C_{\mathbf{son}}$, combining the other two conditions with a continuous sonification.

In general, we expect differences between C_{vis} and C_{sota}, but between C_{son} and the other two in particular, basically in favor of the auditory information types. We expect the differences to manifest both in the users' perceptions (measured by their questionnaire responses and their comments) and the quantitative performance measured in the monitoring task. We expect, in accordance with the literature, that the three mentioned conditions will have no significant effect on the performance of the main task (H1.1). Concerning process monitoring performance and behavior, we expect that additional auditory cues have no effect (H2.1). We do, however, expect continuous sonification to have a significantly positive effect on monitoring performance, compared to the other two conditions (H2.2), as this is the main aim of the sonification design. Concerning the questionnaire responses, we expect items associated with helpfulness, attention switching and performance increase of the respective mode of operation to be more favorably rated in C_{son} compared to C_{vis} and C_{sota} (H3.1). We furthermore believe that users feel more self-assured and in control with C_{son} (H3.2). In contrast, we expect additional information conveyed aurally to increase exhaustion, and therefore C_{sota} and C_{son} to be more exhaustive than C_{vis} (H3.3). For each of these hypotheses, the null hypothesis (H_0) says that there are no significant differences between the groups, while the alternative hypotheses (H_1) assumes the opposite.

5.2.2. Data Recording

In order to answer our research questions, different types of data were recorded during the experiment:

- Main task logs, containing the arithmetic problems and the entered solutions with timestamps.
- Process simulation logs, containing the timestamp, the performed type of interaction, and the current buffer levels and machine status'.
- Questionnaires and verbal comments, encompassing participant demographics, experiences in relevant ares, and satisfaction with respective modes of monitoring.
- Audio and video recordings of the entire experiments from two cameras (from the top and from the side), in order to capture verbal comments and allow for head tracking.

5.2.3. Experiment Plan

The first pre-test based on thinking-aloud (Lewis and Rieman, 1993) was conducted off-site before the experiment to ensure the understandability of the system, the questionnaires, the two tasks and the provided instructions, the validity of the experiment, and the technical functionality of all components of the SoProMon system, especially the logging mechanisms. As a result, we adapted the questionnaire, the experiment procedure and the textual instructions. Furthermore, slight adjustments to the technical implementation were made, mainly in terms of logging and functionalities for the experimenter. Before the experiment, two more pre-tests, this time using the final experiment setup, were conducted. After analyzing the log data, the verbal feedback and video recordings, a few final adjustments were made to the experiment setup and procedure. In order to have three complete sets of condition sequences in this within-subjects design, 18 participants were recruited for the experiment itself. All 6 possible permutations of the three conditions were realized among the runs.

Before signing an informed consent, the participants were give a written introduction into the goals and aims of the study, its duration and experiment procedure.

Next, the participants were handed written instructions for which questions could be asked at any time. The instructions were placed next to the participant during the whole experiment for potential look-ups. The instructions specified the goal of the participants, namely to solve as many arithmetic problems at possible during the three 10-minute experiment segments, while concurrently trying to avoid critical process states as well as possible. The participants furthermore received an explanation of the two tasks and their user interfaces. The experiment goal was also repeated verbally shortly before the experiment.

In turn, the pre-questionnaire containing items on demographic data, visual or auditory impairments and experiences (e.g. in musical or sound design) were handed out. Before starting the process simulation with one of the three conditions (C_{vis}, C_{sota}, C_{son}), the participants were informed regarding how the process status and criticality was conveyed (e.g. no sound, sound in the case or errors, permanent sounds). If the respective part of the experiment was the participant's first, he or she had time to become familiar with the process simulation beforehand, and ask questions. If the participants' first condition was C_{son}, there was an additional sound level calibration phase, during which the individual machine's volumes were adjusted until they were just loud enough to identify. After 10 minutes, the main and secondary task windows closed automatically. The post-condition questionnaire was then distributed, asking for opinions and impressions on the previously presented modes of process operation. These included several items that allowed for comparisons between the three conditions, like the question of whether the mode of process monitoring was exhausting or pleasing, and items that are only related to a specific mode, for example those that concern the understandability of the different mappings from data to sound in C_{son}. The complete questionnaire is available in Appendix A. After a participant completed all three experiment parts, a post-experiment questionnaire was distributed, asking for comparisons of the three presented conditions, furthermore including condition-independent items such as asking for overall feedback on the experiment. In total, the experiment lasted around 65 ± 5 minutes for each participant.

5.2.4. Study Population

The study population had an age median of 26±7.1 years. Table 5.1 shows the experiences of the participants in various relevant fields, as well as their visual and hearing impairments.

Two subjects had slight visual impairments, while three subjects responded that the statement "do you have visual impairments" applied: one of those subjects had no vision in one eye and no impairment in the other. The other 13 subjects had no visual impairments. Most subjects had no or rather little experience with process monitoring; only one had extensive experiences. The majority of the subjects had no or little experience with musical instruments, sound creation, audio editing or programming, or sonification.

Table 5.1.: Experiences and impairments. Exp=Experience with

Item	Not at all	Rather not	Rather	Fully applies
Hearing impairments	17	1	0	0
Visual impairments	13	2	3	0
Exp. process simulation	10	7	0	1
Exp. musical instrument	8	4	3	3
Exp. audio	6	7	4	1
Exp. sonification	8	4	5	1

Table 5.2 shows the participants' opinions and estimations concerning the foundations of the experiment, such as the graphical user interface.

Almost all subjects stated that they understood the foundations of the process simulation (median: 10.0 ± 0.6) as well as the possibilities of interacting with the GUI of the process simulation (supply, maintenance, emptying) (10 ± 1.0). In general, participants did not feel overwhelmed by the process simulation (1.0 ± 2.1) or the main task (1.0 ± 1.7). They felt that the graphical layout of elements of the visual representation of the process simulation was clear and understandable (8.0 ± 2.4) and that understood the visualization principles (10.0 ± 1.4).

However, there was feedback on specific elements of the graphical user interface and the visualization, for example that the visualization of impeding critical states (depicted as color transition) could have clearer or slower, that it was distracting, or that its onset was in some cases not in sync with the acoustic mappings in C_{son}. One participant suggested to visualize only those aspects that are necessary to decide whether a certain intervention is necessary or not (e.g. to not visualize input and output buffers for machines that only offer the maintenance interaction). There was also feedback on the general conduction of the experiment; for example, that the keyboard used for the main task was too stiff or too far away. Two participants criticized that they had to press a key in order to switch the focus of attention to the main task window. In recognition of this possible error source we corrected in post-processing the calculation mistakes that were (supposedly) due to the first digit after an action being ignored. Furthermore, it was mentioned that the mouse positioning favored left-handed users.

Table 5.2.: Opinions on foundations of experiment.

Item	*Median* $\pm$ *IQR*
Understood foundations of process simulation	10.0 ± 0.6
Understood interaction possibilities	10 ± 1.0
Felt overwhelmed by process simulation	1.0 ± 2.1
Felt overwhelmed by main task	1.0 ± 1.7
Visual representation was clear and understandable	8.0 ± 2.4
Understood GUI & process visualization	10.0 ± 1.4

5.2.5. Main Task Measurement

Different measures can be used to compare the performances. The number of calculations that the participants were able to execute in the given experiment time of 10 minutes would be the most obvious choice. On the other hand, not only the number of calculations is relevant, but also their correctness. Therefore, the conditions are also compared concerning their mean *deviation*, which is calculated by averaging the deviation of each entered result from the correct result by

$$\text{deviation}^{\alpha} = \frac{1}{N^{\alpha}} \sum_{i=1}^{N^{\alpha}} \left| 1 - \frac{r_i^{\alpha}}{\hat{r}_i^{\alpha}} \right| ,$$

where α is the participant number, N^{α} the number of questions to which participant α has replied, and r_i (resp. $\hat{r}_i$) are the given (resp. correct) results for the ith question. Furthermore, in order to compare the overall *performance* of the participants between the three conditions, we introduce the 'main task score', a variable that encompasses both the number of solved calculations and their correctness as

$$\text{Main Task Score}^{\alpha} = \frac{N^{\alpha} - \langle N \rangle}{\sigma(N)} - \frac{dev^{\alpha} - \langle dev \rangle}{\sigma(dev)},$$

with $\langle \cdot \rangle$ referring to the sample mean and $\sigma(\cdot)$ being the standard deviation of its variable.

5.2.6. Process Simulation Measurement

An obvious metric to measure the monitoring performance between the different conditions would be to compare how many clicks the users made on average for each condition. Also of interest are the buffer values of the respective buffers at the time of the user's interaction with the simulation (e.g., the input buffer of a certain machine at the time of refilling it). A relatively high average buffer value can, for example, signify that the users do not trust that the respective mode of process monitoring conveys the need for interaction in time, leading the users to switch their attention to the process simulation in regular intervals, and performing interactions *just in case*. A low average buffer, meanwhile, can signify that the users rely on the respective conditions' ability to signal interaction needs. On the other hand, if, for example, an input buffer had already been completely depleted at the time of intervention, this may signify that the respective condition has failed to inform the users in time. In many cases, participants used double clicks for their interactions, while a single click would have been sufficient, a fact that was perhaps not communicated clearly enough to the participants. Therefore, if several clicks were performed directly one after another, only the first click was taken into account.

Anticipation Optimal Rationality

In order to find a measure that enables the evaluation of individual user interactions in terms of adequateness and timeliness, we introduced the *anticipation optimal rationality*. The concept of *anticipation optimal rationality* tries to be true to a real-life production scenario: interactions that come *too early*, and are therefore *unnecessary*, are here punished. This is realistic, as in most real-life production scenarios it is a goal to maximize production and to minimize downtimes as much as possible. As each interaction in the simulation entails a downtime (e.g. for maintaining a machine), it is logical to minimize those interactions. However, the anticipation optimality not only punishes interactions that are too early but also such that are *too late*. *Too late* in this case means that a critical situation had, at the time of interaction, already occurred, or that it was so close, that it could not have been prevented assuming average reaction times. The rationale behind this is that downtimes due to critical states (such as a machine breakdown) are to be avoided even more than 'planned downtimes' due to maintenance.

Thus, in anticipation optimal rationality R_A, we calculate the mean time that participants needed to shift the attention from the main task to the process simulation. This time span included finishing the calculation that they were working on, turning around to the process simulation, and performing the necessary action. If an interaction was performed when the respective machine had already stopped due to a critical state, or when, for example, the buffer value was already so low that an interaction was likely to be performed too late, this interaction was evaluated with a score of 0. Interactions performed exactly at a time at which it would have been, given average reaction times, possible to intervene just before a critical state would have been reached, were rated optimal (1.0). All interactions that occurred at a later point were mapped linearly between 1.0 and 0 from the optimal interaction point (1.0) to the point at which an interaction has been completely *unnecessary* (0), for example when an input buffer was still completely full (see Fig. 5.12).

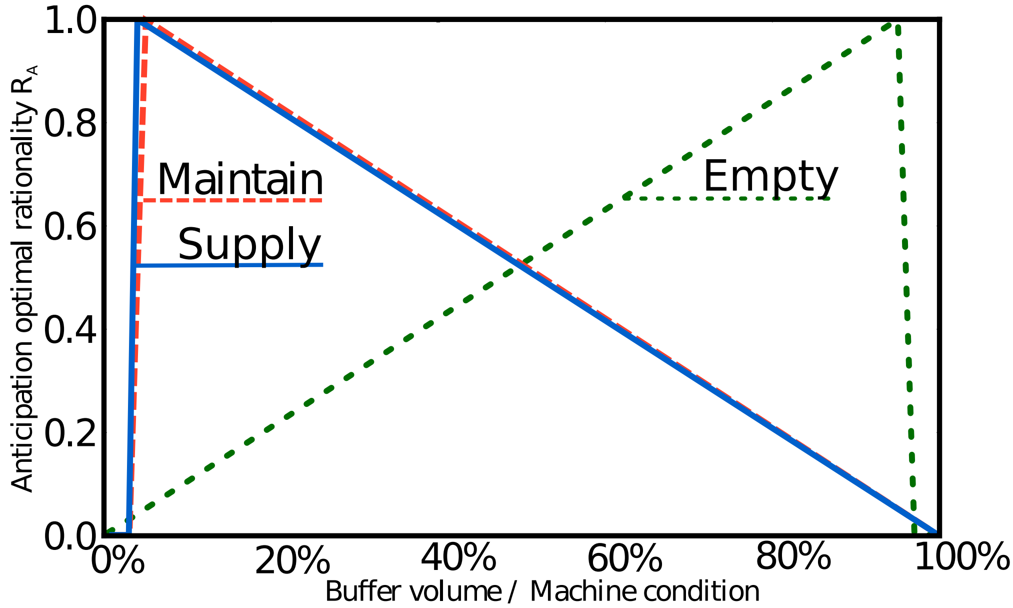


Figure 5.12.: The figure shows for each of the three interaction types - supply (blue, solid line), empty (green, dotted line) and maintain (red, dashed line) - how an interaction at a specific point in time, measured in terms of the current buffer volume or condition of a specific machine (displayed on the x-axis), would be rated concerning the concept of anticipation optimal rationality (displayed on the y-axis).

5.2.7. Preprocessing of Questionnaire Data

In order to test subjects' accuracy in answering and to facilitate detection of random answering, several items of the questionnaire covered the same subject, often on an opposing scale. If the given answers were too contradictory, the respective item pair for the specific participant were marked as outlier/inconsistent data point and removed for the analysis. This way, out of the 1908 individual answers[6], 13 such pairs were removed (e.g. "I was always in full control of the process simulation" vs. "I was overstrained by the process simulation"). As several analyses required complete sets of answers (such as the combination of different items), the removed answers were filled with the mean of the respective item.

In order to allow a more powerful analysis and more representative results, corresponding items have been combined into Likert scales (composite scores).

[6] corresponds to 106 questionnaire items for 18 subjects

The factors have been grouped by average. Subsequently, their consistency has been tested with Cronbach's alpha: only those factors with a reliable consistency outcome (i.e. ≥ 0.8) have been used for data analysis. Unless stated otherwise, the scales range from 10 (fully applies) to 0 (does not apply at all). The complete questionnaire (translated from German) can be found in Appendix A. Questionnaire results concerning individual conditions can be found in Appendix B, and questionnaire results concerning comparisons between conditions in Appendix C. Further detailed results can be found in the online version of the supplementary material[7].

5.2.8. Statistics

All results concerning main task performance have the following format: mean $\pm$ standard deviation, as they are interval-scaled and normally distributed. For the other results (process monitoring performances and questionnaire results), the standard way of communicating results is median $\pm$ IQR (interquartile range), due to the fact that those results have been tested to be not normally distributed, which is why mean values would have little significance. An exception is the discussion of the Likert scales toward the end of the questionnaire results. These scales have been aggregated out of different Likert items and tested for normality distribution.

For scores for which several data points for one experiment part existed (e.g. the buffer values when refilling a machine), for each of the 54 experiment parts (18 subjects x 3 conditions), a mean value has been calculated and compared between the three modes using Friedman tests for repeated measures, respectively repeated measures ANOVA (for the main task results). The reason aggregating using the mean, instead of the median, even though for some variables only a handful of data points may exist, was to avoid losing the information of *outliers* and variance. This information is essential because outliers (e.g. standstills of machines) are crucial in process monitoring and should be avoided when possible. Individual comparisons between two conditions were performed using Wilcoxon signed-rank tests for dependent samples, respectively dependent t-tests (main task results).

For all modes, the false discovery rate has been adjusted with a Benjamini-Hochberg correction. All p-values have been divided by two, if a directed hypothesis has been defined beforehand. The same is true for presented correlations. The results of correlation tests, if not stated otherwise, combine the *r value* of Spearman's Rho with the p-value of Kendall's Tau, as it is more accurate for smaller samples.

[7]https://doi.org/10.4119/UNIBI/2904377

5.2.9. Analysis Pipeline

The data is analyzed using the analysis pipeline described in Sec. 5.1.5. As the first step, the data is loaded, parsed and merged from the different log files using Python and the libraries NumPy and Pandas. For each experiment, three types of logs and data sets are created:

- Main task data: for each experiment condition a subject participates in, a log file with the solved algorithmic problems and their timestamp is created.
- Process simulation data: for each experiment condition a subject participates in, two log files are created. The buffer log file contains entries that show the levels of the different machine resource stacks throughout the experiment. The interaction log shows all the interactions that the participant performed during the experiment with the corresponding timestamps.
- The questionnaire data shows the results of the participant concerning the different questionnaires.

These files are loaded and correlated with each other (see figure 5.13).

At first, the data is preprocessed as described before. The different statistics for the three types of data sets are calculated as described, using mainly the Python packages NumPy and SciPy, but also custom Python code. Graphs are created to visualize the results, using the library Matplotlib but also custom created plots and data visualizations. The calculated statistics as well as the graphs are exported, and some of those exported data are imported into the data mining software RapidMiner[8] for further analysis, mainly for the calculation of aggregated likert scales, where again CSV files and graphs are exported.

5.3. Results of the SoProMon Experiment

This section describes the most interesting and significant experiment results. A summary of the participants' performances in main and second task can be found in Appendix D.

5.3.1. Main Task Results

The number of performed calculations for C_{vis} is slightly higher (mean: 131.22 ± 31.87) than for C_{sota} (127.0 ± 31.43) and C_{son} (124.94 ± 28.48), see Fig. 5.14. However, the differences are not significant (p>0.388). The highest deviations of the results of the arithmetic problems from the correct solution were observed under C_{vis} (0.0098 ± 0.0080), compared to C_{son} (0.0079 ± 0.0051) and C_{sota}

[8] https://RapidMiner.com/

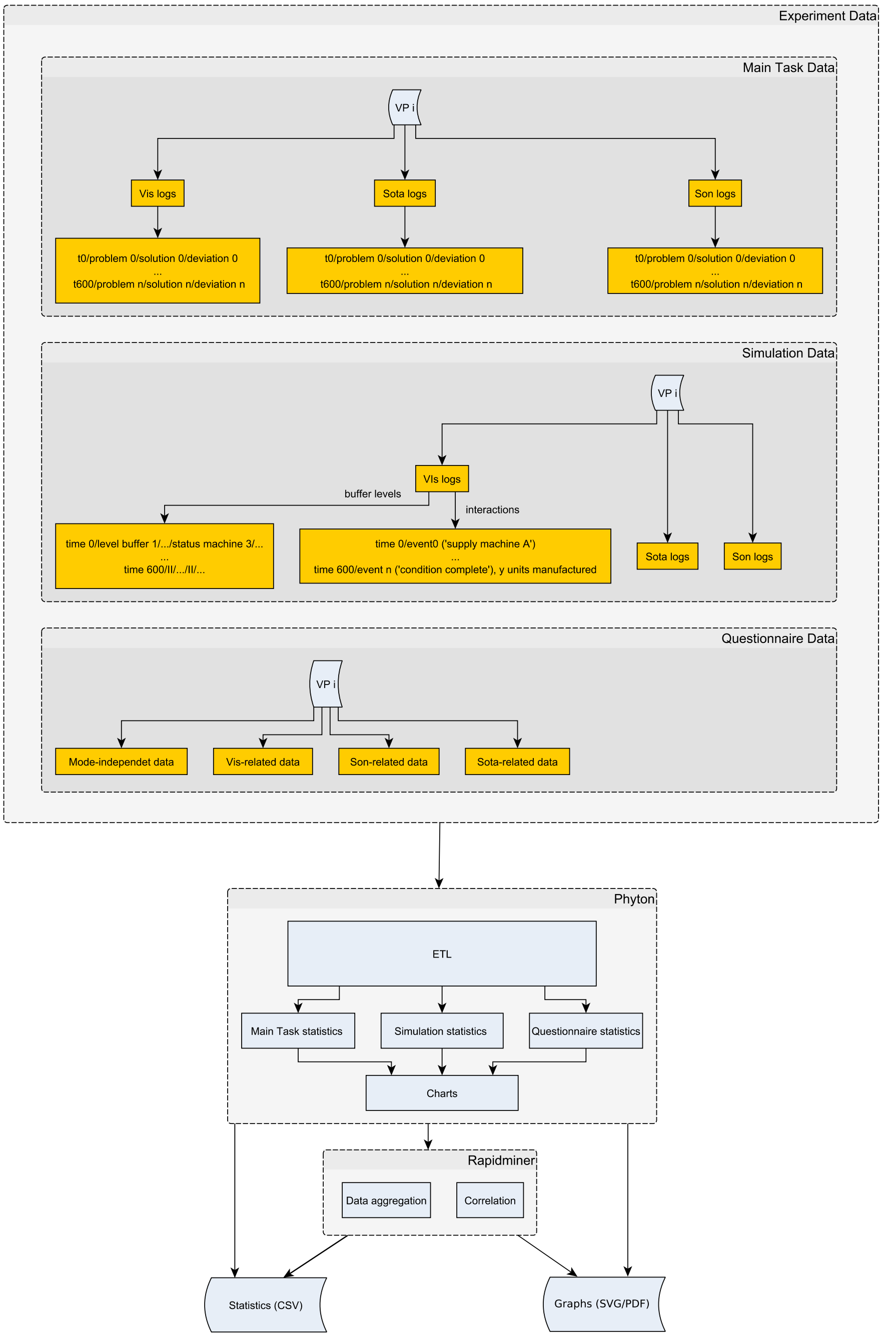


Figure 5.13.: Data analysis pipeline and log data structure of SoProMon system.

(0.0070 ± 0.0078). However, again the differences were not significant (p>0.393, see Fig. 5.15).

Concerning the overall main task score, participants achieved the highest results during C_{sota} (0.1416 ± 1.622), lower scores under C_{son} (-0.0399 ± 1.336), and the lowest scores under C_{vis} (-0.1017 ± 1.655), see Fig. 5.16. The differences are not significant.

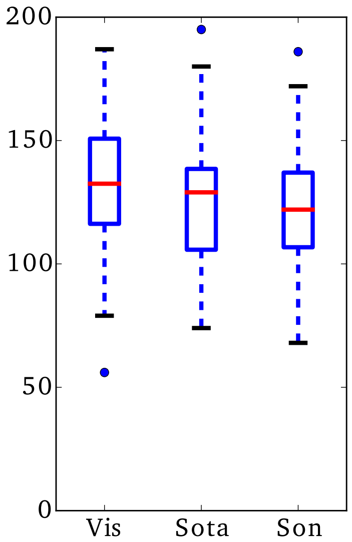


Figure 5.14.: Numbers of solved calculations.

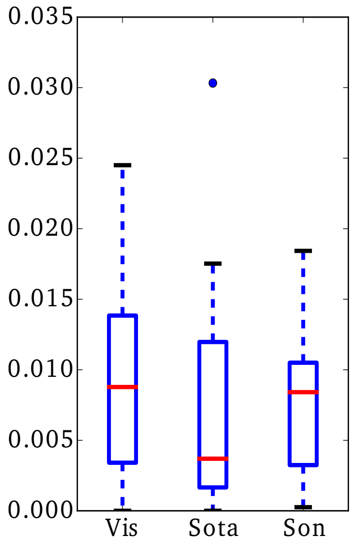


Figure 5.15.: Deviation from correct solution.

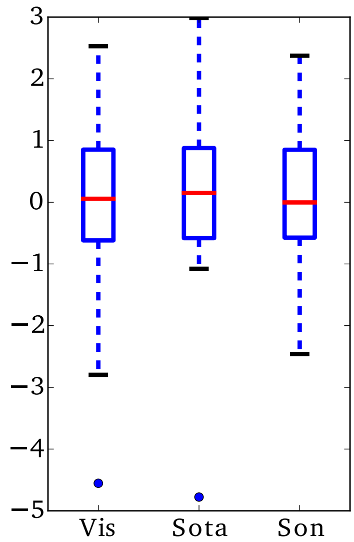


Figure 5.16.: Main task scores, encompassing both numbers of solved calculations and their correctness.

5.3.2. Process Monitoring Results

In total, 1,367 interactions (supply, empty, maintain) were performed, including 483 during the C_{vis} conditions (median number of interactions per participant in C_{vis}: 23.0 ± 7.5), 448 during C_{sota} (23.5 ± 2.75) and 436 in C_{son} (23.5 ± 2.75), see Fig. 5.17. Although these differences are not significant, the number of interactions seems to decrease with the amount of information conveyed aurally. The results of C_{vis} and C_{sota} also deviate more than those of C_{son}.

Analysis of Buffer Values

The median buffer value at the time of clicking 'supply' over all conditions was 0.20 ± 0.184, or at 20% buffer fill level. A substantial number of interactions were performed when the buffer was already depleted.

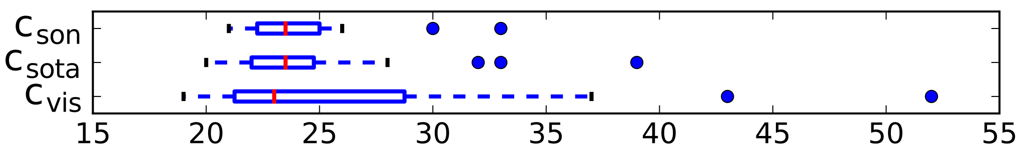


Figure 5.17.: Median number of mouse clicks per experiment part measured for the monitoring task.

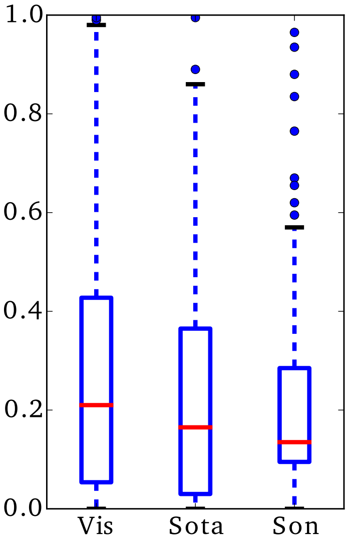


Figure 5.18.: Input buffer levels when clicking 'supply'.

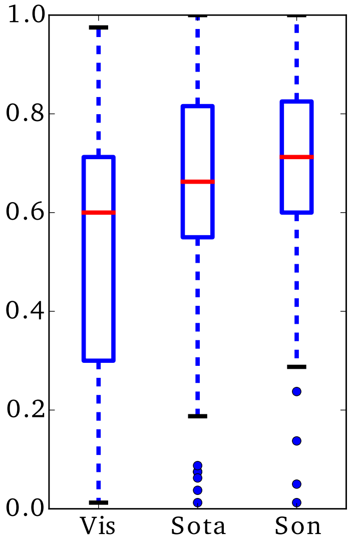


Figure 5.19.: Input buffer levels when clicking 'empty'.

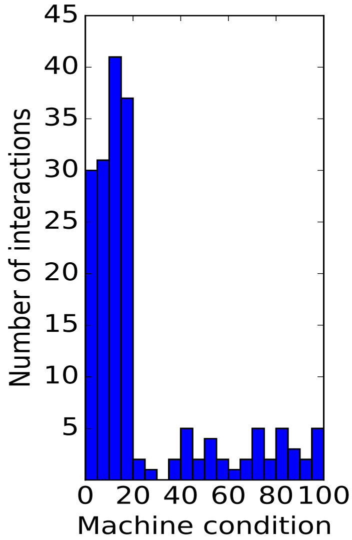


Figure 5.20.: Histogram of machine condition when maintaining machine C.

The aggregated median value when clicking 'supply' was highest during C_{vis} (0.246 ± 0.253), lower in C_{sota} (0.207 ± 0.218), and lowest in C_{son} (0.166 ± 0.127), see Fig. 5.18. The differences, however, are not statistically significant. The number of supplies performed when the respective input buffer was already completely depleted was substantially higher under C_{vis} (29 out of 216 supplies = 13.4 %) and C_{sota} (34 / 199 = 17.1 %) than under C_{son} (2 / 197 = 1 %).

The median buffer value when clicking 'empty' was 0.677 ± 0.139, or 67.7%. Few participants waited until the output buffer was completely full to empty it, but quite a few emptied it when the buffer was still relatively empty. The aggregated median value for clicking 'empty' was lowest under C_{vis} (0.676 ± 0.118), higher under C_{sota} (0.6825 ± 0.133) and highest under C_{son} (0.680 ± 0.104). The differences between the conditions are, however, not significant. Fig. 5.19 shows all buffer levels at the time of emptying.

The median condition of machine C when maintaining it was $11.526\% \pm 7.361$. Fig. 5.20 suggests that most participants intervened only when the machine was about to break down, while many even waited until the machine had stopped. Under C_{vis}, the aggregated median condition of machine C at maintaining was 11.131 ± 7.071, 13.343 ± 12.426 under C_{sota} and 11.526 ± 6.868 under C_{son}. The differences between the conditions are not significant. More participants reacted only after a machine had already stopped under C_{vis} (4 interactions when the machine had already stopped out of 65 total interactions) and C_{sota} (3/61) than under C_{son} (1/56).

Significant differences between the conditions were observed for the maintenance interaction of machine E ($p < 0.038$). The aggregated median machine condition when maintaining it was 18.588 ± 16.925 %. The median condition at maintenance was 14.204 ± 9.315 under C_{vis}, 17.224 ± 12.966 under C_{sota} and 23.754 ± 12.260 under C_{son}. However, there were no significant differences between C_{vis} andC_{sota} ($p > 0.127$), C_{vis} and C_{son} ($p > 0.0789$) or C_{sota} and C_{son} ($p > 0.249$) observed. In C_{vis}, 7 out of 83 maintenances were initiated when the machine had already stopped producing, compared to 6 out of 96 in C_{sota}, and none in C_{son}.

Anticipation Optimal Rationality

The median aggregated anticipation optimal rationality score, independent of the condition, was 0.737 ± 0.477 (see Fig. 5.21). The median anticipation optimal rationality score, aggregated over all conditions, was 0.670 ± 0.209. There are highly significant differences ($p < 0.001$) between C_{vis} (0.632 ± 0.175), C_{sota} (0.578 ± 0.111) and C_{son} (0.787 ± 0.074). The results under C_{son} deviate less than under the other conditions (see Fig. 5.22), something that has been observed for many other metrics in this experiment as well.

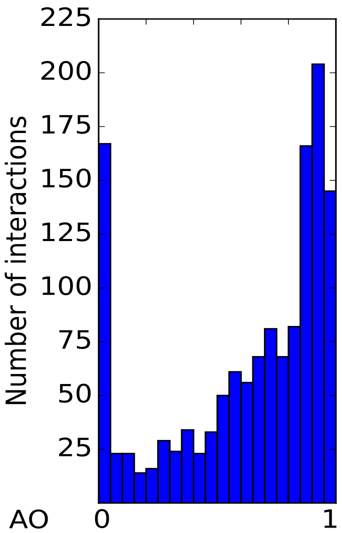


Figure 5.21.: All interactions with respective anticipation optimal rationalities.

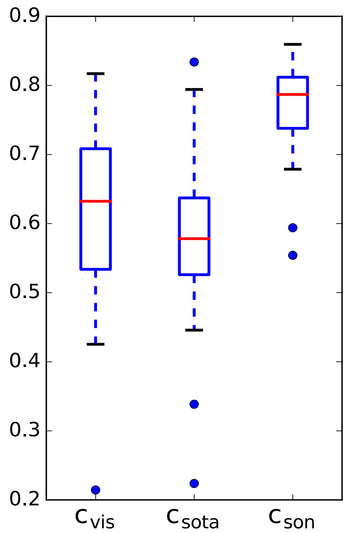


Figure 5.22.: Median anticipation optimal rationalities of the experiment parts.

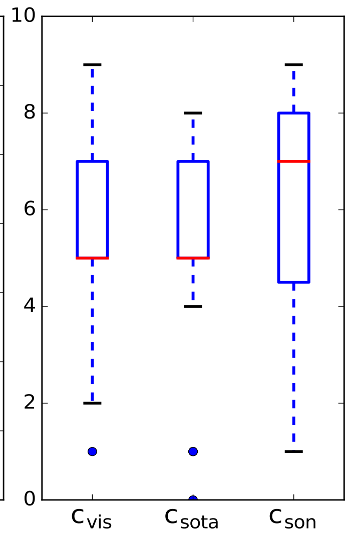


Figure 5.23.: Monitoring task performance improvement.

The difference between C_{vis} and C_{sota} is not significant ($p > 0.25$), but between C_{vis} and C_{son} ($p < 0.001$) as well as between C_{sota} and C_{son} ($p < 0.001$) significant differences have been observed. As can be seen from Fig. 5.24, the difference between C_{son} and C_{vis} and C_{sota} was higher when these conditions were not the participants' respective first condition.

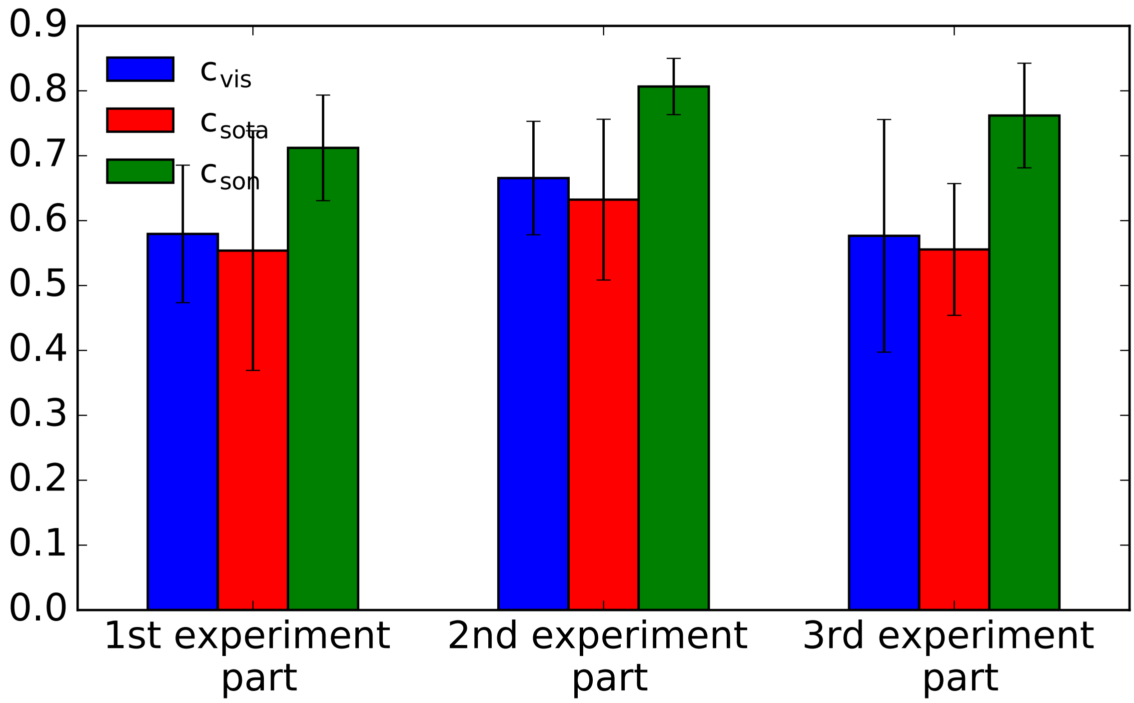


Figure 5.24.: Aggregated mean anticipation optimal rationality R_A, depending on if the respective condition was the participants first (left), second (middle) or third (right) condition of the experiment. The three sub figures compare the results of C_{vis} (left), C_{sota} (middle) and C_{son} (right).

5.3.3. Questionnaire Results

There seems to be a clear preference for C_{son} concerning the responses to those items that compare the three conditions and for which significant differences between the conditions were observed:

- Item "I have improved my performance in the process simulation over time." ($p<0.009$, Fig. 5.23):
  C_{son} $(7.0 \pm 3.5) > C_{\text{vis}}$ (5.0 ± 2.0) $p<0.048$
  C_{son} $(7.0 \pm 3.5) > C_{\text{sota}}$ (5.0 ± 2.0) $p<0.048$
- Item "I have improved my performance in the main task over time." ($p<0.040$)
  C_{son} $(5.5 \pm 4.0) > C_{\text{sota}}$ (4.0 ± 2.0) $p<0.040$
- Item "I have been informed in time about potential problems during process simulation." ($p<0.001$, Fig.5.25):
  C_{son} $(9.0 \pm 4.5) > C_{\text{vis}}$ (2.5 ± 4.75) $p<0.002$
  C_{son} $(9.0 \pm 4.5) > C_{\text{sota}}$ (3.0 ± 4.5) $p<0.003$
- Item "How helpful were the different modes of process monitoring that have been presented in the respective parts of the experiment?" ($p<0.002$, Fig. 5.26):
  C_{son} $(8.5 \pm 1.0) > C_{\text{vis}}$ (6.0 ± 3.5) $p<0.018$
  C_{sota} $(7.5 \pm 1.0) > C_{\text{vis}}$ (6.0 ± 3.5) $p<0.028$

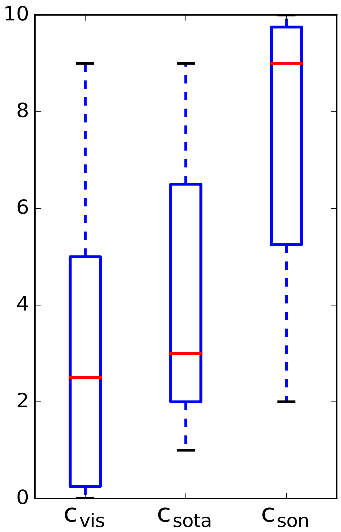


Figure 5.25.: Timely information about problems in process simulation.

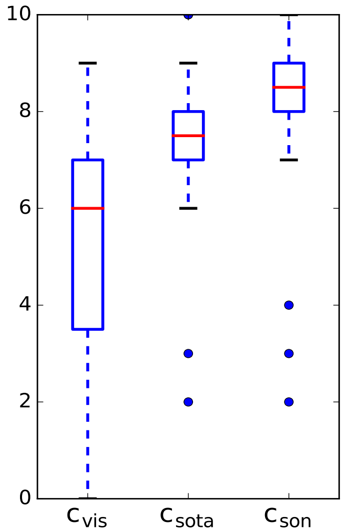


Figure 5.26.: Perceived helpfulness of C_{son}.

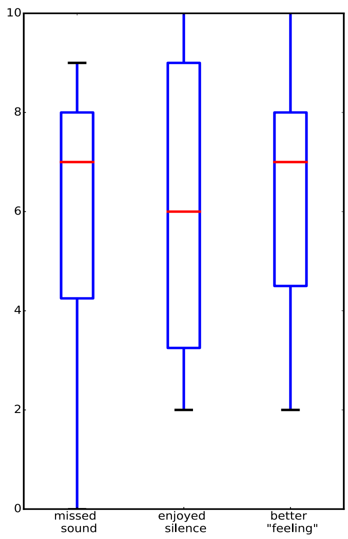


Figure 5.27.: Results for questions on self assurance versus obtrusiveness of C_{son}.

Participants who started their experiment with a condition that included sound, when asked if the sounds in later C_{sota} or C_{son} modes helped them to obtain a better *feeling* for the process, in general agreed (7.0± 3.5). Fig. 5.27 depicts more statistical details.

It is interesting that the mean of items related to the intrusiveness of C_{son} (see Fig. 5.28) is slightly higher (5.0± 2.75) than of those associated with the sound design being pleasing (4.0± 2.0). However, the feedback related to information aspects was on average quite positive (7.0± 1.75). Almost all participants stated that they believe sounds can generally be helpful for process monitoring (9.0± 2.0). Perhaps not surprisingly, participants indicated significantly more often ($p<0.009$) that they oriented themselves at the sounds in C_{son} (8.5 ± 3.0) than in C_{sota} (3.5 ± 4.0). Furthermore, almost all participants stated that they were able to differentiate (8.0± 2.5) and assign the various machine sounds (7.5± 1.0). Furthermore, there is a very strong correlation (Pearson: 0.900, $p<0.001$) between the understanding of the mappings from data to sound applied in C_{son}, and the belief that the sonification of C_{son} is helpful. Not surprisingly, there is also a very strong correlation between the understanding of the C_{son} mappings and the perceived informativeness of the sounds of C_{son} ($r=0.884$, $p<0.001$). Fig. 5.29 shows the participants' assessment concerning how easy it was to perceive the different aspects of the sound design of C_{son} and its mappings.

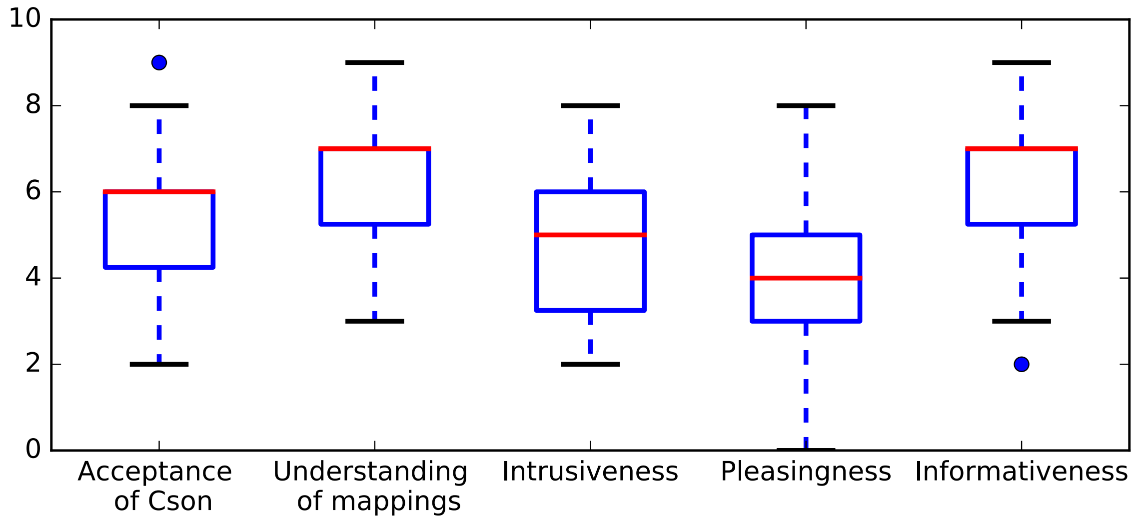


Figure 5.28.: Results of Likert scales related to the sound design of C_{son}.

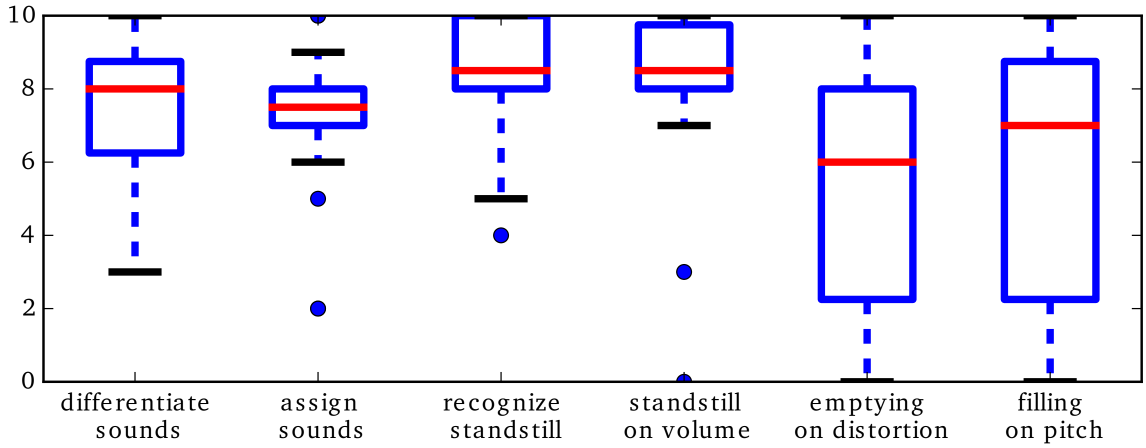


Figure 5.29.: Results for questions concerning the 'ease of perception' of the various aspects and mappings of C_{son}.

5.4. Experiment Discussion

5.4.1. Main Task Results

It seems that participants generally solved slightly fewer calculations with increasing monitoring data conveyed multi- modally (from C_{vis} to C_{sota} to C_{son}), even though the calculation results in C_{vis} were slightly less accurate than in C_{sota} and C_{son}. When taking both the number of solved calculations and their correctness into account, the highest results can be observed in C_{sota}, followed by C_{son} and C_{vis}. As no statistical differences between the three conditions were observed, neither in terms of the number of solved calculations, nor their correctness, nor the overall main task score, the null hypothesis of H1.1 can be accepted. Fig. 5.30 depicts the average main task scores under the three conditions, depending on whether the respective condition was the first, second or third part of the experiment for the respective participant.

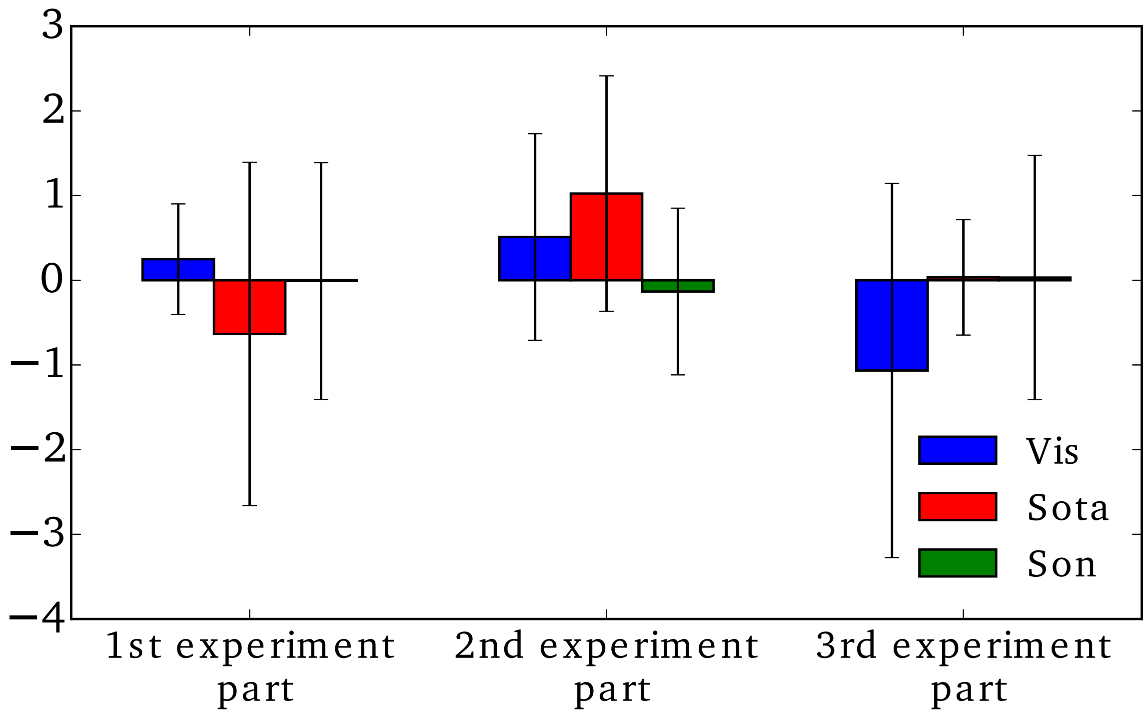


Figure 5.30.: Main task scores, depending on if the respective condition was the subject's first (left), second (middle) or third part (right) of the experiment. The three sub figures compare the results of C_{vis} (left), C_{sota} (middle) and C_{son} (right).

There are some tendencies that can be observed, for example, that when the participants performed monitoring under C_{vis} as their last experiment part, significantly lower main task performances were achieved compared to when it was their first or second experiment part. This could be caused by the fact that the participants had to shift their attention between the two tasks more often, as the process status was not conveyed aurally as well. Since the data shows that

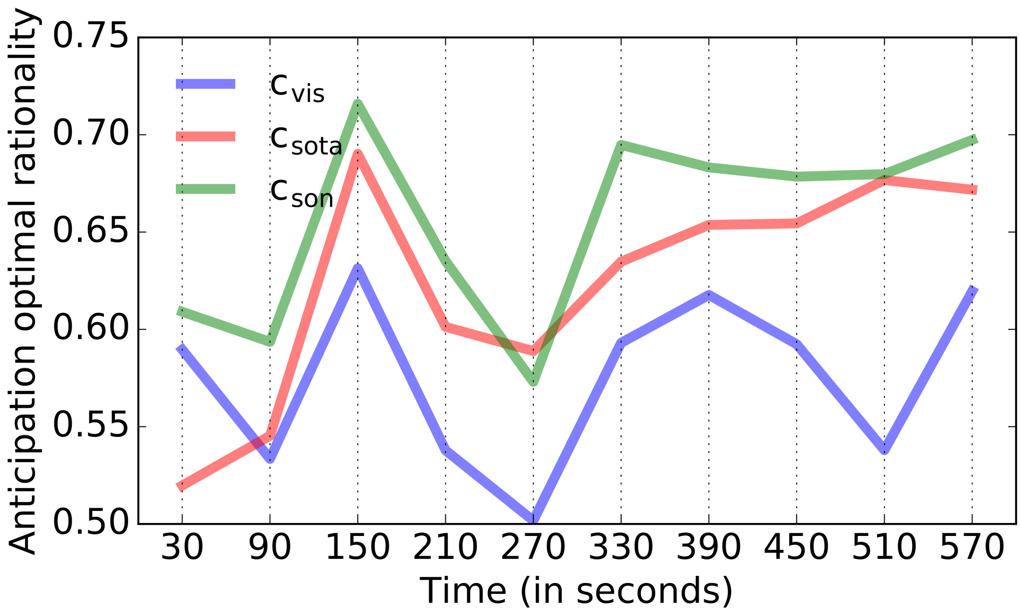


Figure 5.31.: Mean anticipation optimal rationalities over the course of the experiments' time in 10 time slots. Blue=C_{vis}, red=C_{sota}, green=C_{son}.

C_{son} achieves its best result as the respective subjects' third mode of operation, it can potentially be assumed that C_{son} would especially benefit from a habituation effect.

5.4.2. Process Monitoring Results

Participant performance in the process monitoring task was significantly higher under C_{son} than under C_{vis} and C_{sota}, while performance under C_{sota} was on average not significantly different than that under C_{vis}. Thus, the null hypothesis of H2.1 can be accepted, while it can be rejected for H2.2. In tendency, substantially fewer interactions were performed under C_{son} compared to C_{vis} and C_{sota}, when a machine had already reached a critical state.

In general, the results of C_{son} seem to posses a lower variability than those of C_{sota} and C_{vis} concerning almost all aspects, as the results typically deviate less and contain fewer outliers. An explanation for this might be that participants in C_{son} interacted less often *too late*, thus avoiding critical states more often, but also less often *too early*.

Fig. 5.31 reinforces these observations by showing the average R_A values for the three conditions within 10 time slots. What can be seen is that the participants achieved higher R_A values in C_{vis} and C_{son} compared to C_{sota} in the beginning. After a short while, the R_A values for C_{sota} increase significantly and are on average higher than those of C_{vis} for the rest of the experiment. The highest scores throughout the experiment, except for a short period during the middle of the experiment, can be observed for C_{son}. In general, all three conditions show

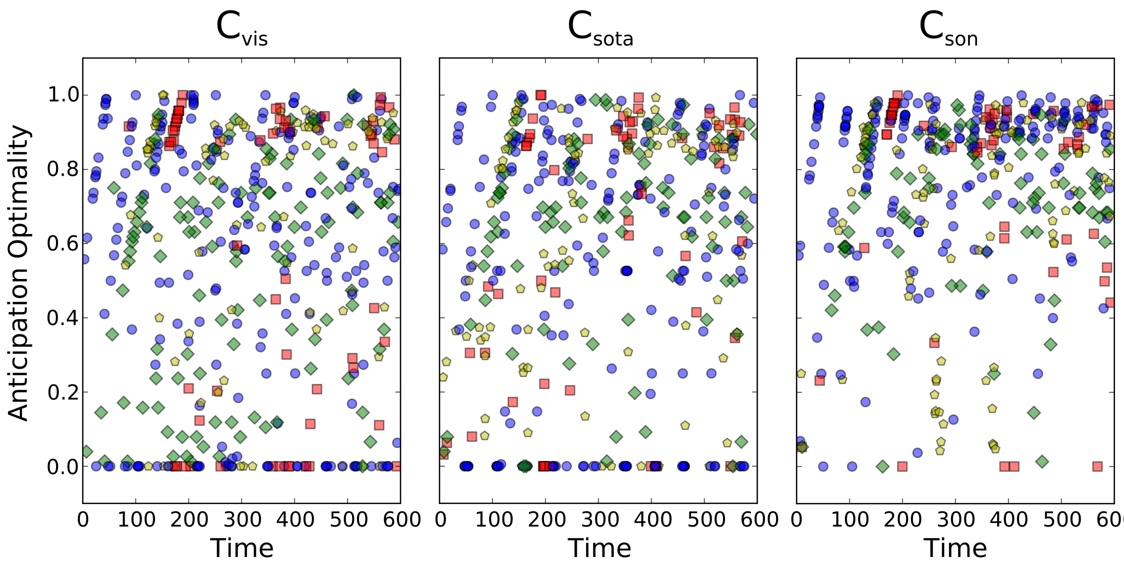


Figure 5.32.: All interactions of all participants during C_{vis}(left), C_{sota}(middle) and C_{son}(right) throughout the experiment time (x-axis), evaluated according to anticipation optimal rationality (y-axis). Supplies are symbolized with a blue circle, emptyings with a green diamond, maintenances of machine C with a red square, and maintenances of machine E with a yellow pentagon.

the same ups and downs throughout the experiment. However, towards the end of the experiment, a sudden drop in the R_A values of C_{vis} has been observed. Independent of these ups and downs, in tendency the R_A values of C_{sota} and C_{son} seem to steadily improve throughout the experiment until the end, while the R_A values for C_{vis} remain more or less constant.

Fig. 5.32 shows all user interactions performed by all participants for the conditions C_{vis}, C_{sota} and C_{son}. The different event types are highlighted by different symbols and colors, as explained in the caption. The plots for C_{vis} and C_{sota} look fairly similar, with events clustered more or less equally along the y-axes, however with larger clusterings at the lower end of the chart, indicating interactions that have been evaluated with an R_A of zero. In C_{son}, on the other hand, there are almost no symbols at the bottom, but more towards the upper part, indicating higher R_A values. Furthermore, the C_{son} chart looks more organized and *tidy*. This is likely because of two factors: (a), in C_{son} there have been fewer interactions in general, and (b) the events of the different types seem more grouped together, indicating that an interaction of the same type has been executed more often at around the same time in the experiment, and with a similar R_A value. This is partly due to the fact that as the different interactions are required at around the same intervals within the experiment (apart from a small random factor), interactions were more often performed at the 'optimal time'.

5.4.3. Questionnaire Results

C_{sota} was considered significantly more helpful than C_{vis}, as was C_{son}, thus in this regard the null hypothesis of H3.1 can be rejected, although C_{son} was not considered to be significantly more helpful than C_{sota}. There were no significant differences in terms of attention switching.

Although there was also no significant difference in terms of estimated process-monitoring-performance increase between C_{vis} and C_{sota}, C_{son} scored significantly higher than both. Concerning H3.2, the null hypothesis cannot be rejected, as participants did not feel significantly more in control during C_{son} (8.0± 5.5), than during C_{vis} (5.5±5.0) or C_{sota} (4.0±4.75). As no statistically significant differences in terms of perceived exhaustion could be observed, the null hypothesis of H3.3 also cannot be rejected. However, several participants stated that *it would be better not to use sounds that become continuously stronger, but to combine the problem sounds with additional warning sounds (e.g. when the input stack is at only 10%)*. Another participant stated that he found the continuous sonification very exhausting and found the silence in the later C_{vis} pleasing. Two other participants stated that they would have preferred auditory-cue-based sonifications over continuous sonifications. All participants that commented negatively on C_{son} further indicated in their questionnaire responses that they found C_{son} to be more intrusive, less pleasing and less euphonious than the median participant (one subject even stated that he would *go crazy* listening to our sound design for a longer period of time.).

On the other hand, one participant noted after C_{sota} that *due to the fear of the appalling sounds, one automatically tries to observe the graphical representation more, thus it feels like the sounds distract more from the actual task*. Another participant mentioned that

> "this mode (C_{sota}) was very shocking. In the mode with all sounds (C_{son}) it was a bit problematic to differentiate all sounds, but in principle this mode is the most pleasing one, as there is a smooth transition from what you are actually doing, to the monitoring."

The same participant stated that in C_{vis} one would probably make the most mistakes unless one would have the second screen within one's line of vision, while this participant could imagine using C_{son} for a longer period of time, although he/she would still have to test this. Verbal comments of two participants further suggest that their performance in C_{son} would most likely increase over time and the intrusiveness of the sounds would decrease. The experiment data further suggests that there is room for improvement concerning the sound selection and mapping strategies.

Three subjects stated that the two machines using water-based sounds were quite difficult to distinguish. Another participant suggested that some of the sounds were too similar, especially in terms of their frequency spectrum. Furthermore, it was stated by one subject that the breaking stick was harder to perceive than the other machines. Almost all participants were able to understand the mapping of urgency on volume (8.5$\pm$ 1.75), while most (however with a large deviation among participants) stated that they were able to understand the mapping on distortion (6.0$\pm$ 5.75) and rising pitch (7.0$\pm$ 6.5). Two subjects stated that they had not been aware of those two mappings (even though they were explained in the introductory text). Such improvements might further increase both monitoring performance and the participants' acceptance.

There seems to be a trade-off between helpfulness and exhaustion – with C_{son} being considered the most exhaustive (6.0 $\pm$ 2.0), but also the most helpful condition, and C_{sota} being less exhausting (4.0 $\pm$ 4.0), but also less helpful. C_{vis} (5.0$\pm$ 4.0) seems to be considered more exhausting than C_{sota} while at the same time less helpful. Participants found significantly ($p<0.007$) that C_{son} (4.0 ± 5.75) required a longer training period than C_{sota} (0.5 ± 2.75).

Table 5.3 provides an overview of the hypotheses described in Sec. 5.2.1, and their answers.

Table 5.3.: Hypotheses of Sec. 5.2.1 and their answers.

Hyp	H_0/H_1	Result
1.1	H_0	No differences concerning main task performance
2.1	H_0	No differences between C_{vis} and C_{sota} in monitoring
2.2	H_1	Higher monitoring performance under C_{son}
3.1	H_1	C_{son} considered sig. more helpful than C_{vis}
3.2	H_0	Participants did not feel more in control with C_{son}
3.3	H_0	C_{sota}, C_{son} not considered more exhausting than C_{vis}

5.5. Experiment Conclusion

We wanted to find out how well *continuous sonification* can direct attention in comparison to *alert-based sonification* for process monitoring as a secondary task, something that, to our best knowledge has not been investigated experimentally before. Furthermore, there already exist approaches based on continuous sonification for peripheral monitoring, but most employed sonification techniques left room to believe that they would not be considered pleasing (as they e.g. are based on synthesized sounds that are not very complex), and could lead to fatigue if listened to over a long period of time, like a complete workday. Our approach, therefore, featured an event-based forest soundscape. We have developed a system that allows its users to compare the effectiveness of different sonifications for process monitoring in a fine-grained manner that extends beyond the typically used reaction times and binary correctness measures. The main task is simulated by means of simple arithmetic problems that have to be solved, whereas process monitoring involved a simplified simulated production process requiring several user interactions. An experiment with 18 subjects was conducted, comparing three conditions in a within-subjects design: C_{vis} (visual only), C_{sota} (visuals + auditory alerts after reaching a critical state) and C_{son} (combining the two former with a continuous, event-based sonification that applies a forest soundscape). Each of the three experiment parts was conducted for 10 minutes.

5.5.1. Results

The main results are:

- Participants were significantly more effective in process monitoring with C_{son} compared to C_{vis} and C_{sota}.
- There were no significant differences in terms of main task performance observed between the conditions.
- Participants found C_{son} significantly more helpful for monitoring, with C_{sota} being less helpful and C_{vis} the least helpful.
- There seems to be a strong polarization concerning whether C_{son} or C_{sota} can be considered more intrusive and distracting. It is probably safe to assume that a more pleasing sound design (e.g. more carefully selected sounds and improved mappings) could increase the acceptance of C_{son}.
- The experiment failed to prove that participants feel 'more in control of the process' under C_{son} compared to the two other conditions, but participants stated that the two modes that include audio were also not significantly more exhausting than C_{vis}.

In general, as Fig. 5.33 shows, it seems that a trade-off between main task and process monitoring task has to be made. If the main task is of the highest importance while process monitoring can be neglected, C_{sota} seems to be the mode of choice, as it shows the highest main task scores (but the lowest monitoring performance).

On the other hand, if is not absolutely crucial that employees are not to be disturbed during their main task, while process monitoring is also important, C_{son} seems to be the best suited mode: C_{son} leads to slightly lower main task scores than C_{sota}, but by far the highest process monitoring results. C_{vis} ,on the other hand, seems to be unsuitable for most cases, as it has a lower main task performance and at the same time a significantly lower process monitoring effectiveness than C_{son}.

5.5.2. Limitations and Future Work

In general, the process simulation of this experiment was designed to simulate a real-life production process. Such processes are typically 'bigger' and more complex, for example, concerning the number of machines and the number of values

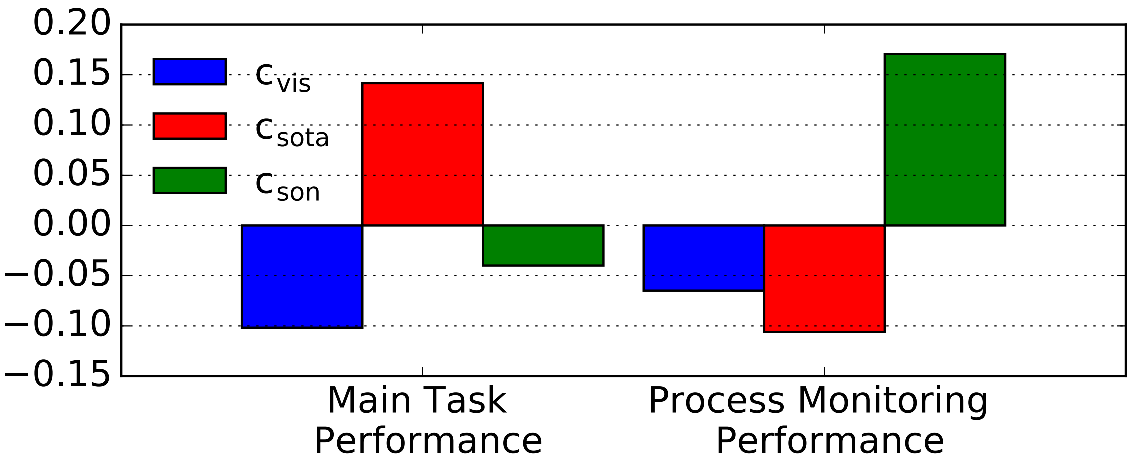


Figure 5.33.: Main task performance (left) versus process monitoring performance (right) for C_{vis}, C_{sota} and C_{son}. All numbers have been scaled to 0, compared to the respective mean experiment score in main and secondary task.

that can and should be observed (e.g. temperature measures). Therefore, in real production scenarios there can exist a significantly higher number of potentially critical states and situations, often making it difficult or nearly impossible to define them all beforehand. Such a complex scenario would potentially be less feasible with C_{sota}, as the situations and states that would issue an auditory cue would have to be defined beforehand. A continuous sonification that does not rely on pre-defined values and states, but instead conveys all interesting events and values, might be able to better handle such a scenario. However, on the one hand, the sound design of C_{son} would have to be adjusted to account for the fact that, when not all critical states and situations can be known beforehand, one also cannot map the approaching of such situations to volume. Implications of this might, for example, be that a continuous sonification would have to be designed so aesthetically pleasing that it is also acceptable when played at *normal* volumes over a complete work day. On the other hand, in a scenario with dozens of different machines, it would be difficult to distinguish and assign a unique sound to each machine. In such a scenario, the sonifications would either need to be based on aggregated, process-level data (such as e.g. so called KPIs - Key Performance Indicators), or the individual machines/data points to be sonified would have to be interactively selectable by the user. With such techniques, however, both a mapping of approaching critical states to volume in a continuous sonification, and an auditory-cue based sonification might be suitable again. Thus, for both approaches, different mappings would have to be tested and compared for more complex scenarios that better represent real-life working conditions as well. This could be achieved, for example, by modeling a more complex process simulation involving more machines, more data attributes requiring attendance, and more different interaction possibilities.

However, an even better way would probably be to install a sonification system in a real-world monitoring context (e.g. in the control room of a factory), and let users actually use the system for a longer period of time (e.g. for several work days or weeks). With a mix of questionnaires and semi-structured interviews, aspects of long-term usability and intrusiveness could be answered in more detail than was possible in this study. Furthermore, as requested by two users, further experiments will be conducted that include a condition with more fine-grained warnings that are more pleasing in the beginning, but gradually become more intrusive. At the same time, different – potentially more pleasing – C_{son} designs will be tested to find out (a) if they are considered more pleasing and participants could imagine using them for a longer period of time and (b) if they would still enable the same level of effectiveness as C_{son} of this study. Such sonification designs could, for example, be based on continuous soundscapes that are not based on short, repetitive events (such as in this experiment), but on longer, looped samples, or even *musical* concepts.

Furthermore, the literature suggests that the results of domain experts could differ from those of process monitoring novices, which would have to be evaluated as well. During this study, head tracking data was collected. The data will be analyzed and presented elsewhere. One hypothesis is that the participants in C_{son} had to shift their focus of attention significantly less frequently than under the other conditions. Although C_{sota} would have alerted the users in case of critical situations, it can be expected that participants did not always trust the system to convey an alarm in time; therefore, users may have checked the visual display to be safe.

To conclude, continuous sonifications, like our forest-soundscape sonification, enhance the adequacy of interactions in peripheral process monitoring better than displays based on auditory cues and systems that rely solely on visual means, while they do not significantly affect the main task performance.

5.6. Concept for Motif-Based Monitoring

As the previously presented SoProMon system specifically focuses on process monitoring as a secondary task, this section presents an approach that mainly aims to support users in direct monitoring. As direct monitoring is performed with the users' attention on process monitoring, more information can be perceived than during monitoring as a secondary task. At the same time, non-obtrusiveness seems less important, as there is no primary task on which users must concentrate. This prototype therefore bases on motif-based earcons, as suggested in the literature survey in Sec. 4.4.

A first prototype has been developed in which sound events are short melodic sequences based on the principles of musical contour (the direction and shape in which musical notes move) in order to increase recognizability. The developed prototype uses different instruments to convey the information regarding which activity corresponds to which event. Fig. 5.34 illustrates this mapping scenario.

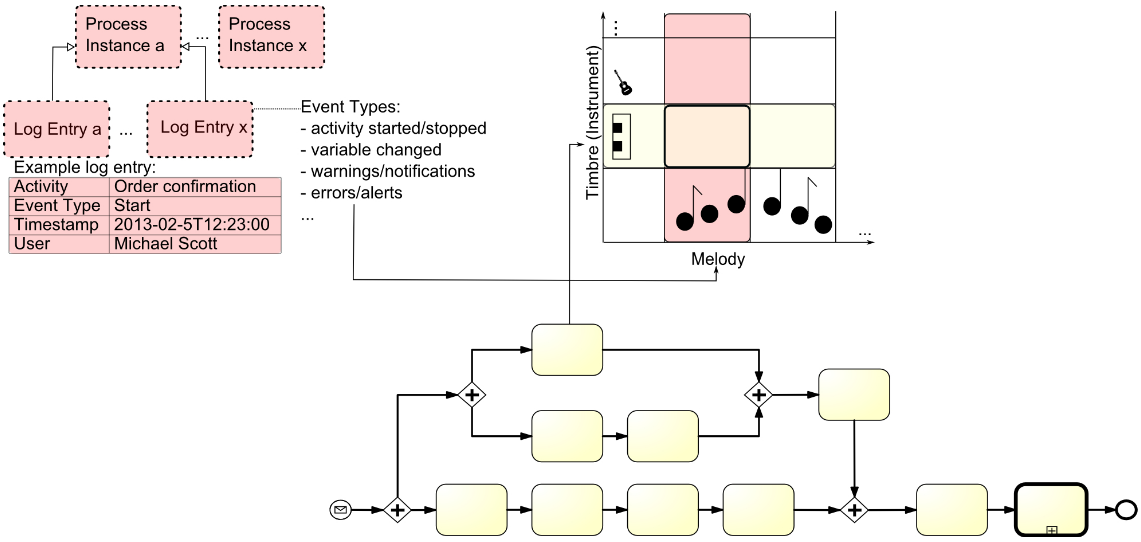


Figure 5.34.: Illustration of event mapping concept.

A sound event consists of a melody and an instrument, where the instrument is determined by the corresponding activity type (e.g. "invoice received"), and the melody is determined by the event type (e.g. "activity started").

5.6.1. GUI and Interaction Design

Fig. 5.35 shows a mock-up of a customization interface where users are able to define their mappings. The user interface should allow adjustments in how process events are sonified, as well as customization of which information is conveyed and in which level of detail during run time. The left side of the screen shows how detailed settings for the event sonifications (specifically, in this case, events related to the data flow) could look like. On the right side, the overview for the settings for the sonification of KPIs can be found.

5.6.2. Technical Architecture

From a technical standpoint, there are several requirements that a sonification component for real-time process monitoring needs to fulfill. Figure 5.36 shows a proposed multi-modal monitoring concept that tries to fulfill those requirements. The mapping of the sound events is based on the concept presented in Fig. 5.34. The general idea of the sonification concept we developed is as follows:

- The proposed system taps into event notifications issued from a PAIS during process execution in real-time (e.g. starting/stopping of activities, errors etc) over its API.
- KPIs that are calculated based on these individual events are received from a PAIS as well.

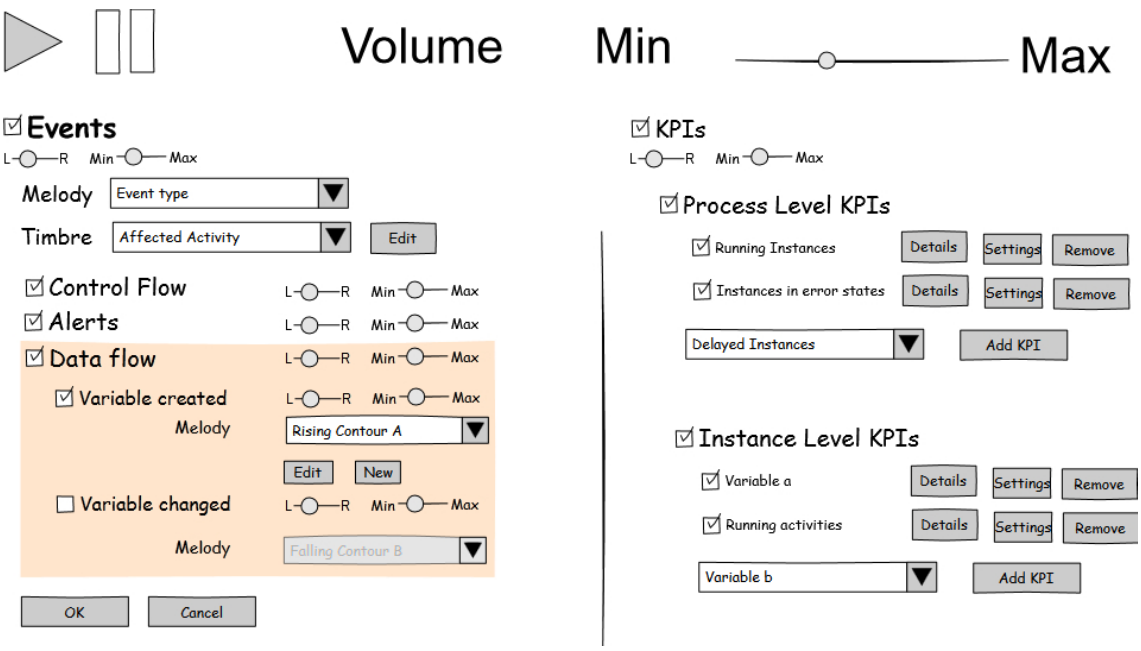


Figure 5.35.: Possible menu options for multi-modal process monitoring.

- Both data types, events and KPIs, are sent to the mapping component, which translates the data into sonification commands, depending on the users' wishes - both in terms of what data in what level of detail he or she is interested in (filtering) and how it should be translated to sound (mapping). These options are be customized by the user over a web-based user interface. Independent of the user settings, the sonifications of events and KPIs (or other quantitative data, such as sensor values) bases on different concepts:
- Events are sonified whenever they occur, for example by playing a short melody that conveys information such as the related activity and the type of event.
- Quantitative parameters, such as KPIs or sensor values, differ from discrete events in that they are continuously recalculated (or measured), therefore they are mapped onto continuous sound streams.

The system will be built in such a way as to enable the exchange of different modular sonification components, which can be based on different sonification techniques and methods. Therefore, it will be possible to flexibly select the sonification techniques that are best suited for a specific organization and its individual users.

Figure 5.37 shows a conceptual view of the proposed system architecture. The central component will be the Monitoring Component, which collects occurring process events from different sources (like the execution engine CPEE), pre-processes and sends them, according to the users' settings, over the messaging protocol OSC (Open Sound Control) to different Sonification Components. OSC messages adhere to a URI-style naming scheme. Each user can access a customized web interface where he or she can adjust the mappings from data

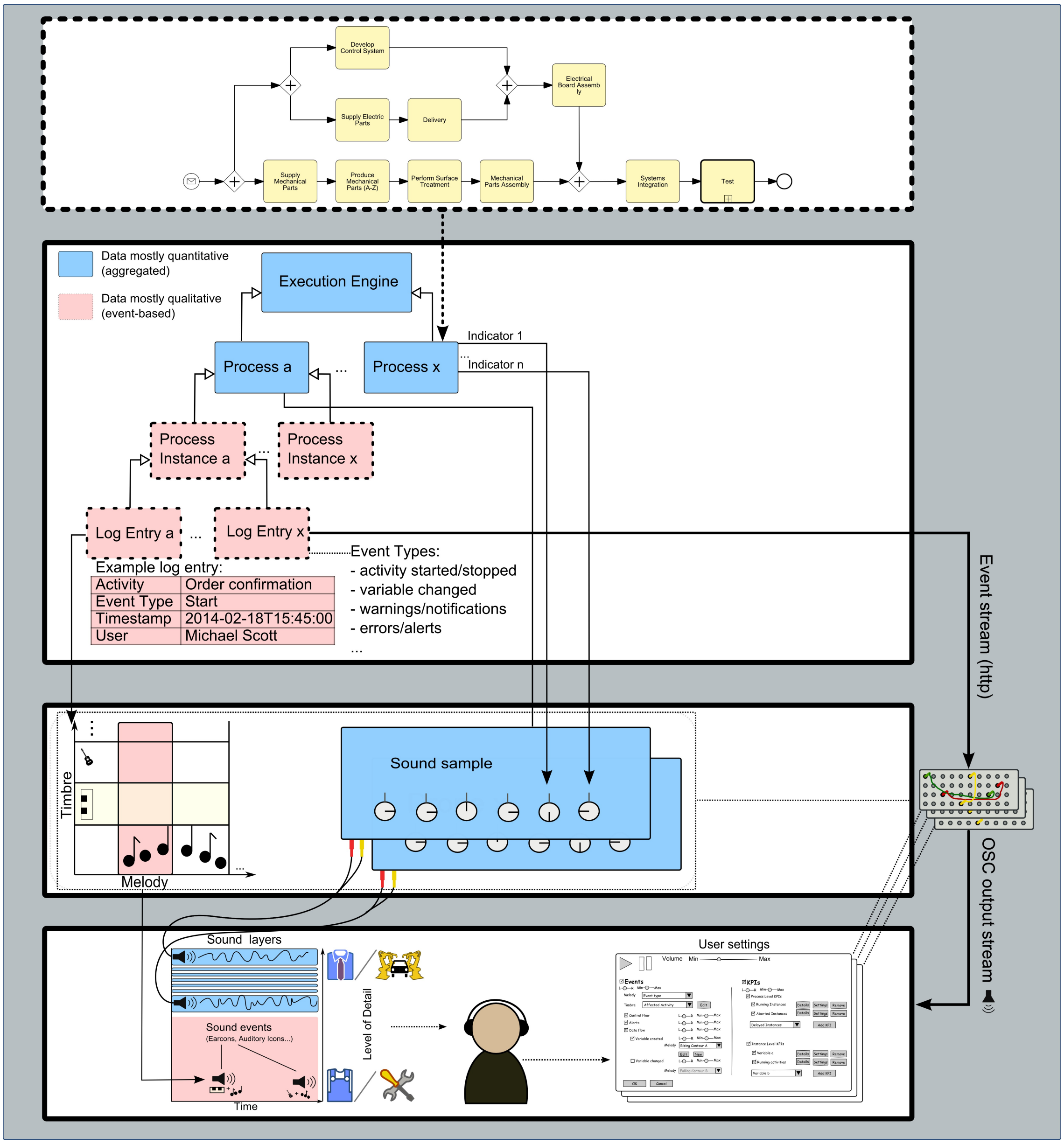


Figure 5.36.: Multi-modal monitoring concept.

to sound, filters and other settings that will directly effect his or her personal sonification.

A potential message schema for OSC messages that are sent to the sonification components could consist of the following hierarchy:

- /Event: an event that needs to be sonified has occurred. The following example shows how the start of an activity could be conveyed: /Event/Process A/Instance 1/Activity/"construction"/Started
  The following example shows how the change of a variable could be conveyed:
  /Event/Process A/Instance 1/variable/"variable a", 20

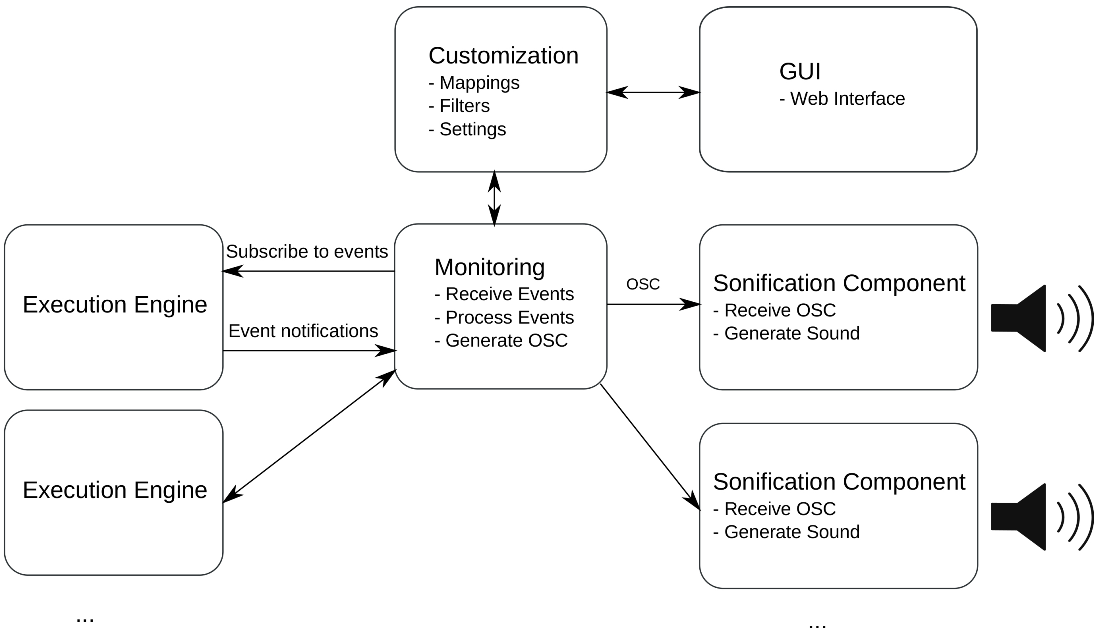


Figure 5.37.: Architecture proposal for multi-modal process monitoring.

- /KPI: a KPI has been recalculated, for example:
  KPI/Process A/relationOverdueOrders, 5.4
- /Mapping: the user has changed mapping settings in the GUI, like mappings to melodies and instruments
- /Settings: the user has changed general settings in the GUI; for example volumes for individual channels.

Different sonification components could be implemented in different languages; for example one for sonifying events, and another one responsible for sonifying KPIs. The only technical requirement for the programming language is that a library that implements the OSC specification exists. Different programming languages have been successfully tested in terms of their suitability as sonification components, like SuperCollider, the Java library JFugue (in combination with either Java or Processing) or Pure Data. Chuck had also been tested; however at the time of testing (early 2013) it had not yet supported the full OSC protocol, only basic properties of it.

5.6.3. Preliminary Implementation

Example recordings of this prototype can be found at:
http://soundcloud.com/process_sonification/

The recording "Event types"[9] shows a sonification of five different event types in sequential order ("activity started", "variable changed", "warning occurred", "error

[9] direct link: https://soundcloud.com/process_sonification/event-types

occurred" and "activity finished"). Most melodies in this example were played on the same instrument (piano), except for "variable changed" and "error occurred", signifying that the events all occurred while executing the same activity. The second example, "Activities"[10] shows different events of the same type ("activity started") that are all played with different instruments, meaning that these events are related to different activities. The "Event sonification examples" 1[11], 2[12] and 3[13] show different sonifications of a sample process' execution. The first example contains a warning in the second activity. In the second example, an error occurs instead of the warning; therefore, the second activity does not finish. In the third example, no errors or warnings occur; subsequently, the second activity is finished before the first. First informal user evaluations concerned with distinction and recall of event notifications suggest that if the same instrument is always used, the different event types can be distinguished and memorized even after only rudimentary instructions. However, with a high frequency of occurring events in particular, the distinguishing of the different event types becomes more difficult as soon as different activities (and thus different instruments) are involved. However, user performance seems to increase with training time.

The presented examples only represent a possible mapping for the sonification of execution events; KPIs have not yet been considered.

5.7. Concept for Soundscape-Loop-Based Monitoring

One of the outcomes of the experiment to evaluate the SoProMon system was that the presented continuous sonification was found to be only moderately pleasing. One assumption for this cause lies in the fact that the sound files are short and, due to the fact that they are played for each production step, the resulting sonification can sound quite repetitive. In the process simulation, this challenge was somewhat alleviated by the random factor that was introduced - however, if the production in a real production setting is more regular, the resulting sonification will be as well. Our approach for a redesign is to sacrifice accuracy on the detail level of machine executions, and instead to use longer sound samples with lengths of several minutes to represent each machine. This new sonification concept, therefore, abstracts from the individual events that occur during process execution, and only presents high-level KPIs and aggregated data. In this way, less information can be conveyed, but the sonification can be designed to be more subtle and potentially less distracting from the main task than the previously presented approaches. At the same time, by being based on aggregated data and abstracting from individual events, an additional user

[10] direct link: https://soundcloud.com/process_sonification/activities
[11] https://soundcloud.com/process_sonification/event-sonification-example-1
[12] https://soundcloud.com/process_sonification/event-sonification-example-2
[13] https://soundcloud.com/process_sonification/event-sonification-example-3

requirement not covered by the sound design used in the SoProMon study has been fulfilled.

A possible sonification within a forest scenario could, for example, consist of 3-4 sound streams, each representing the status of one machine/production line or one specific production-related KPI. For each of those streams (e.g. consisting of rain or waterfall sounds), different samples for different intensities (e.g. rain in different stages, from light dripping to a thunderstorm) can be exchanged in real-time, depending on the value of the corresponding KPI. Instead of a simple exchange of audio samples, a smooth fading could make the transition more natural. Alerts or exceptional situations could be conveyed in this scenario be by the introduction of sounds that are unnatural to the respective soundscape, like chainsaw sounds in the forest scenario.

An audio recording of the mock-up prototype can be accessed on the internet[14]:
S1-forest.mp3.

5.8. Conclusion

Multi-modal sonifications for process monitoring need to fulfill different requirements, depending on factors such as the amount and type of data that needs to be conveyed, or whether process monitoring is primarily conducted as the only task of the user, or as a background activity in parallel to other work. The user study based on the SoProMon system proved that a multi-modal sonification can improve user performance during process monitoring as a secondary task. However, the sonification design is crucial for both the effectiveness of attention allocation and user acceptance. The presented continuous soundscape sonification based on short forest sounds was deemed effective overall but only moderately pleasing, which is why a design using longer, loop-based samples has been developed as well. For direct monitoring as a primary (or only) task, more information can be conveyed, as there is no primary task from which to distract. A presented approach combines motif-based melodic earcons with drone sounds to present event notifications and KPI developments at the same time. However, pleasingness and user acceptance should be measured for all monitoring use cases to enable long-term suitability.

[14]see http://doi.org/10.4119/unibi/2752965, file: S1-forest.mp3

6. Visualization and Sonification for Business Process Analysis

This chapter describes the development of a concept and multi-modal prototype to support users in analyzing historical process execution data, as well as first evaluations of that prototype (see Fig. 6.1).

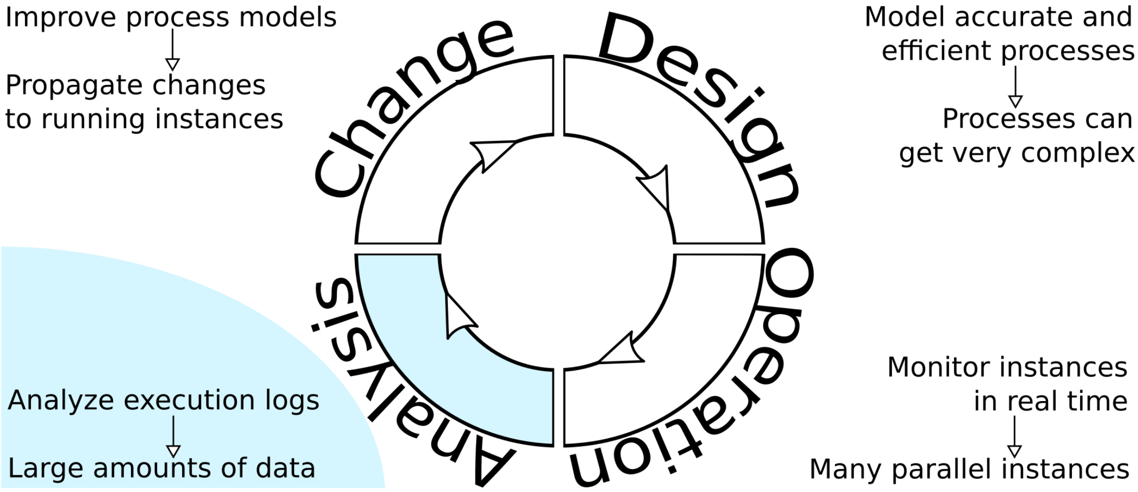


Figure 6.1.: Analysis phase of the business process life cycle.

This chapter contains ideas, results, figures, and text from my previous publications. The description of the ProSon-Plugin and its evaluation contains text, figures, and results that were published in (Hildebrandt, Amerbauer, and Rinderle-Ma, 2016).

6.1. Concept of the ProSon-Plugin

As process mining techniques have proven beneficial for the analysis of process execution data, we do not suggest replacing machine learning methods with multi-modal visualization and sonification, but rather complementing these means. The concept of visual analytics - where visualization is used to obtain a first overview of the data, to find interesting areas in the data for further analysis, or to generate hypotheses - can serve as a guideline for the integration of (semi) automatic data processing with human perception and interaction. The approach presented here follows the same principle, but further extends visual analytics to include sonification, resulting in audiovisual analytics (or multi-modal analytics).

As several approaches for the visualization of qualitative data already exist, and visualization research is generally more established and mature, the main emphasis in the development of the ProSon-Plugin is placed on sonification design. As sonification for exploratory data analysis is a relatively young discipline, there is still a lack of established guidelines for best practices concerning sound and mapping design in most areas.

Although the developed plugin is generic enough to allow for most use cases of process mining and process data analysis, including process discovery, such as play-in (creating a process model out of an event-log), we assume that the strengths of sonification will lie in cases where a process model has already been defined, and the goal is to detect anomalies and deviations instead.

The usability and thus user acceptance of systems for multi-modal process data analysis is based on the design decisions that are made in several areas. One crucial factor is probably the design of the interactive process that enables users to create sonification mappings. This is because most potential users will not be experienced with sonification, as at least in western culture, the visual sense is predominant and sonification is still a research topic in many areas. Therefore, the interaction design should support users in their process analysis workflow as much as possible.

As log files typically contain qualitative as well as quantitative data, suitable mappings for all data types should be available. Table 6.1 presents the log data attributes that are currently supported by our concept. There are three basic types of mappings that need to be supported. *Timestamp* refers to the concept that the time intervals between individual events should be represented. *Nominal* variables are not represented by numbers and thus contain no inherent ordering, such as different activity names. *Continuous* variables, on the other hand, are represented by numerical values.

Table 6.1.: Supported concepts

Concept name	XES	Mapping
Trace name	concept:name	nominal
Event time	time:timestamp	timestamp
Event name	concept:name	nominal
Event life cycle state	life cycle:transition	nominal
Event organizational resource	org:resource	nominal
Event organizational role	org:role	nominal
Event organizational group	org:group	nominal
Event cost	cost:total	continuous

6.1.1. Technical Implementation of the ProSon-Plugin

ProM is based on the Java graphic library Swing which is build on the Java graphic library AWT. Therefore, the developed prototype is also based on Swing and AWT, as well as on the slickerbox library, which is also used in ProM, in order to give the user interface a ProM look and feel.

The sonification is based on JFugue[1], a Java library for music programming. JFugue is an abstraction over the MIDI capabilities of the JVM (Java Virtual Machine). MIDI, short for "Musical Instrument Digital Interface", is a protocol for communication and synchronization of electronic instruments, musical software and other electronic equipment. In the beginning, it was mainly used to connect different electronic music instruments together. One common example of its usage was to use one keyboard to control several (hardware) synthesizers. However, it is increasingly being used to connect external devices (such as keyboards) to computer software such as software synthesizers, or for communication between different music and audio software. The MIDI protocol does not transmit any audio information, just event messages and control parameters such as played notes, volume and pitch. Computer software based on the MIDI-protocol sends information that the sound card interprets. The sounds that are currently available for playback are nowadays in many cases based on the *General MIDI*-standard, a standardized set of 128 instrument and drum sounds (all of which are available in different pitches). The quality of the played sound depends heavily on the sound card. The actual sounds that are played upon triggering have been generated from basic sound waves by early sound cards, while current sound cards contain the sounds as recordings of real instruments already in their memory.

JFugue abstracts from the General MIDI standard by allowing the translation of a string of musical instructions into MIDI commands. In such a way, the loaded event log, together with user settings such as filters, mapping options and other parameters such as playback speed, is translated into a JFugue music string, which in turn generates MIDI commands that are sent to the user's sound card, which plays back the generated sonification. In general, a JFugue MusicString consists of control information (such as the channel, the instrument, volume and speed) and a pattern of musical information (such as the notes to be played, octaves and durations of the notes). In contrast to the Java MIDI-package[2], it is possible to enter high-level concepts, such as musical chords, directly.

6.1.2. Visualization

The main purpose of the visualization in our multi-modal concept is to show the effects of sonification mappings that the users create, and demonstrate the combination of the two modalities. Therefore, the visualization follows the sonification,

[1]http://www.jfugue.org/
[2]https://docs.oracle.com/javase/tutorial/sound/overview-MIDI.html

meaning that changes performed in the sonification mapping are reflected in the visualization. On a conceptual level, the chosen visualization operates on the same abstraction level as the sonification. The basic principles of the visualization are based on the dotted chart visualization available in ProM, which can be seen as a baseline for visualizing business process event data (Song and Aalst, 2007). It intends to offer the user an overview concerning the distribution of traces and events over time at a glance. However, not all aspects of the dotted chart visualization have been implemented, as the intention of our prototype was not to replicate all functionalities, but instead to demonstrate the principles of multi-modal visualization in conjunction with sonification.

The main visualization concepts that are used in our prototype are positioning on the x- and the y-axis, color and shape (e.g. to visualize the type of an activity or trace). If no mappings are applied, events are visualized by a black dot. Mappings to color and shape are only applied when both an instrument and a melody are mapped to attribute values contained in the event, as a melody without an instrument or vice versa would result in an incomplete sonification mapping. Five different shapes (square, circle, star, plus and rhombus) have been implemented. During playback of a sonification, a progress indicator shows the playback position, so that the user can know which event is being sonified while the sonification is running. At the moment, the ordinal value sonification mappings *volume* and *panning* have no visual equivalence. Table 6.2 presents the previously explained types and how they can be mapped onto visual and auditory properties. While the mapping of *timestamp* is implicit, and thus cannot be changed by the user, mappings concerning nominal and continuous variables can be customized.

Table 6.2.: Visualization and sonification mappings

Type	Visualization	Sonification	Assignment mandatory
Timestamp	position (x-axis)	playback time	implicit
Nominal	shape, color	instrument, melody	yes
Continuous	not implemented	volume, panning	no

6.1.3. Sonification

Based on the literature analysis presented in Sec. 4.4, our proposed concept is based on parameterized motif earcons, similar to the ones described in (Cullen and Coyle, 2005) or our concept for motif-based monitoring presented in Sec. 5.6.

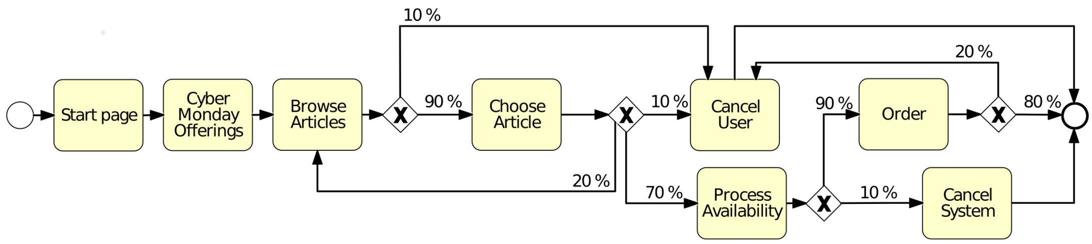
Before starting the playback of the generated sonification, the respective music string is generated. Sound events that are to be played in parallel are each played on a different channel. Gaps between subsequent events are translated to rest events in the music string. Nominal event data (e.g. the name of an activity or of a user) can either be mapped to one of the 128 available instruments, or to a succession of tones, like a short melody - the user can freely enter assigned melodies over the GUI. By default for each data entry a different chord is predefined. For the available continuous variables (volume and panning), the assigned variable is mapped to a value between 0 and 127, as numerical values in the General MIDI standard are usually contained in this range. The timestamp is implicitly mapped according to how the respective earcon is placed within the temporal structure of the resulting sonification. The user can adjust the total playback speed of the sonification in two different ways: he/she can adjust the overall playback speed, which influences the temporal gaps in between the different process events (and subsequently the assigned musical events), or the event playback speed (which in effect influences the duration of the assigned musical event). Both factors can be adjusted in conjunction or separately.

6.2. Use Case-Based Evaluation

This chapter presents a fictive use case modeled to represent a realistic situation, in order to test the suitability of the developed approaches for typical scenarios. It shows the usability of our concept for event logs with a high data density, such as those generated from processes with a high level of automation, like production-or e-commerce processes. We base our use case on a fictive user and his/her information needs and evaluate if and how our proposed concept can best support him/her in this task. Ultimately, this will lead to assumptions regarding which use cases can benefit most from introducing sonification, and for which types of data and questions it is most suitable.

In general, as we do not suggest using sonification as a replacement for visualization, data processing, filtering or statistical analysis, we do not compare our approach to existing visualization or process mining approaches. For this first informal evaluation, we decided against a process-discovery use case. The reason for this decision is that process discovery is an area already targeted by sophisticated process mining algorithms. Furthermore, we believe that due to the characteristics of sonification, our approach is more suitable for use cases based on already discovered or defined process models, such as anomaly detection, process improvement, root cause analysis of irregularities or errors, and retrospective process performance monitoring. The main reason for this is that our auditory perception is especially suitable for hearing deviations and irregularities, particularly in otherwise regular or repeating sound streams. Process discovery, on the other hand, is typically more focused on discovering *normal* and *typical* behavior instead of outliers. However, as our plugin might nonetheless help to some extent with the task of process discovery, we plan to evaluate its suitability for this task at a later point in time. Thus, the use case at hand concentrates on anomalies and irregularities, especially those that are not targeted by *classical*

Figure 6.2.: BPMN model of the order process with probabilities of path selections used in the simulation.



process anomaly detection, such as anomalies that are not based on a single trace compared to normal execution, but instead on, for example, instance-spanning trends and developments.

6.2.1. Use Case

In order to evaluate and present our concept for log files with a high event density in which many events occur within a short period of time, our fictive evaluation use case is based on the order process of a web shop. The process (see Fig. 6.2) models the steps that customers take when using the web shop. The numbers on the exclusive gateways describe the probabilities that a process instance follows a certain path. The execution time of each activity varies within a given range. With this information, the process was simulated 1,650 times, and the pauses between the first activities (*Start Page*) of subsequent traces generally vary between one second and five minutes. The generated log file contains web shop activity between 27/11/2015, 6 PM and 28/11/2015, 6 AM. Within this 12 hour period, 11,918 events were produced. As the path of each trace depends to a certain degree on a random variable, several anomalies are contained in the generated execution log *natively*, such as, for example, traces with an unusual number of loops. We furthermore explicitly applied two artificial anomalies to the order process:

1. The generated log file contains the time period between 6 AM and 6 PM. We assumed that activity on the shop's homepage decreases throughout the night, hence this activity decline has been simulated for the log. However, 9 hours after the start of the log file, we modeled a burst of high activity. In this period of 5 minutes, new traces were generated with a pause that ranged from 200 milliseconds to 5 seconds.
2. We simulated a system malfunction between 6 and 8 hours after the beginning of the log file. This was simulated by changing the probabilities of the XOR decision between the activities *Process Availability* and *Order* and *Cancel System*, resulting in 100% of availability checks leading to the *Cancel System* activity.

In this scenario, we assume the owner of the web shop to be the fictive user of our prototype. The web shop owner is, of course, interested in different types of process-related data, mainly those that are directly related to business performance, such as the number of orders in a given time period. On the other hand, the owner is also interested in a constant quality of service throughout the ordering process, thus he/she is interested in potential service outages or technical problems that might, for example, have prevented a sale from being completed. In our example, we assume a process model that has already been defined, and that is being enacted by a business process management system, or supported by some other kind of PAIS. Thus, it can be assumed that all events related to that process are logged in a suitable format, like the XES format. In our fictive example, the web shop owner performs process data analysis every morning based on logs from the previous night. The owner wants to answer the following questions:

- Have there been problems or irregularities with the technical execution of the process, or with the availability of the server? This might be the case if, for example, activities take substantially longer than usual to finish, or if there has been a high number of cancellations by users (which of course can also be a sign of other problems, e.g. that users did not find what they were looking for).
- Have there been phases of particularly high or low activity?
- Have there been users who were browsing and selecting more articles than the average user, possibly indicating that they did not find what they were looking for?
- Have there been phases with an increased frequency of cancellations by the system?
- Have there been other noteworthy/irregular occurrences in individual traces?

6.2.2. Evaluation

Audio recordings of the prototype can be accessed on our project homepage[3] and on SoundCloud[4]. The XES file used in the evaluation can be downloaded on the project homepage as well[5].

The source code of the plugin, as well as a standalone version of the prototype, and instructions for the integration into ProM are available as well[6].

In order to analyze the event log of last night, the web shop owner first loads the generated log file into the prototype (see Fig. 6.3). On the right side, the menu with different options for playback, filtering and mapping, as well as detailed

[3] Project website: http://cs.univie.ac.at/project/sopromon
[4] https://soundcloud.com/tobias_hildebrandt/sets/sonifications-of-business-process-data
[5] XES file: log.xes
[6] Prototype: https://github.com/felixamerbauer/business_process_sonification_package

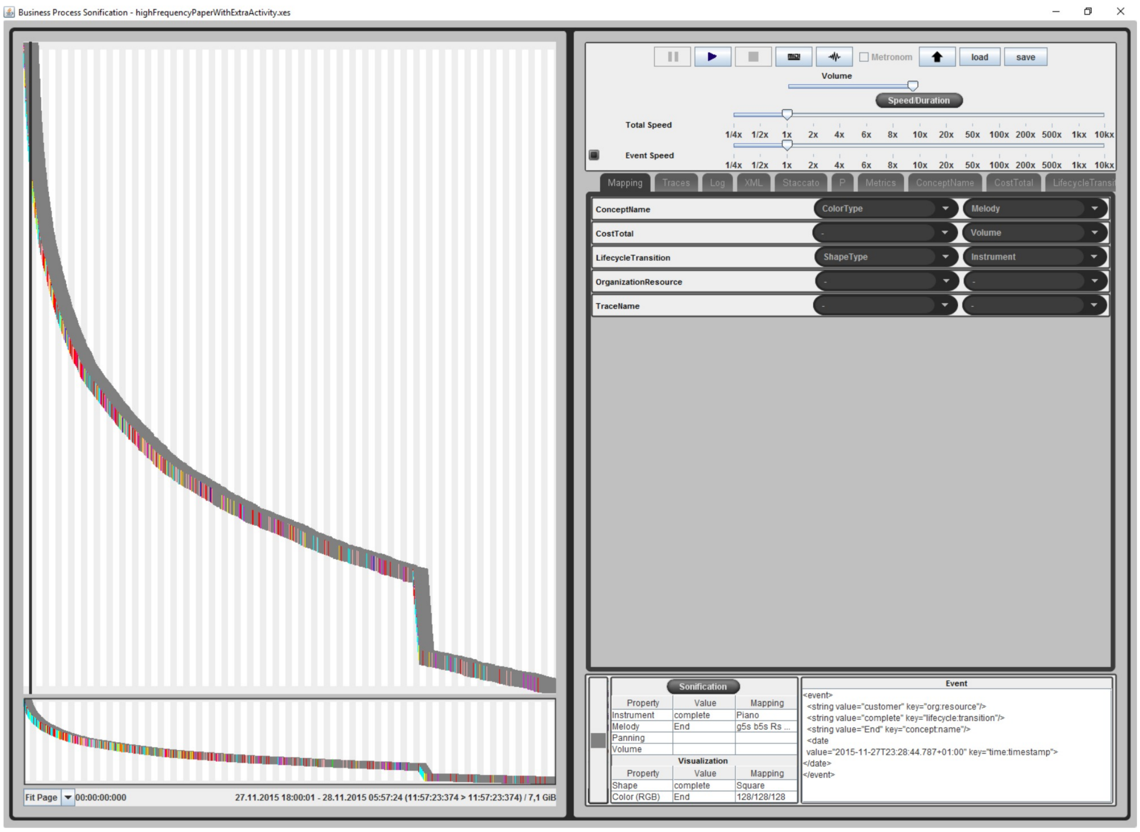


Figure 6.3.: ProSon GUI with unfiltered event log loaded.

event information, can be found. There is such a high number of events that one cannot derive much information from the visualization without filtering or zooming in, except that towards the end of the log, a high number of traces started, a variation visible by the *step*.

Playing back the events at the same speed as they were logged is not a viable option in this case, as the playback would have a duration of 12 hours (as does the log file). Thus, when one wants to obtain a overview of the data in a short amount of time, the playback speed has to be increased.

Development of Traffic

The fictive user first wants to analyze the development of user activity of the previous night, and determine if there have been activity peaks. Thus, at the beginning, the web show owner is not interested in all activity types, just in those of the type *start page*, which is why he/she filters for this activity type using the menu on the right. As, even after filtering, a high number of events remain, the owner enters a melody that consists only of one note (a C in 3rd octave) to be played whenever the event appears. A longer melody consisting of 3 or 4 short notes, perhaps with different durations and short breaks between them, would be more recognizable. However, as such a melody would take more time to be

played, this would restrict the possible increase in playback speed. A single note, on the other hand, can be played in a short period of time.
As *start page* is the only activity enabled, it is not necessary to be able to differentiate different melodies (i.e. activities). This first mapping thus only intends to give a first impression of the development of traces over time by the number and density of notes played. As no variable is mapped onto the concept *instrument*, all notes are played using the default instrument "piano" (which of course can be adapted in the menu).

tracesshort.mp3 demonstrates the concept on a short version of the log file. However, the sonification of the 12 hour log would also take 12 hours to play at regular playback speed. Thus, the owner increases the playback speed, and the event playback speed accordingly. This way, not only are the pauses between the events decreased, but the playback speed of the events is increased as well, meaning that the duration of the notes is decreased to *fit* all events into the shortened time frame.

tracescomplete.mp3 shows the same mapping applied to the original, full log file; however, it is based on a compressed playback with a duration of just 20 seconds, constituting a compression factor of 1:2160. As in these 20 seconds, 1650 events of the type *start page* are contained (82.5 per second), individual piano notes can hardly be discerned anymore. Instead, as the short piano notes are played whenever the respective event occurs, several events of simultaneous traces often overlay each other. The resulting sonification has similarities to that of a Geiger counter, both in terms of the resulting output and the principles behind the sonification mapping. Just as a Geiger counter measures the level of radiation, this resulting sonification measures the level of process activity.

What can be detected from the sonification (and from the visualization) is that there is a short burst of activity towards the end of the log file. What can be heard also, but not so easily seen in the visualization without zooming and scrolling, is that the density of activities decreases more or less steadily towards the end of the log file (apart from the mentioned anomaly).

If the user is interested in analyzing a smaller section of the log in detail, such as the short activity burst, the user may use other plugins available in ProM to filter out regions that are not of interest to him/her. Another option would be to stay in the ProSon plugin and use the zoom functionality. This way, the user can click on individual events that are of interest to obtain details, such as timestamp, in order to conduct root cause analysis.

Development of Orders

The web shop owner now knows about the development of traffic over time. In the next step, he/she would also like to know if the development of orders is similar to that of new traces, as not every user is placing an order. Therefore, the owner uses the menu to filter out all activities except those of the type *order*.

order.mp3 shows a sonification with its playback speed increased in such a way that the final sonification is only 10 seconds long (see Fig. 6.4).

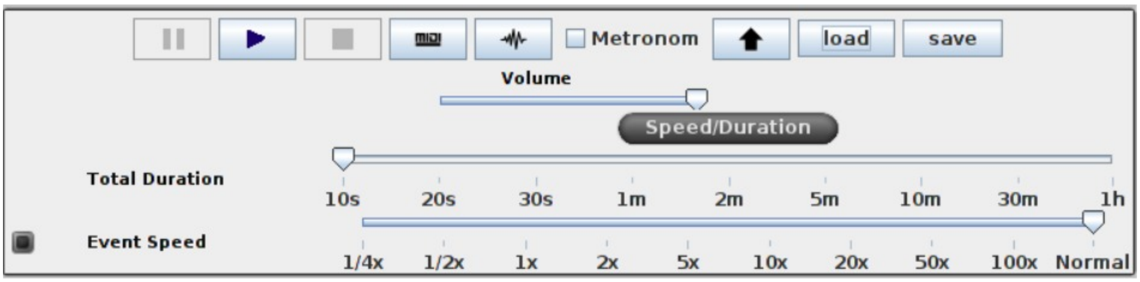


Figure 6.4.: Transport menu

What can be deducted from this sonification is that the placement of orders generally follows the development of traces. However, the web shop owner can already see in the small bottom left overview visualization (see Fig. 6.5) that there is a relatively large period of time towards the end where no orders have been placed. The user could then zoom into that period and, by activating the other activity types again, find out which traces are concerned to find out which traces are concerned to determine potential causes of the interruption of orders.

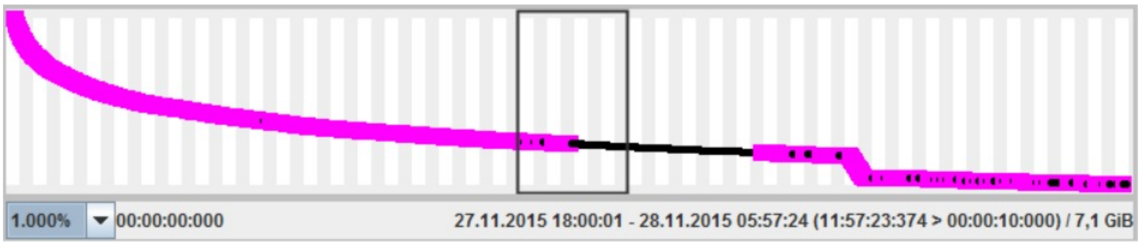


Figure 6.5.: Overview visualization of log filtered for event type *order*.

Orders and User Cancellations

As the unusual number of traces that contained no orders could have been caused by a high number of users canceling the order process, the web shop owner is interested in the distribution of orders in relation to that of user cancellations. Therefore, the owner uses the GUI to filter out all but those two activity types. Now, as there are two different activity types to be heard, he/she needs to assign two different melodies to distinguish these. As a high number of traces still exist, the user decides again for single notes. He/she assigns a relatively low note (a C in the 3rd octave) to *cancel user* events, and a higher, shorter note (an E in the 6th octave) to *order* events. In order to demonstrate the two different notes, ordercancelshort.mp3 shows the mapping applied to the shortened version of the log file, played back in real time. As there are more high notes than low notes, it can be deduced that there are more users who place orders than who cancel the process. However, due to the aforementioned reasons, for the complete log, the playback speed has to be significantly increased.

ordercancel.mp3 shows a very condensed sonification of the complete log file that is 10 seconds long. What can be heard is the previously discovered period without orders, notable by an absence of high notes. As the low notes are still played during that phase, the user can deduce that the lack of orders in that phase can mostly be explained by a high number of user cancellations. Furthermore, the user can hear that the distribution of orders and cancellations follows the general traffic trend but is, except of the short anomalous period, distributed quite regularly. The user again now has the option to either use other plugins for root cause analysis, or to use our plugin by zooming in and clicking on particular events.

Browsing and Choosing

Another point of interest for the web shop owner is the presence of users who browsed and selected more articles than average, indicating that they did not find what they were looking for in the shop. Again, the web shop owner filters only for the two activity types he/she is interested in, in this case *browse* and *choose*, and assigns two single notes, this time a lower one for *browse* (C in 4th octave) and a higher one for *choose* (D in 6th octave). This follows a logical analogy, as the activity *browse* always precedes the activity *choose*; therefore, the increase in pitch follows the process path.

browsechoosetraces.mp3 presents two sonifications of different single traces to demonstrate the concept. Next, to be able to mentally assign an activity to a trace, the user maps trace to the concept *instrument*. This means that all events of a particular trace are played using the same instrument. Thus, if, for example, trace 1 is mapped to the instrument "piano", a *browse* event in trace 1 is sonified by playing the note C in 4th octave on a piano.

browsechooseshort.mp3 demonstrates the mapping at hand of a short example, with playback set to real-time. For each trace, a different instrument is mapped, until after 127 traces the available instruments repeat themselves. In the short example, one can already imagine that it might be difficult to remember if an instrument has already been played or to mentally group events that belong to the same trace. A possible solution would be to load the log file in the *untangled* mode, which means that, while loading, the log files are sorted in such a way that a new trace only starts after the previously started trace has completed, with a pause of a few seconds in between. For a log file with a relatively low number of traces it is in such a way possible to detect anomalies or deviances, such as traces with higher numbers of browsing and choosing than usual, by listening to one trace after another. However, again for high-density log files, such as the one at hand, playback has to be sped up.

browsechoose.mp3 presents a sonification of the full log condensed to 20 seconds, with the concept *instrument* mapped, and loaded in the regular fashion. As could be anticipated, individual events can no longer be discerned at this speed. However, by the rate of the instruments repeating themselves (every 127 traces) one can estimate the frequency of new traces.

browsechoosesequence.mp3 on the other hand, presents the same mapping loaded in the untangled fashion. The resulting sonification sounds a bit more ordered and less chaotic, as each trace is played before the next one, but the basic challenge remains the same: at this playback speed, it is impossible to make out individual traces or even events. Slowing down the playback speed, on the other hand, would make a sonification of all traces in a short time unfeasible.

Other Irregularities

Finally, the user wants to detect other types of irregularities and anomalies for which he/she was not actively looking. Thus, the user filters out any activity types that are not of interest, and maps the visual and auditory properties of the others. Fig.6.6 shows the activities our fictive user is interested in, and how they are mapped to short notes. Again, the pitch increases with advancing process state. Increasing melodies symbolize positive events; decreasing pitches negative ones.

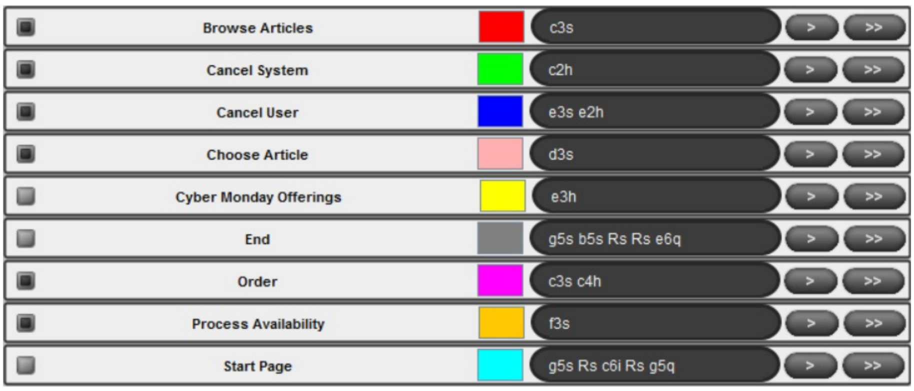


Figure 6.6.: Mapping of melodies to activities.

Fig. 6.7 shows the visualization of a small log file filtered for those event types.

To demonstrate the concept, the log is further filtered for two individual traces to show how the resulting sonifications of different execution paths sound (see Fig. 6.8).

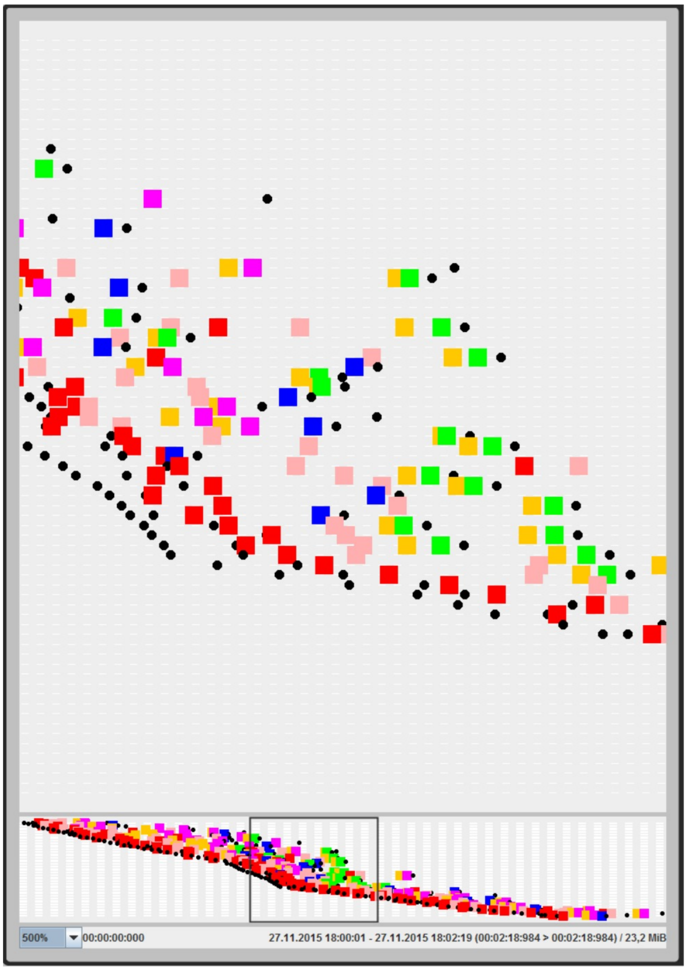


Figure 6.7.: Visualization of small log file filtered for certain activity types.

completetraces.mp3 presents the two traces played in a sequence. As this example present single traces, the mapping of instrument to trace is not yet necessary.

completeshortsequenced.mp3 presents a short log loaded in the untangled mode. To hear when one trace ended and the next started, *trace* is mapped onto *instrument*. At the selected playback speed (4x), it is audible if individual traces show interesting or deviant behavior, thus such a sonification is suitable for either very small log files, or for those that represent a longer time period, but exhibit low event density. For high frequency logs, the speed has to be increased again.

complete.mp3 presents the same sonification for the complete log file with the playback speed increased to 500x.

completereversed.mp3 presents a reversed mapping, where the activity type is mapped onto *instrument* instead of onto *melody*. *Trace* is not mapped. In

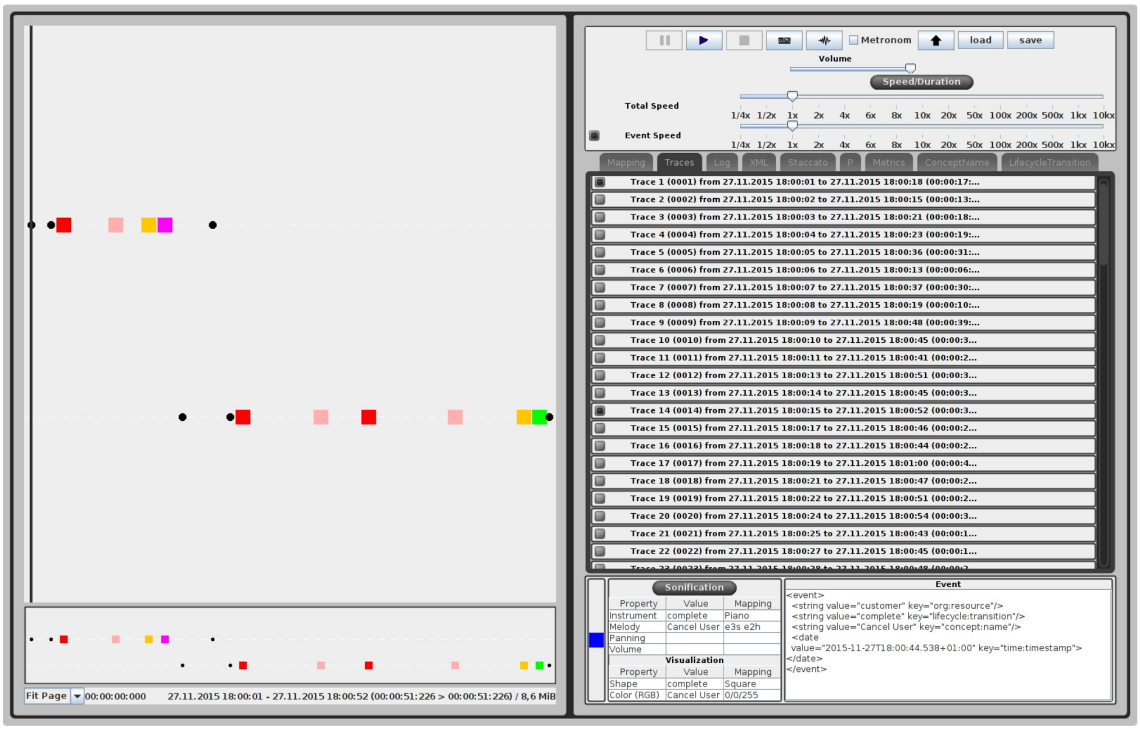


Figure 6.8.: Small log file filtered for two individual traces and certain activity types.

this example, one can hear instantly that during the period without orders, many users have canceled the process.

6.3. Conclusion

Though a large variety of algorithms for process mining and automated processing of event data exists - for example, in the form of ProM plugins - visual analytics is an important addition to process analysis. To further build on our cognitive abilities of pattern recognition, we propose a method of enhancing visualization with techniques from sonification, forming a concept for multi-modal process analysis. Although sonification has been widely researched for data analysis and real-time monitoring, no comprehensive approaches for its application in process event data analysis exist yet. As there seem to exist, independent of the application domain, very few approaches that investigate the usage of sonification in analyzing qualitative data, we believe that some of our results can also be transferred to other domains in which data of similar structure needs to be examined. To demonstrate our multi-model approach, we developed a prototype as a ProM plugin, which combines visualization and sonification. It has been evaluated by way of a fictive use case in order to answer our research questions.

To summarize, it seems that sonification for the purpose of process data analysis is suitable for the same tasks as visualization: both modalities suggest themselves to gather initial rough impressions about the data structure and distributions, and see/hear general trends in the data in a short time. Thus, many of the fictive users' questions, particularly those that concern discovering general trends or phases in which a large number of traces show deviant behavior, can be answered with our proposed concept. This is an area where *traditional* conformance checking has certain limitations, as it typically tries to detect anomalous traces that differ from pre-defined process models or *typical* traces, but is not focused on detecting instance-spanning tendencies and irregularities.

However, with both visualization and sonification, it seems difficult to detect small anomalies or deviations that concern only single traces in large amounts of event data, such as detecting users who were browsing and choosing more than usual. Thus, it is often necessary to perform filtering and selection, but methods from process mining and other algorithmic processing will also help to detect data elements and events of interest. After resorting to those methods of data preparation and processing, sonification and visualization are again suitable to find differences and anomalies between single traces, or to find out what caused these deviations. For small event logs or such with a low event density, such intermediate steps might often not be necessary, as demonstrated in this section at hand of shortened log files.

As an example, a typical process analysis workflow for high-frequency data sets might consist of the following steps, in accordance with the earlier presented information-seeking mantra (Shneiderman, 1996):

1. Load the log file into the ProSon plugin to get a first overview, and see/hear interesting segments and anomalies.
2. Filter and sort the trace concerning either specific activity types, or time periods in which anomalies have already been detected in the previous step. Depending on the type of process data, users might also want to filter the trace to only contain orders exceeding a certain order value, or only traces that have not been completed. Such filtering is partially supported by our proposed plugin, but other plugins can also be used.
3. Load the filtered log file into ProSon, detect more irregularities or more specific irregularities, perform root cause analysis by filtering, selecting, and zooming, or, if necessary:
4. For root cause analysis or automatic processing, use other ProM plugins or other systems, if necessary.
5. Repeat steps until all questions are answered.

In general, most of the anomalies and trends that can be detected by our multi-modal approach can be also be found using process mining or other types of algorithmic data processing, at least if one knows what one is looking for. However, the same argument can be brought forward concerning single-modal visualization, and tools such as *Dotted Chart analysis* are amongst the most used ProM plugins (Claes and Poels, 2013). We are not suggesting multi-modal approaches as a replacement for the state of the art, neither for process mining and other algorithmic processing, nor for visual analytics. Instead, we merely propose complementing and enhancing current approaches with sonification, which we believe will help users in detecting additional features in their process logs.

As sonification is not yet widespread, outside of a few areas such as medical science, the typical user will not have experiences in using sonification for data analysis. Thus, it can be expected that most users will need a training period to familiarize themselves with our prototype, but also with the concept of sonification in general, before being productive. Based on findings from research, it can be expected that this training period may be shorter for people with musical and/or acoustic training. Furthermore, we expect productivity to be greater when the user is not only accustomed to our tool, but also to how his/her processes *sound*. Thus, if a user analyzes the events of the same process regularly, he/she will be even more efficient in detecting deviations and irregularities after a while. This is because in the beginning, the user has no comparison to *normal* behavior.

While the sonification of individual events seems to work well for discovering general trends, we also plan to enhance our approach to visualize and sonify the development of KPIs and other trace-spanning measurements over time. This might require other types of sonification (such as continuous sound streams instead of singular sound events) and mapping techniques, which will be analyzed. We further plan to enhance our system by extending the concepts that can be mapped to visual and acoustic properties.

In general, even though the presented use case evaluation showed the usefulness of our approach, this of course can serve only as a preliminary study and cannot produce any representative results. Therefore, more formal long term evaluations with professional data analysts are planned.

Finally, even though we expect the main benefits of multi-modal approaches to be in areas such as anomaly detection, future evaluations will also analyze their potential for process discovery. Although the prototype itself is built to analyze business process data, parts of the design concept and sonification techniques may be transferable to the analysis of other data of similar structure (i.e. event-based data that is mostly qualitative).

7. Summary and Discussion

The ability to monitor business processes in real-time, as well as to analyze historical business process execution data, is becoming increasingly crucial. On the one hand, more and more processes are becoming automated, making it important to anticipate and prevent critical situations in a timely manner. On the other hand, the ability to find patterns and trends in the growing amount of process execution data might give companies a competitive edge over their competitors. The tasks of real-time monitoring of business process data and retrospective analysis of process execution data both benefit greatly from, and rely on, visualization and visual analytics. Even though many information needs of the users can be answered with such approaches, several open challenges remain, such as the inability to focus on other work while performing process monitoring as a second task, or the limited number of visual dimensions onto which data can be mapped. A combination of visualization and visual analytics with sonification (the presentation of data using non-speech sound) might alleviate some of these challenges, like the need for constant visual focus. Therefore, this thesis mainly deals with the question of whether the combination of visualization and sonification to form multi-modal displays can alleviate some of those challenges.

Even though sonification has already been applied for process monitoring and data analysis, it had not previously been explicitly applied for these use cases in the business process domain. Therefore, several open questions and research challenges remained; for example concerning how to effectively combine visualization and sonification to form multi-modal systems for business process monitoring and analysis, or which sonification techniques and mappings are best suited for the tasks and requirements at hand. Business process execution data is peculiar and different than that of most other domains, as it is mainly qualitative and structured by discrete events.

To answer some of these questions, the SoProMon system, a modular evaluation framework for multi-modal sonifications for process monitoring as a second task, has been developed. A quantitative and qualitative experiment based on this system was devised, which aimed to test the SoProMon framework to find out how different types of sonification can help in guiding attention in a dual-task scenario for process monitoring compared to a visual-only system.

Subsequently, an approach for the multi-modal analysis of historical process execution data was presented. This approach was implemented in the form of a software prototype, which combines methods from visual analytics and sonification to form a multi-modal, audiovisual analytics framework. The prototype was implemented as a plugin for the process mining framework ProM.

7.1. Reflection

In the following, the research questions stated in the introduction are answered:

> What are the peculiarities of business process execution data and what implications do these have with regard to sonification?

What makes business process execution data different from many other types of data is the fact that it is event-based, and primarily of qualitative nature. Unlike, for example, data stemming from technical processes, which are often based on continuous, quantitative data, business process execution data is primarily based on discrete events. Therefore, unlike in other data that is primarily quantitative, parameter mapping is only suited for certain aspects of the data, while auditory icons and earcons are more suitable than they are for the analysis of other types of data.

> Which user requirements exist concerning multi-modal solutions during the process operation and analysis phases?

The main user requirements are that sonifications for process monitoring should not distract users from their main tasks, and that they should be adaptable to their preferences. Approaches for real-time monitoring need to take the type and frequency of data into account, and should be unobtrusive during normal operation but attention-grabbing when necessary.

For process analysis, it is less crucial for multi-modal systems to be unobtrusive as, unlike process monitoring, analysis is often performed with users' full attention and typically not for a complete work day. However, studies have shown that sonifications are more likely to be accepted by users when the sounds are aesthetically pleasing. As this might differ for every user, it is recommended to enable the user to customize his/her sonification.

> What implications do these requirements have on possible sonifications in this area?

For peripheral process monitoring in particular, it is important to find unobtrusive ways to convey process data aurally. This is especially crucial, as several different user groups, most of whom perform monitoring as a second task in the background, have different information needs. Musical sonifications are often considered to be pleasing, but depending on their design, especially for real-time monitoring they can become intrusive after a while. Sonifications based on nature-based soundscapes, like forest sounds, are often less intrusive, but they also typically allow less data to be conveyed simultaneously.

> Which sonification techniques and mappings are best suited to the requirements defined for real-time data analysis and the analysis of historical instance execution data?

Depending on the context, parameterized auditory icons or earcons (especially those based on musical motifs) combined with parameter mapping seem beneficial in particular. Parameterized earcons seem to be suitable to sonify individual events. For process analysis, in such a way it is possible to obtain an overview of large log files in a short amount of time, discover general trends over time, or find anomalies in individual traces in smaller log files, or in those with a low event density. However, the sonification (and visualization) of individual events in large log files has certain limitations, considering the number of overlaying sound events. For peripheral process monitoring in particular, a sonification of individual process events might become intrusive, depending on the frequency of occuring events. Instead, a mapping of aggregated KPIs and quantitative parameters to continuous soundscapes might, in combination with alert notifications, be more sensible for many use cases.

> What are the strengths and weaknesses of visualization and sonification techniques during the operation and analysis phases?

For the operation phase, the two modalities complement each other well - the strength of visualization is that it can be more easily ignored than sonification, when concentrating on other tasks. On the other hand, sonification is very good at attracting attention when desired, which is a challenge of visualization. For multi-modal sonification as a second task, it seems that a trade-off between main task and process monitoring task has to be made. If the main task is of the highest importance while process monitoring can be neglected, visualization in combination with auditory alerts seems to be the best fitting mode, as users seem to achieve the best main task performance under this condition (but also the lowest monitoring performance). On the other hand, a continuous, multi-modal sonification seems to be the most suitable mode for situations in which process monitoring is still important, but it is not absolutely crucial that the users are not to be disturbed during their main task. It leads to slightly lower main task performances than visual monitoring enhanced with auditory alerts, but leads to the highest process monitoring results. A visual-only solution, on the other hand, seems to be unsuitable for most cases, as it leads to lower main task performances and a significantly lower process monitoring effectiveness than a continuous multi-modal sonification.

In the analysis phase, the strength of visualization is its ability to present detailed data and concrete information, while it is hard to detect trends over time. For sonification, it is generally the other way round. In general, as demonstrated in the use case evaluation, multi-modal approaches in this domain ideally couple both modalities tightly, such that changes in the sonification mapping are directly applied to the visual part. Such systems should further allow the user to customize the visual and acoustic mappings as easily and directly as possible, in order to allow the user to quickly try out different mappings and settings.

> How can visualization and sonification techniques best be combined to support users during the real-time monitoring and post-hoc analysis of business process execution data?

For (peripheral) real-time monitoring, a combination that works in many cases involves peripheral sonification and a visual or multi- modal display. The peripheral sonification enables users to identify or anticipate undesired or unusual process states, while the visual or multi-modal display allows users to conduct root-cause analysis to find out exactly what has happened, why it happened and its exact effects. Therefore, a good approach for peripheral monitoring could be to visualize all data while enabling the user to select filters and generally navigate through the data. Sonification should be used as little as possible, but as much as necessary to ensure the user is fully aware of system status and able to anticipate failures. The amount of data conveyed aurally depends on both the use case and the user. In some cases, this can mean sonifying all occurring events; in others, sonifying only a few carefully selected parameters.

For direct monitoring and retrospective analysis, comprehensive studies concerning which aspects are most effectively conveyed by which modality are still lacking. However, the presented prototype for multi-modal process analysis - in which the user chooses what to visualize and sonify, and how to do so - might be a suitable strategy for most situations. Every user might have different preferences for what to visualize and what to sonify, along with different perceptual abilities. Preliminary user feedback suggests that there may not be data types that can only be effectively conveyed aurally, but that both modalities together may help to get a *feel* for the data faster. Further studies in this regard will have to be conducted.

> Do multi-modal solutions better match the requirements than visualization or sonification alone during the phases *operation* and *analysis*?

For the operation phase, at least for peripheral monitoring as second task, a significant advantage of multi-modal approaches over *pure* visualization has been proven. An additional continuous, event-based sonification significantly enhances process monitoring performance compared to visualization alone, or to a system that combines visualization with auditory alerts that are played upon reaching a critical state. At the same time, additional auditory information does not negatively effect the performance in the main task in a significant way. The participants also found multi-modal systems with continuous sonification significantly more helpful for monitoring as a second task than purely visual monitoring systems, or those that combine visualization with auditory alerts.

However, for the presented sonification design, a strong polarization was observed - while a few of the participants perceived the presented continuous sonification as intrusive and distracting, and would prefer a condition that combines visualization with more advanced auditory alerts, others preferred the presented continuous multi-modal sonification. Participants did not feel significantly more in control using the continuous multi-modal sonification, but it was also not deemed to be significantly more exhausting than the other conditions.

For the analysis phase, such effects can only be assumed thus far; however, concrete studies in this regard still have to be conducted. Each modality is better suited to convey different types of data features. Furthermore, user perceptions concerning different modalities can vary, depending on each user's training and abilities. Therefore, it can safely be assumed that a combination of both modalities is usually preferable to visualization alone.

Tasks that seem to suggest themselves for multi-modal analysis include obtaining overviews over large event logs, discovering general trends and large scale anomalies, and comparing selected individual events and traces for anomaly detection and root cause analysis. Limitations of both visualization and sonification have been discovered through the analysis of small-scale anomalies in large amounts of unfiltered event data.

7.2. Transfer of Results

Although the approaches presented in this thesis were geared towards business process execution data, and specifically towards the tasks of real-time monitoring and analysis of historical data, many of the results might be transferable to use cases of other domains with similar data structure and/or tasks. As mentioned previously, there are several domains in which data of similar structure exists, i.e.

data that is mostly event-based - like computer program executions, computer security or web server log files. Furthermore, as the results showed that sonification is particularly suitable for the task of real-time monitoring, many of the results can be transferable to other use cases of real-time monitoring. This includes, for example, all use cases based on a control center, such as power plants, space operations, electric grids or traffic and transit operations, including smart city control centers. Situations in which monitoring is performed as a second task, such as network or computer security monitoring, or monitoring of cyber-physical systems, could especially benefit from the results of this thesis. Pertinent results here could include the trade-off between information that can be conveyed, and potential distraction from the main task; for example, the fact that a purely visual system may rarely be the best choice. Instead, systems promising the best results for most use cases feature either continuous sonification or fine-grained auditory alerts or alarms designed to inform users even before a critical situation occurs.

7.3. Limitations and Future Work

A detailed analysis of the limitations of the research and its results has been presented for the operation (Sec. 5.5.2) and analysis phases (Sec. 6.3), resulting in detailed plans for future work in the respective areas. For the monitoring phase, one main limitation is that the results have been obtained in a laboratory experiment, and not under real-life working conditions. Therefore, future studies are planned, and will allow potential users to evaluate the approaches for a longer period of time in realistic settings. The conducted experiment has further shown that the presented conditions featuring audio have certain flaws; therefore, further studies featuring improved conditions will be conducted.

One main limitation in the analysis prototype is its concentration on individual events. Further improvements of the prototype will, therefore, include sonifications of high-level parameters, such as KPIs. Another main limitation is the lack of user evaluations to test acceptance and effectiveness. Such evaluations will be conducted in the form of focus groups (and potentially quantitative experiments) with both sonification and process analysis experts. The results will help determine if additional sonification provides a benefit over *traditional* visual analytics. The existence of such studies is expected to greatly increase the acceptance of multi-modal sonification in the BPM community, and perhaps subsequently lead to its adoption in organizations. Such tests will also help researchers investigating sonification for data of similar structure or for similar use cases, as it seems that the majority of sonification mappings and techniques created by researchers have not been formally evaluated yet. A wider adoption of such evaluation methods could lead to best practice catalogs of sonification techniques and mappings for many domains and data types.

Bibliography

Aalst, Wil M. P. van der (2011). *Process Mining - Discovery, Conformance and Enhancement of Business Processes*. Springer. ISBN: 978-3-642-19344-6. URL: http://dx.doi.org/10.1007/978-3-642-19345-3 (cit. on pp. 20, 22, 29).

Aalst, Wil M. P. van der (2012). "A Decade of Business Process Management Conferences: Personal Reflections on a Developing Discipline." In: *Business Process Management*. Ed. by Alistair Barros, Avigdor Gal, and Ekkart Kindler. Lecture Notes in Computer Science 7481. Springer Berlin Heidelberg, pp. 1–16. ISBN: 978-3-642-32884-8 978-3-642-32885-5. URL: http://link.springer.com/chapter/10.1007/978-3-642-32885-5_1 (visited on 08/12/2013) (cit. on p. 3).

Aalst, Wil et al. (2012). "Process Mining Manifesto." In: *Business Process Management Workshops*. Ed. by Wil Aalst et al. Vol. 99. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 169–194. ISBN: 978-3-642-28107-5 978-3-642-28108-2. URL: http://www.springerlink.com/content/q8446vr556p28n20/ (visited on 02/21/2012) (cit. on pp. 4, 13, 14, 20, 22).

Alty, James L. (1995). "Can we use music in computer-human communication?" In: *Proceedings of the HCI'95 conference on People and computers X*. HCI '95. New York: Cambridge University Press, pp. 409–423. ISBN: 0-521-56729-7. (Visited on 07/04/2012) (cit. on p. 36).

Amerbauer, Felix (2017). "Multi-modal Sonification of Business Process Event Logs." MA thesis. Austria: University of Vienna (cit. on p. 11).

Bakker, Saskia, Elise van den Hoven, and Berry Eggen (2010). "Exploring Interactive Systems Using Peripheral Sounds." en. In: *Haptic and Audio Interaction Design*. Lecture Notes in Computer Science 6306. Springer Berlin Heidelberg, pp. 55–64. ISBN: 978-3-642-15840-7 978-3-642-15841-4. (Visited on 06/24/2014) (cit. on p. 37).

Ballora, Mark, Robert J. Cole, et al. (2012). "Use of sonification in the detection of anomalous events." In: vol. 8407, 84070S. DOI: 10.1117/12.918688. URL: http://adsabs.harvard.edu/abs/2012SPIE.8407E..0SB (visited on 04/27/2016) (cit. on p. 46).

Ballora, Mark, Nicklaus A. Giacobe, and David L. Hall (2011). "Songs of cyberspace: an update on sonifications of network traffic to support situational awareness." In: *Proc. SPIE 8064, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2011, 80640P (7 June 2011)*. Vol. 8064, 80640P–80640P–6. DOI: 10.1117/12.883443. URL: http://dx.doi.org/10.1117/12.883443 (visited on 03/18/2015) (cit. on pp. 39, 40).

Ballora, Mark and David L. Hall (2010). "Do you see what I hear: experiments in multi-channel sound and 3D visualization for network monitoring?" In: vol. 7709, 77090J–77090J–7. DOI: 10.1117/12.850319. URL: http://dx.doi.org/10.1117/12.850319 (visited on 03/18/2015) (cit. on pp. 39–41).

Ballora, Mark, Brian Panulla, et al. (2010). "Preliminary Steps in Sonifying Web Log Data." In: Washington: International Community for Auditory Display (cit. on pp. 36, 39, 40, 46).

Barra, M. et al. (2002). "Multimodal monitoring of Web servers." In: *Multimedia, IEEE* 9.3, pp. 32–41. ISSN: 1070-986X. DOI: 10.1109/MMUL.2002.1022857 (cit. on pp. 39–41, 44).

Berman, L. I. and K. B. Gallagher (2006). "Listening to program slices." In: *Proceedings of the 12th International Conference on Auditory Display (ICAD2006)*. London, UK: Department of Computer Science, Queen Mary, University of London, UK, pp. 172–175. URL: Proceedings/2006/BermanGallagher2006.pdf (cit. on p. 44).

Berman Irwin, Lewis (2011). "Program Comprehension Through Sonification." PhD thesis. Durham University. URL: http://etheses.dur.ac.uk/1396/ (cit. on p. 44).

Bezerra, Fabio, Jacques Wainer, and W. M. P. van der Aalst (2009). "Anomaly Detection Using Process Mining." en. In: *Enterprise, Business-Process and Information Systems Modeling*. Ed. by Terry Halpin et al. Lecture Notes in Business Information Processing 29. DOI: 10.1007/978-3-642-01862-6_13. Springer Berlin Heidelberg, pp. 149–161. ISBN: 978-3-642-01861-9 978-3-642-01862-6. URL: http://link.springer.com/chapter/10.1007/978-3-642-01862-6_13 (visited on 02/06/2017) (cit. on p. 21).

Biesheuvel, Simon (1949). *Psychological tests and their application to non-European peoples*. en. Evans (cit. on p. 27).

Bly, Sarah (1982). *Sound and Computer Information Presentation (PhD Thesis)*. University of California (cit. on p. 30).

Boardman, D.B. et al. (1995). "LISTEN: A tool to investigate the use of sound for the analysis of program behavior." In: *Computer Software and Applications Conference, 1995. COMPSAC 95. Proceedings., Nineteenth Annual International*, pp. 184–189. DOI: 10.1109/CMPSAC.1995.524778 (cit. on p. 44).

Bobrik, Ralph and Thomas Bauer (2007). "Towards Configurable Process Visualizations with Proviado." In: *Proc. of the 16th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE 2007*, pp. 367–369 (cit. on p. 28).

Bock, D. S. (1994). "ADSL: An auditory domain specification language for program auralization." In: *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Ed. by G. Kramer and S. Smith. Santa Fe, NM, U.S.: International Community for Auditory Display, pp. 251–256. URL: Proceedings/1994/Bock1994.pdf (cit. on p. 44).

Brock, D, JA Ballas, and JL Stroup (2002). "Using an auditory display to manage attention in a dual task, multiscreen environment." In: *Proceedings of the Eighth International Conference on Auditory Display*. Kyoto, Japan (cit. on p. 37).

Brock, D., B. McClimens, and M. McCurry (2010). "Virtual Auditory cueing revisited." In: *Proceedings of the 16th International Conference on Auditory Display*. Washington, USA. URL: http://www.nrl.navy.mil/itd/aic/sites/www.nrl.navy.mil.itd.aic/files/pdfs/brock-etal-icad2010.pdf (cit. on p. 37).

Brown, A. et al. (2009). "Poster: Towards Music-Assisted Intrusion Detection." In: Oakland, USA. URL: http://faculty.uoit.ca/kapralos/publications/ieee_ssp2009.pdf (visited on 03/18/2015) (cit. on p. 40).

Caldwell, Barrett S. and Jacob E. Viraldo (2014). "Improving Control Room State Awareness through Complex Sonification Interfaces." en. In: *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*. Ed. by Sakae Yamamoto. Lecture Notes in Computer Science 8522. Springer International Publishing, pp. 317–323. ISBN: 978-3-319-07862-5 978-3-319-07863-2. (Visited on 07/07/2014) (cit. on p. 37).

Campo, Alberto de (2007). "Toward a data sonification design space map." In: *Proceedings of the 13th International Conference on Auditory Display (ICAD2007)*. Ed. by Gary P. Scavone. Montreal, Canada: Schulich School of Music, McGill University, pp. 342–347. URL: Proceedings/2007/deCampo2007.pdf (cit. on p. 66).

Chambers, J M, M V Mathews, and F R Moore (1974). *Auditory data inspection, Technical Memorandum 74-1214-20* (cit. on p. 30).

Ciardi, C. (2004). "sMax: A multimodal toolkit for stock market data sonification." In: *Proc. of the 10th International Conference on Auditory Display (ICAD2004)*. Ed. by S. Barras and P. Vickers. International Community for Auditory Display (ICAD) (cit. on p. 29).

Claes, Jan and Geert Poels (2013). "BPM 2012 International Workshops. Revised Papers." In: ed. by Marcello La Rosa and Pnina Soffer. Springer Berlin Heidelberg. Chap. Process Mining and the ProM Framework: An Exploratory Survey, pp. 187–198. ISBN: 978-3-642-36285-9. DOI: 10.1007/978-3-642-36285-9_19 (cit. on p. 127).

Cohen, J. (1992). "Monitoring background activities." In: *Auditory Display*. Ed. by G. Kramer. New York: Addison Wesley Longman, pp. 499–531 (cit. on p. 43).

Cohen, J. (1994). "Out to lunch: Further adventures monitoring background activity." In: *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Ed. by G. Kramer and S. Smith. Santa Fe, NM, U.S.: International Community for Auditory Display, pp. 15–20. URL: Proceedings/1994/Cohen1994a.pdf (cit. on pp. 37, 43, 44).

Crawford, Jennifer et al. (2002). "Multimodal displays for anaesthesia sonification: Timesharing, workload, and expertise." In: *PROCEEDINGS OF THE ESA/CHISIG CONFERENCE ON HUMAN FACTORS (HF2002*, pp. 27–29 (cit. on p. 38).

Csapo, Adam and Peter Baranyi (2012). "A Unified Terminology for the Structure and Semantics of CogInfoCom Channels." In: *Acta Polytechnica Hungarica* 9, pp. 85–105. ISSN: 17858860 (cit. on p. 77).

Cullen, Charlie and E. Coyle (2005). "TrioSon: A Graphical User Interface for Pattern Sonification." In: *Proc. of the 11th International Conference on Auditory*

Display. International Community for Auditory Display (ICAD) (cit. on pp. 46, 116).

Cullen, Charlie and Eugene Coyle (2006). "Musical Pattern Design Using Contour Icons." In: vol. International Computer Music Conference Proceedings. Louisiana, USA: Ann Arbor, Michigan: Scholarly Publishing Office, University of Michigan University Library. URL: http://arrow.dit.ie/dmccon/21 (cit. on p. 64).

deButts, Benjamin (2014). "Network Access Log Visualization & Sonification." Student Project. Medford, USA: Tufts University. URL: https://tuftsdev.github.io/DefenseOfTheDarkArts/students_works/final_project/fall2014/bdebutts.pdf (cit. on pp. 40, 46).

DiGiano, Christopher J., Ronald M. Baecker, and Russell N. Owen (1993). "LogoMedia: a sound-enhanced programming environment for monitoring program behavior." In: *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*. CHI '93. New York, NY, USA: ACM, pp. 301–302. ISBN: 0-89791-575-5. DOI: 10.1145/169059.169229. (Visited on 07/20/2012) (cit. on p. 44).

Eggen, B. et al. (2008). "Soundscapes at Workspace Zero–Design Explorations into the Use of Sound in a Shared Environment." In: *Proc. Sixth Int. Conf. on Pervasive Computing* (cit. on p. 43).

Fairfax, Tom, Christopher Laing, and Paul Vickers (2015). "Network Situational Awareness: Sonification and Visualization in the Cyber Battlespace." In: *Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance:* ed. by Maria Manuela Cruz-Cunha. IGI Global. ISBN: 978-1-4666-6324-4 978-1-4666-6325-1. URL: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-6324-4 (visited on 03/18/2015) (cit. on p. 40).

Francioni, Joan M., Larry Albright, and Jay Alan Jackson (1991). "Debugging parallel programs using sound." In: *SIGPLAN Not.* 26.12, pp. 68–75. ISSN: 0362-1340. DOI: 10.1145/127695.122765. (Visited on 09/10/2012) (cit. on p. 45).

Frauenberger, C. (2006). "Sonifikation: Stärken, Schwächen und Anwendungen. Kritik zum Artikel von C. Dayé, M. Egger de Campo und A. de Campo." In: *Angewandte Sozialforschung* 24, pp. 56–60 (cit. on p. 27).

Gadatsch, Andreas (2009). "Integriertes Geschäftsprozess- und Workflow-Management." de. In: *HMD Praxis der Wirtschaftsinformatik* 46.2, pp. 35–42. ISSN: 1436-3011, 2198-2775. DOI: 10.1007/BF03340341. URL: http://link.springer.com/article/10.1007/BF03340341 (visited on 01/11/2017) (cit. on p. 3).

Gaver, William W. (1989). "The SonicFinder: an interface that uses auditory icons." In: *Hum.-Comput. Interact.* 4.1, pp. 67–94. ISSN: 0737-0024. DOI: 10.1207/s15327051hci0401_3. (Visited on 02/13/2013) (cit. on p. 44).

Gaver, William W., Randall B. Smith, and Tim O'Shea (1991). "Effective sounds in complex systems: the ARKOLA simulation." In: *Proc. of the SIGCHI conference on Human factors in computing systems: Reaching through technology (CHI'91)*. ACM, pp. 85–90 (cit. on pp. 36, 43, 44).

Gaver, William, Thomas Moran, et al. (1992). "Realizing a video environment: EuroPARC's RAVE system." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. New York, NY, USA: ACM, pp. 27–35. ISBN: 0-89791-513-5. DOI: 10.1145/142750.142754. (Visited on 02/06/2013) (cit. on p. 43).

Gilfix, Michael and Alva L. Couch (2000). "Peep (The Network Auralizer): Monitoring Your Network with Sound." In: *Proceedings of the 14th USENIX Conference on System Administration*. LISA '00. Berkeley, CA, USA: USENIX Association, pp. 109–118. URL: http://dl.acm.org/citation.cfm?id=1045502.1045526 (visited on 03/18/2015) (cit. on pp. 36, 40, 41, 44, 56).

Giot, Rudi and Yohan Courbe (2012). "InteNtion – Interactive Network Sonification." In: *Proceedings of the 18th International Conference on Auditory Display*. Atlanta, GA, USA: The International Community for Auditory Display, pp. 235–236. URL: https://smartech.gatech.edu/handle/1853/44419 (visited on 03/18/2015) (cit. on pp. 40, 41).

Gopinath, M. (2004). "Auralization of Intrusion Detection System using JListen." Master's Thesis. Pilani (Rajasthan), India: Birla Institute of Technology and Science. URL: https://www.cs.purdue.edu/homes/apm/listen/JListen-1.0-BITS/Gopinath/MS-thesis/thesis-final.doc (cit. on pp. 40–42).

Grigori, Daniela et al. (2004). "Business Process Intelligence." In: *Computers in Industry* 53.3, pp. 321–343. ISSN: 0166-3615. DOI: 10.1016/j.compind.2003.10.007. URL: http://www.sciencedirect.com/science/article/pii/S0166361503001994 (visited on 02/14/2012) (cit. on p. 14).

Grossmann, Wilfried and Stefanie Rinderle-Ma (2015). *Fundamentals of Business Intelligence*. Data-Centric Systems and Applications. Springer. ISBN: 978-3-662-46530-1. URL: http://dx.doi.org/10.1007/978-3-662-46531-8 (cit. on p. 28).

Gulden, Jens (2014). "Towards a Generalized Notion of Audio as Part of the Concrete Syntax of Business Process Modeling Languages." en. In: *Business Process Management Workshops*. Ed. by Fabiana Fournier and Jan Mendling. Lecture Notes in Business Information Processing. DOI: 10.1007/978-3-319-15895-2_35. Springer International Publishing, pp. 420–425. ISBN: 978-3-319-15894-5 978-3-319-15895-2. URL: http://link.springer.com/chapter/10.1007/978-3-319-15895-2_35 (visited on 01/13/2017) (cit. on p. 6).

Handfield, Robert et al. (2014). *Key Trends Shaping the Global Logistics Environment*. Literature Series: Supply Chain Management. BVL International (cit. on p. 1).

Hao, Ming C. et al. (2006). "Business process impact visualization and anomaly detection." In: *Information Visualization* 5.1, pp. 15–27. ISSN: 1473-8716 (cit. on p. 28).

Hauer, Wolfgang and Katharina Vogt (2015). "Sonification of a Streaming-Server Logfile." In: *International Conference on Auditory Display (ICAD) 2015*. Proceedings of the 21st International Conference on Auditory Display. Graz, Austria, pp. 315–316. ISBN: 978-3-902949-01-1. URL: http://iem.kug.ac.at/fileadmin/media/institut-17/proceedings150707_3.pdf (cit. on p. 42).

Hermann, Thomas (2002). *Sonification for Exploratory Data Analysis (PhD Thesis)*. Universität Bielefeld (cit. on pp. 5, 29, 30, 32, 33).

Hermann, Thomas, Tobias Hildebrandt, et al. (2015). "Optimizing Aesthetics and Precision in Sonification for Peripheral Process-Monitoring." In: *International Conference on Auditory Display (ICAD) 2015*. Proceedings of the 21st International Conference on Auditory Display. Graz, Austria, pp. 317–318. ISBN: 978-3-902949-01-1. URL: http://eprints.cs.univie.ac.at/4395/ (cit. on pp. 10, 11, 70).

Hermann, Thomas, Andy Hunt, and John G. Neuhoff (2011). *The Sonification Handbook*. Logos Berlin. ISBN: 3-8325-2819-9 (cit. on pp. 31, 32).

Hermann, Thomas, Christian Niehus, and Helge Ritter (2003). "Interactive visualization and sonification for monitoring complex processes." In: *Proceedings of the 2003 International Conference on Auditory Display*. ICAD. International Community for Auditory Display (ICAD) (cit. on p. 36).

Hevner, Alan R. et al. (2004). "Design Science in Information Systems Research." In: *MIS Q.* 28.1, pp. 75–105. ISSN: 0276-7783. URL: http://dl.acm.org/citation.cfm?id=2017212.2017217 (visited on 01/09/2017) (cit. on p. 7).

Hildebrandt, Tobias (2013). "Towards Enhancing Business Process Monitoring with Sonification." In: *Intl. Workshop on Theory and Applications of Process Visualization (TAProViz 2013)*. Lecture Notes in Business Information Processing. Berlin, Germany: Springer (cit. on pp. 11, 70).

Hildebrandt, Tobias, Felix Amerbauer, and Stefanie Rinderle-Ma (2016). "Combining Sonification and Visualization for the Analysis of Process Execution Data." In: *2016 IEEE 18th Conference on Business Informatics (CBI)*. Vol. 02, pp. 32–37. DOI: 10.1109/CBI.2016.47 (cit. on pp. 11, 113).

Hildebrandt, Tobias, Thomas Hermann, and Stefanie Rinderle-Ma (2014). "A sonification system for process monitoring as secondary task." In: *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 191–196. DOI: 10.1109/CogInfoCom.2014.7020444 (cit. on pp. 1, 10, 70).

Hildebrandt, Tobias, Thomas Hermann, and Stefanie Rinderle-Ma (2016). "Continuous sonification enhances adequacy of interactions in peripheral process monitoring." In: *International Journal of Human-Computer Studies* 95, pp. 54–65. ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2016.06.002. URL: http://www.sciencedirect.com/science/article/pii/S107158191630074X (cit. on pp. 10, 11, 70).

Hildebrandt, Tobias, Simone Kriglstein, and Stefanie Rinderle-Ma (2012a). "Beyond Visualization: On Using Sonification Methods to Make Business Processes More Accessible to Users." In: *18th International Conference on Auditory Display (ICAD 2012)*. Georgia Institute of Technology, pp. 248–249 (cit. on pp. 10, 49).

Hildebrandt, Tobias, Simone Kriglstein, and Stefanie Rinderle-Ma (2012b). "On Applying Sonification Methods to Convey Business Process Data." In: *CaISE 2012 Forum*. CaISE Forum. CEUR. URL: http://eprints.cs.univie.ac.at/3449/ (cit. on pp. 9, 49).

Hildebrandt, Tobias, Jürgen Mangler, and Stefanie Rinderle-Ma (2014). "Something doesn't sound right: Sonification for monitoring business processes in manufacturing." In: *LABEM 2014 : 1st International Workshop: Lowering the Adoption Barrier of Enterprise Modelling*. Geneva, Switzerland: IEEE. URL: http://eprints.cs.univie.ac.at/4066/ (cit. on pp. 10, 49).

Hildebrandt, Tobias and Stefanie Rinderle-Ma (2013). "Toward a Sonification Concept for Business Process Monitoring." In: *19th International Conference on Auditory Display (ICAD 2013)*. Lodz, pp. 323–330 (cit. on pp. 9, 25, 49).

Hildebrandt, Tobias and Stefanie Rinderle-Ma (2015). "Server sounds and network noises." In: *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pp. 45–50. DOI: 10.1109/CogInfoCom.2015.7390562 (cit. on pp. 9, 25).

Horton, Richard (2015). *The robots are coming - A Deloitte Insight report*. Tech. rep. Deloitte. URL: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/finance/deloitte-uk-finance-robots-are-coming.pdf (cit. on p. 1).

Hudson, Scott E. and Ian Smith (1996). "Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems." In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. CSCW '96. New York, NY, USA: ACM, pp. 248–257. ISBN: 0-89791-765-0. DOI: 10.1145/240080.240295. URL: http://doi.acm.org/10.1145/240080.240295 (visited on 07/04/2012) (cit. on p. 63).

Jameson, D. H. (1994). "The run-time components of sonnett." In: *Proceedings of the 2nd International Conference on Auditory Display (ICAD94)*. Ed. by G. Kramer and S. Smith. Santa Fe, NM, U.S.: International Community for Auditory Display, pp. 241–250. URL: Proceedings/1994/Jameson1994.pdf (cit. on p. 44).

Kilander, F. and P. Loennqvist (2002). "A whisper in the woods - an ambient soundscape for peripheral awareness of remote processes." In: *Proceedings of the 8th International Conference on Auditory Display (ICAD2002)*. Kyoto, Japan (cit. on pp. 37, 43).

Kimoto, Masahiko and Hiroyuki Ohno (2002). "Design and Implementation of Stetho—Network Sonification System." In: *Proceedings of the 2002 International Computer Music Conference*, pp. 273–279 (cit. on pp. 40, 41).

Knuth, Donald E. (2011). *BCS/IET Turing lecture*. University of Manchester (cit. on p. 36).

Kramer, Gregory (1994). *Auditory Display: Sonification, Audification, And Auditory Interfaces*. en. Google-Books-ID: yzCEQgAACAAJ. Avalon Publishing. ISBN: 978-0-201-62604-9 (cit. on p. 4).

Kramer, G. et al. (1999). *Sonification Report: Status of the Field and Research Agenda - Report prepared for the National Science Foundation by members of the International Community for Auditory Display*. International Conference on Auditory Display (cit. on pp. 4, 25, 28).

Kriglstein, Simone, Margit Pohl, et al. (2016). "Visual Analytics in Process Mining: Classification of Plug-ins of the ProM Framework." In: *EuroVA*. (accepted) (cit. on pp. 22, 29).

Kriglstein, Simone, Günter Wallner, and Stefanie Rinderle-Ma (2013). "A Visualization Approach for Difference Analysis of Process Models and Instance Traffic." In: *Int'l Conference on Business Process Management 2013*. Germany: Springer. URL: https://eprints.cs.univie.ac.at/3739/ (visited on 11/19/2013) (cit. on p. 28).

Lewis, Clayton and John Rieman (1993). *Task-centered User Interface Design: A Practical Introduction*. en. University of Colorado, Boulder, Department of Computer Science. URL: http://hcibib.org/tcuid/tcuid.pdf (cit. on p. 84).

Malandrino, Delfina et al. (2003). "NeMoS: Network monitoring with sound." In: *Proceedings of the 9th International Conference on Auditory Display (ICAD2003)*. Boston, USA: International Community for Auditory Display. URL: https://smartech.gatech.edu/handle/1853/50475 (visited on 03/18/2015) (cit. on p. 40).

McClimens, B., J. L. Stroup, et al. (2004). *The design of mixed-use virtual auditory displays: Recent findings with a dual-task paradigm*. Proceedings. URL: https://smartech.gatech.edu/handle/1853/50920 (visited on 06/02/2015) (cit. on p. 37).

McClimens, Brian and Derek Brock (2010). "Sonification of NRL Dual-Task Data." In: *16th International Conference on Auditory Display (ICAD2010)*. Washington, D.C., USA (cit. on p. 37).

McClimens, Brian, Derek Brock, et al. (2011). *Modeling Attention Allocation in a Complex Dual Task With and Without Auditory Cues* (cit. on p. 37).

Miguel A Garcia-Ruiz, Miguel Vargas (2006). "Towards a multimodal human-computer interface to analyze intrusion detection in computer networks." In: *Proceedings of the First Human- Computer Interaction Workshop (MexIHC)*. Puebla, Mexico: University of the Americas (cit. on pp. 40, 41).

Muñoz-Arteaga, Jaime et al. (2009). "A Method to Design Information Security Feedback Using Patterns and HCI-Security Criteria." en. In: *Computer-Aided Design of User Interfaces VI*. Ed. by Victor Lopez Jaquero et al. London: Springer London, pp. 283–294. ISBN: 978-1-84882-205-4 978-1-84882-206-1. URL: http://link.springer.com/10.1007/978-1-84882-206-1_26 (visited on 03/18/2015) (cit. on pp. 40, 41).

Papadopoulos, Christos et al. (2004). "CyberSeer: 3D Audio-visual Immersion for Network Security and Management." In: *Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security*. VizSEC/DMSEC '04. New York, NY, USA: ACM, pp. 90–98. ISBN: 1-58113-974-8. DOI: 10.1145/1029208.1029223. URL: http://doi.acm.org/10.1145/1029208.1029223 (visited on 03/19/2015) (cit. on pp. 40, 41).

Petzmann, Alexander et al. (2007). "Applying MDA Concepts to Business Process Management." In: *2007 BPM & Workflow Handbook*. Ed. by Layna Fischer (ed.) 1st. Future Strategies Inc., ISBN: 0-9777527-1-2 (cit. on p. 18).

Poguntke, M. and K. Ellis (2008). "Auditory attention control for human-computer interaction." In: *2008 Conference on Human System Interactions*, pp. 231–236. DOI: 10.1109/HSI.2008.4581439 (cit. on p. 38).

Pollack, I. and L. Ficks (1954). "The Information of elementary multimensional auditory displays." In: *The Journal of the Acoustical Society of America* 26.1, pp. 155–158 (cit. on p. 30).

Qi, Lei et al. (2007). "Toward Sound-Assisted Intrusion Detection Systems." en. In: *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS*. Ed. by Robert Meersman and Zahir Tari. Lecture Notes in Computer Science 4804. Springer Berlin Heidelberg, pp. 1634–1645. ISBN: 978-3-540-76835-7 978-3-540-76843-2. URL: http://link.springer.com/chapter/10.1007/978-3-540-76843-2_36 (visited on 03/18/2015) (cit. on p. 40).

Rauterberg, Matthias and Erich Styger (1994). "Positive Effects of Sound Feedback During the Operation of a Plant Simulator." In: *Selected papers from the 4th International Conference on Human-Computer Interaction*. EWHCI '94. London: Springer, pp. 35–44. ISBN: 3-540-58648-2. (Visited on 07/04/2012) (cit. on p. 36).

Renwick, John T. and Paul E. Babson (1985). "Vibration Analysis — A Proven Technique as a Predictive Maintenance Tool." In: *IEEE Transactions on Industry Applications* IA-21.2, pp. 324–332. ISSN: 0093-9994. DOI: 10.1109/TIA.1985.349652 (cit. on p. 34).

Rinderle, Stefanie et al. (2006). "Businesss Process Visualization - Use Cases, Challenges, Solutions." In: pp. 204–211. URL: http://dbis.eprints.uni-ulm.de/112/ (visited on 02/14/2012) (cit. on p. 28).

Sanderson, Penelope et al. (2004). "Visual and auditory attention in patient monitoring: a formative analysis." en. In: *Cognition, Technology & Work* 6.3, pp. 172–185. ISSN: 1435-5558, 1435-5566. DOI: 10.1007/s10111-004-0159-x. URL: http://link.springer.com/article/10.1007/s10111-004-0159-x (visited on 06/01/2015) (cit. on pp. 37, 38).

SAP SE (2015). *Monitoring PP/DS Alerts - SCM Basis - SAP Library*. URL: https://help.sap.com/saphelp_scm50/helpdata/en/a5/cf893744ce9c71e10000009b38f842/content.htm (visited on 08/18/2015) (cit. on p. 18).

Sauer, Olaf (2004). "MODERN PRODUCTION MONITORING IN AUTOMOTIVE PLANTS." In: *World Automotive Congress. Book of abstracts : 23-27 May 2004, Barcelona*. Barcelona (cit. on p. 17).

Sayal, Mehmet et al. (2002). "Business Process Cockpit." In: *Proceedings of the 28th International Conference on Very Large Data Bases*. VLDB '02. Hong Kong, China: VLDB Endowment, pp. 880–883. URL: http://dl.acm.org/citation.cfm?id=1287369.1287445 (visited on 11/04/2016) (cit. on p. 28).

Schmandt, C. and G. Vallejo (2003). ""Listenin" to domestic enviroments from remote locations." In: *Proceedings of the 9th International Conference on Auditory Display (ICAD2003)*. Ed. by E. Brazil and B. Shinn-Cunningham. Boston, pp. 220–223 (cit. on p. 43).

Schönhage, Bastiaan, Alex van Ballegooij, and Anton Elliëns (2000). "3D gadgets for business process visualization - a case study." In: *Proc. of the 5th Symposium*

on Virtual Reality Modeling Language (Web3D-VRML). ACM, pp. 131–138 (cit. on p. 28).
Seagull, F. Jacob, Christopher D. Wickens, and Robert G. Loeb (2001). "When is Less More? Attention and Workload in Auditory, Visual, and Redundant Patient-Monitoring Conditions." en. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 45.18, pp. 1395–1399. ISSN: 1071-1813, DOI: 10.1177/154193120104501817. (Visited on 07/19/2012) (cit. on p. 36).
Shneiderman, Ben (1996). "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In: VL '96. IEEE Computer Society, pp. 336–. ISBN: 0-8186-7508-X (cit. on pp. 57, 127).
Siemens AG (2007). *Siemens SIMATIC PCS 7*. Tech. rep. URL: http://w3.siemens.com/mcms/process-control-systems/SiteCollectionDocuments/efiles/pcs7/pdf/76/kg_stpcs7_v70_en.pdf (cit. on p. 18).
Song, Minseok and Wil MP van der Aalst (2007). "Supporting process mining by showing events at a glance." In: *Proceedings of the 17th Annual Workshop on Information Technologies and Systems (WITS)*, pp. 139–145 (cit. on p. 116).
Sonnenwald, D.H. et al. (1990). "InfoSound: an audio aid to program comprehension." In: *Proceedings of the Twenty-Third Annual Hawaii International Conference on System Sciences*. Vol. ii, 541–546 vol.2. DOI: 10.1109/HICSS.1990.205229 (cit. on p. 44).
Speeth, S. (1961). "Seismometer sounds." In: *The Journal of the Acoustical Society of America* 33, pp. 909–916 (cit. on p. 30).
Spehr, Georg (2008). *Funktionale Klänge: Mehr als ein Ping*. transcript Verlag (cit. on p. 5).
Suntinger, Martin et al. (2008). "The Event Tunnel: Interactive Visualization of Complex Event Streams for Business Process Pattern Analysis." In: *Proc. of the IEEE Pacific Visualisation Symposium*. IEEE Computer Society, pp. 111–118 (cit. on p. 28).
Supper, Alexandra and Karin Bijsterveld (2015). "Sounds Convincing: Modes of Listening and Sonic Skills in Knowledge Making." In: *Interdisciplinary Science Reviews* 40.2, pp. 124–144. ISSN: 0308-0188. DOI: 10.1179/0308018815Z.000000000109. URL: http://dx.doi.org/10.1179/0308018815Z.000000000109 (visited on 01/11/2017) (cit. on p. 5).
Tran, Quan and Elizabeth Mynatt (2000). "Music Monitor: Ambient Musical Data for the Home." In: *Extended Proceedings of HOIT: Home Oriented Informatics and Telematics 2000*, pp. 85–92 (cit. on pp. 36, 43).
Vaishnavi, Vijay K. and William Kuechler (2015). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology, 2nd Edition*. en. Google-Books-ID: OOE_CQAAQBAJ. CRC Press. ISBN: 978-1-4987-1526-3 (cit. on pp. 7, 8).
Varner, Philip E. and John C. Knight (2002). "Security Monitoring, Visualization, and System Survivability." In: Vancouver. URL: http://www.cs.virginia.edu/~jck/publications/security.visualization.isw-2001.pdf (cit. on pp. 40, 41).

Varner, Philip E. and John C. Knight (2004). *Monitoring and Visualization of Emergent Behavior in Large Scale Intrusion Tolerant Distributed Systems*. Tech. rep. Pennsylvania State University. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.403&rep=rep1&type=pdf (cit. on pp. 40, 41).

Vickers, P and J. Alty (1996). "CAITLIN: A musical problem auralisation tool to assist novice programmers with debugging." In: *Proceedings of the International Conference on Auditory Display ICAD'96*. Palo Alto (cit. on pp. 44, 45).

Vickers, Paul (2011). "Sonification for Process Monitoring." In: *The Sonification Handbook*. Ed. by Thomas Hermann, Andy Hunt, and John G. Neuhoff. Logos Berlin, pp. 455–492. ISBN: 3-8325-2819-9 (cit. on pp. 6, 34–36, 43).

Vickers, Paul and James L. Alty (1998). "Towards some organising principles for musical program auralisations." In: *Proceedings of the 1998 international conference on Auditory Display*. ICAD'98. Swinton, UK, UK: British Computer Society, pp. 29–29. URL: http://dl.acm.org/citation.cfm?id=2227605.2227634 (visited on 02/22/2013) (cit. on pp. 44, 64).

Vickers, Paul and James L Alty (2002). "Musical program auralisation: a structured approach to motif design." In: *Interacting with Computers* 14.5, pp. 457–485. ISSN: 0953-5438. DOI: 10.1016/S0953-5438(02)00004-8. (Visited on 08/28/2012) (cit. on pp. 44, 45).

Vickers, Paul and James L. Alty (2005). "Musical program auralization: Empirical studies." In: *ACM Trans. Appl. Percept.* 2.4, pp. 477–489. ISSN: 1544-3558. DOI: 10.1145/1101530.1101546. (Visited on 08/28/2012) (cit. on pp. 36, 45).

Vickers, Paul, Chris Laing, and Tom Fairfax (2014). "Sonification of a Network's Self-Organized Criticality." In: arXiv: 1407.4705. URL: http://arxiv.org/abs/1407.4705 (visited on 03/18/2015) (cit. on pp. 39, 40).

Walker, Bruce N. and Michael A. Nees (2011). "Theory of Sonification." In: *The Sonification Handbook*. Ed. by Thomas Hermann, Andy Hunt, and John G. Neuhoff. Logos Berlin, pp. 455–492. ISBN: 3-8325-2819-9 (cit. on pp. 32, 35, 63).

Watson, M., P. M. Sanderson, et al. (2003). "Operating Theatre Patient Monitoring: The Effects of Self Paced Distracter Tasks and Experimental Control on Sonification Evaluations." In: Ergonomics Society of Australia, pp. 128–137. URL: http://espace.library.uq.edu.au/view/UQ:99479 (visited on 07/19/2012) (cit. on pp. 37, 38).

Watson, Marcus and Penelope Sanderson (2004). "Sonification Supports Eyes-Free Respiratory Monitoring and Task Time-Sharing." en. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46.3, pp. 497–517. ISSN: 0018-7208, 1547-8181. DOI: 10.1518/hfes.46.3.497.50401. URL: http://hfs.sagepub.com/content/46/3/497 (visited on 06/01/2015) (cit. on p. 38).

Weske, M. (2007). *Business Process Management: Concepts, Languages, Architectures*. Springer (cit. on pp. 1, 2).

Weske, Mathias, Wil M. P. van der Aalst, and H. M. W. (Eric) Verbeek (2004). "Advances in business process management." In: *Data Knowl. Eng.* 50.1, pp. 1–8. URL: http://dblp.uni-trier.de/rec/bibtex/journals/dke/WeskeAV04 (visited on 07/25/2012) (cit. on pp. 2, 3).

Wieringa, R. J. (2014). *Design science methodology for information systems and software engineering*. Springer Verlag. URL: http://eprints.eemcs.utwente.nl/25449/ (visited on 03/22/2017) (cit. on p. 7).

Worrall, David (2015). "Realtime Sonification and Visualization of Network Metadata (The NetSon Project)." In: *International Conference on Auditory Display (ICAD) 2015*. Proceedings of the 21st International Conference on Auditory Display. Graz, Austria, pp. 337–339. ISBN: 978-3-902949-01-1. URL: http://iem.kug.ac.at/fileadmin/media/institut-17/proceedings150707_3.pdf (cit. on p. 42).

Yan, Zhiqiang and Paul Grefen (2011). "A Framework for Business Process Model Repositories." In: *Business Process Management Workshops: BPM 2010 International Workshops and Education Track, Hoboken, NJ, USA, September 13-15, 2010, Revised Selected Papers*. Ed. by Michael zur Muehlen and Jianwen Su. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 559–570. ISBN: 978-3-642-20511-8. DOI: 10.1007/978-3-642-20511-8_51. URL: http://dx.doi.org/10.1007/978-3-642-20511-8_51 (cit. on p. 16).

Appendix A.

Questionnaire

VP-ID: __

Condition sequence: ____________________

Personal data

Study course, respectively occupation _______________________________

Age: _______ Years

Gender: ☐ male ☐ female ☐ do not want to specify

First language ☐ German ☐ other ___________________

Id	Question	Does not apply 1	Rather not applies 2	Rather applies 3	Fully applies 4	No comment
A1	Do you have hearing impairments?					
A2	Do you have visual impairments?					
A3	Are you experienced with the development or application of process simulation?					
A4	Are you able to play an instrument?					
A5	Are you experienced with sound production, audio editing or – programming?					
A6	Are you experienced with the application or development of sonifications?					

Visualization condition
VP-ID: ___

Please indicate on a scale of 0 (does not apply at all) to 10 (does fully apply) to what extent you agree to the following statements.

7. I have understood the principles of the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

8. I have understood the interaction possibilities (supply, empty, maintain).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

9. I have felt overwhelmed by the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

10. I have felt overwhelmed by the arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

11. It was difficult to switch my attention between the arithmetic problems and the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

12. I was in control of the process simulation at all times.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

13. I have reacted in time when complications during simulation (e.g. an input buffer running empty) occurred.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

14. I have intervened in the process simulation more often than necessary.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

15. I have produced a large quantity of units.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

16. I have solved a large amount of arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

17. I have improved my performance at process simulation over time.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

18. I have improved my performance at arithmetic problems over time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

19. I have been informed about potential problems during process simulation in time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

20. The elements of the process simulation were arranged clearly and understandable.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

21. I have understood the foundations of the process visualization (e.g. the coloring of the filling levels).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

22. The previous experiment condition was exhausting.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

23. Over time, my performance at the process simulation has decreased (e.g. due to difficulties in concentrating).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

24. Over time, my performance at the arithmetic problems has decreased (e.g. due to difficulties in concentrating).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

Do you have comments, notes and suggestions concerning the previous experiment condition, and the presented mode of process monitoring?

____
____
____
____
____
____

SOTA condition
VP-ID: ___________

Please indicate on a scale of 0 (does not apply at all) to 10 (does fully apply) to what extent you agree to the following statements.

25. I have felt overwhelmed by the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

26. I have felt overwhelmed by the arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

27. It was difficult to switch my attention between the arithmetic problems and the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

28. I was in control of the process simulation at all times.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

29. I have reacted in time when complications during simulation (e.g. an input buffer running empty) occurred.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

30. I have intervened in the process simulation more often than necessary.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

31. I have produced a large quantity of units.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

32. I have solved a large amount of arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

33. I have improved my performance at process simulation over time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

34. I have improved my performance at arithmetic problems over time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

35. I have been informed about potential problems during process simulation in time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

36. The previous experiment condition was exhausting.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

37. Over time, my performance at the process simulation has decreased (e.g. due to difficulties in concentrating).

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

38. Over time, my performance at the arithmetic problems has decreased (e.g. due to difficulties in concentrating).

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

39. A longer familiarization period with the sounds would have been necessary.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

40. The sounds have helped me to perceive already occurred problems.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

41. The sounds were intrusive.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

42. My reaction times have been improved by the sounds.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

43. I have based my decisions during process simulation mainly on the sounds.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

44. I would be able to perform the mode of process monitoring presented in the previous part of the experiment for a longer period of time (without parallel arithmetic problems),

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

45. The volume of the sounds has been too low.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

46. I have based my decisions during process simulation mainly on the graphical representation.

[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

Do you have comments, notes and suggestions concerning the previous experiment condition, and the presented mode of process monitoring?

Soundscape condition
VP-ID: ___________

Please indicate on a scale of 0 (does not apply at all) to 10 (does fully apply) to what extent you agree to the following statements.

47. I have felt overwhelmed by the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

48. I have felt overwhelmed by the arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

49. It was difficult to switch my attention between the arithmetic problems and the process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

50. I was in control of the process simulation at all times.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

51. I have reacted in time when complications during simulation (e.g. an input buffer running empty) occurred.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

52. I have intervened in the process simulation more often than necessary.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

53. I have produced a large quantity of units.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

54. I have solved a large amount of arithmetic problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

55. I have improved my performance at process simulation over time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

56. I have improved my performance at arithmetic problems over time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

57. I have been informed about potential problems during process simulation in time.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

58. The previous experiment condition was exhausting.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

59. Over time, my performance at the process simulation has decreased (e.g. due to difficulties in concentrating).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

60. Over time, my performance at the arithmetic problems has decreased (e.g. due to difficulties in concentrating).
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

61. A longer familiarization period with the sounds would have been necessary.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

62. The sounds have helped me to perceive already occurred problems.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

63. The sounds have helped me to perceive potential disruptions before they have occurred.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

64. The sounds were intrusive.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

65. My reaction times have been improved by the sounds.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

66. I have based my decisions during process simulation mainly on the sounds.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

67. I would be able to perform the mode of process monitoring presented in the previous part of the experiment for a longer period of time (without parallel arithmetic problems),
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

68. The volume of the sounds has been too low.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

69. I have based my decisions during process simulation mainly on the graphical representation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

70. The previously presented sounds are informative.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

71. ...helpful for incidental process monitoring.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

72. ...intrusive.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

73. ...pleasing.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

74. ...understandable.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

75. ...euphonious.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

76. … irritating.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

77. I find the interplay of the individual sounds coherent.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

78. … logical and intuitive.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

79. I had to concentrate very hard to perceive and understand the sounds and their meaning.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

80. Over time, I had to concentrate less hard to perceive and understand the sounds and their meaning.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

81. I was able to distinguish the individual sounds.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

82. I was able to assign the sounds to the respective machines.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

83. The acoustic alerts of machines that have broken down were easy to understand.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

84. … helpful.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

85. The mapping of potential standstills onto volume was easy to understand.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

86. … helpful.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

87. The mapping of emptying input buffers onto the distortion of the sound was easy to understand.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

88. … helpful.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

89. The acoustic representation of filling output buffers trough an increase in pitch was easy to understand.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

90. … helpful.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

Do you have comments, notes and suggestions concerning the previous experiment condition, and the presented mode of process monitoring?

____
____
____

Concluding questions
VP-ID: ________

Please indicate on a scale of 0 (does not apply at all) and 10 (does fully apply) to what extent you agree to the following statements.

91. How helpful have the different presented modes of process monitoring been?

91a. Purely visual monitoring.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

91b. Additional auditory display of alerts.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

91c. Extended, continuous auditory representation during process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

92. How exhausting have the different presented modes of process monitoring been?

92a. Purely visual monitoring.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

92b. Additional auditory display of alerts.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

92c. Extended, continuous auditory representation during process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

93. How pleasing have the different presented modes of process monitoring been?

93a. Purely visual monitoring.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

93b. Additional auditory display of alerts.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

93c. Extended, continuous auditory representation during process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

94. How distracting have the different presented modes of process monitoring been?

94a. Purely visual monitoring.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

94b. Additional auditory display of alerts.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

94c. Extended, continuous auditory representation during process simulation.
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

> If the first condition included sound:

95. Did you miss the sounds later during the purely visual part of the experiment?
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

96. Did you enjoy the silence later during the purely visual part of the experiment?
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

> If the first condition did not include sound:

97. Did the sound during later parts of the experiment help you to get a better “feeling” of the occurrences during process simulation?
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

98. Do you in general agree, that sound can be helpful for process monitoring (independent of the application areas presented in this experiment)?
[0 ☐] [1 ☐] [2 ☐] [3 ☐] [4 ☐] [5 ☐] [6 ☐] [7 ☐] [8 ☐] [9 ☐] [10 ☐]

Do you have comments, notes and suggestions concerning the previous experiment condition, and the presented mode of process monitoring?

____

Appendix B.

Questionnaire Results Part 1

Questionnaire results concerning individual conditions

All data have been computed in python (NumPy/SciPy) and have been exported to xls for the sake of presentation.

Item	Mean	Std	Med	IQR
Age			26.00	3.75
A1 Hearing impairments			1.00	0
A2 Visual impairments			1.00	0.75
A3 Experiences process simulation			1.00	1
A4 Musical instrument			2.00	2
A5 Experiences Sound Creation			2.00	1.75
A6 Experiences Sonification			2.00	2
A7 understood foundations of simulation			10.00	1
A8 understood interaction possibilities			10.00	0.75
A20 simulation layout clear			8.00	2
A21 visulization principles understood			10.00	1
A63 Sounds helped to perceive potential problems in advance			8.00	0.75
A70 sounds are informative			6.50	3
A71 ...helpful for incidental process monitoring			7.50	1.75
A72 ...obtrusive			6.00	2.75
A73 ...pleasing			3.00	3.5
A74 .. comprehensible			7.00	3
A75 ... euphonious			4.00	2.75
A76 ... irritating			5.00	4.75
A77 I found the interplay of the individual sounds coherent			6.00	5
A78 ... logical and intuitive			6.50	4
A79 I had to concentrate hard in order to perceive and understand the sounds and their meaning			5.50	5
A80 Gradually, I had to concentrate less in order to perceive and understand the sounds and their meaning			7.00	2
A81 I was able to distingusih the different sounds			8.00	2.5
A82 I was able to assign the respective machines to the sounds			7.50	1
A83 The acoustic error notifications at standstill of machines were easy to understand			8.50	2
A84 ... helpful			8.00	3.5
A85 The mapping of potential standstills on the volume was easy to understand			8.50	1.75
A86 ... helpful			9.00	2
A87 The mapping of emptying input buffers on a distortion of the sound was easy to understand			6.00	5.75
A88 ... helpful			6.50	4.5
A89 The acoustic display of filling output buffers on a rising pitch was easy to understand			7.00	6.5
A90 ... helpful			8.00	5
A95 did you later, during the purely visual part of the experiment, miss the sounds?			5.00	0
A96 did you later, during the purely visual part of the experiment, enjoy the silence?			6.00	0
A97 Did the sounds help you in later parts of the experiment, to obtain a better "feeling" for the occurences in the process simulation?			6.00	0
A98 Do you in general think, that sounds can be helpful for process monitoring (independent of the areas of usage presented in this experiment)?			9.00	2
Positive towards SON	5.84	1.59		
Negative towards SON	4.79	1.79		
SON helpful	6.71	1.85		
SON unpleasant	5.26	1.75		
Understanding SON mappings	6.61	1.78		
Acoustic Experiences (1-4)	2.00	0.86		
Selfassesment	6.14	1.50		
SON sound design	4.32	2.23		
SON informative	6.73	1.86		

Appendix C.

Questionnaire Results Part 2

Questionnaire results concerning comparisons between conditions

All data have been computed in python (NumPy/SciPy) and have been exported to xls for the sake of presentation.

Item	VIS	IQR	SOTA	IQR	SON	IQR	p*	p* Vis/Sota	p* vis/son	p* sota/son	Vis/Sota #	vis/son #	sota/son #
A9/A25/A47 overwhelmed by the process simulation	1.00	2.50	1.00	1.00	2.00	3.00	0.808						
A10/26/48 overwhelmed by main task	1.00	1.00	1.00	1.75	1.50	1.75	0.872						
11(VIS)/27(SOTA)/49(SON): I found it difficult to switch my attention between the main-task and the process simulation	4.00	3.75	3.00	2.75	3.50	3.00	0.180						
12(VIS)/28(SOTA)/50(SON): I had the process simulation under control at all times	5.50	5.00	4.00	4.75	8.00	5.50	0.138						
A13/29/51 Reacted in time to complications	7.00	3.75	6.00	2.00	8.50	3.50	0.106						
A14/30/52 intervened more often than necessary	3.50	5.00	4.00	3.50	4.00	4.00	0.893						
17(VIS)/33(SOTA)/55(SON): I have improved my performance in the process simulation over time	5.00	2.00	5.00	2.00	7.00	3.50	0.009	0.633	0.057	0.064	0.633	0.048	0.048
18(VIS)/34(SOTA)/56(SON): I have improved my performance in the arithmetic problems over time.	5.00	1.00	4.00	2.00	5.50	4.00	0.078	0.040	0.814	0.053	0.080	0.407	0.040
19(VIS)/35(SOTA)/57(SON): I have been informed in time about potential problems during process simulation	2.50	4.75	3.00	4.50	9.00	4.50	0.001	0.482	0.001	0.004	0.241	0.002	0.003
22(VIS)/36(SOTA)/58(SON): The previous part of the experiment was exhausting	5.00	4.00	4.00	4.00	6.00	2.00	0.247						
39(SOTA)/61(SON): A longer period of familiarization with the sounds would have been necessary			0.50	2.75	4.00	5.75	0.013						
40(SOTA)/62(SON): The sounds helped me in perceiving already occurred problems			8.50	2.00	8.50	1.00	0.928						
41(SOTA)/64(SON): The sounds were distracting/disturbing/annoying			3.50	6.50	6.00	4.00	0.491						
42(SOTA)/65(SON): My reaction time has been improved through the sounds			7.50	2.75	8.00	1.75	0.972						
43(SOTA)/66(SON): For process simulation, I have mainly oriented myself at the sounds			3.50	4.00	8.50	3.00	0.009						
44(SOTA)/67(SON): could carry out process monitoring for longer period of time			7.00	2.75	5.50	4.75	0.196						
91a(VIS)91b(SOTA)/91c(SON): How helpful were the different modes of process monitoring	6.00	3.50	7.50	1.00	8.50	1.00	0.002	0.036	0.012	0.390	0.027	0.017	0.195
92a(VIS)92b(SOTA)/92c(SON): How exhausting were the different modes of process monitoring	5.00	4.50	5.00	4.75	5.50	3.75	0.869						
93a(VIS)93b(SOTA)/93c(SON): How pleasing (pleasant?) were the different modes of process monitoring	5.50	3.50	5.00	3.75	5.00	3.75	0.637						
94a(Vis)/b(Sota)/c(Son): How distracting were the different modes of process monitoring?	4.50	6.50	5.50	5.00	5.00	4.50	0.597						

* Differences between groups, Friedman test for comparisons between three groups, Wilcoxon test for comparisons between two groups

p-values corrected for multiple comparisons with Benjamini-Hochberg correction, divided by two if directed hypothesis was made beforehand

Appendix D.

Participant Summary

# vpnr	mode	Sequence	no. calculations	mean dev from solution in %	mainTaskScore	no. clicks	mean AO	mean buffer at supply	at empty	at maintain c	at maintain e
3001	vis	0	136	0.8767	0.1948	33	0.5271	0.4708	0.4679	59.3208	43.2948
3001	sota	1	130	1.4180	-0.7453	33	0.5686	0.3263	0.4250	65.5653	33.8026
3001	son	2	120	0.9653	-0.4431	25	0.7814	0.1959	0.6400	31.7551	23.5493
3002	sota	0	74	0.1803	-0.8426	39	0.5157	0.2350	0.7500	28.7380	66.3746
3002	vis	1	127	0.4539	0.4879	22	0.8171	0.2673	0.7531	11.3276	15.5206
3002	son	2	134	0.9426	0.0395	21	0.8124	0.0740	0.8563	11.4485	12.4298
3003	son	0	150	0.2886	1.4572	25	0.6787	0.3023	0.5646	10.0802	32.7094
3003	sota	1	180	0.0935	2.6932	22	0.6089	0.3264	0.6425	22.4077	7.7075
3003	vis	2	187	0.3762	2.5288	21	0.7170	0.2718	0.7175	1.8768	6.8816
3004	vis	0	136	0.0528	1.3314	24	0.7211	0.2300	0.7000	8.8252	11.5741
3004	son	1	133	0.9698	-0.0303	22	0.8594	0.0795	0.7150	8.3502	14.0351
3004	sota	2	124	1.5186	-1.0776	21	0.6251	0.0685	0.7438	13.5658	23.8706
3005	sota	0	128	1.2736	-0.6106	22	0.8339	0.1735	0.6675	15.6717	16.7237
3005	son	1	157	0.7394	1.0609	26	0.7411	0.1465	0.7625	15.6219	52.9710
3005	vis	2	155	2.0495	-0.8108	22	0.7350	0.0705	0.6750	13.2735	16.3095
3006	son	0	126	0.0256	1.0466	23	0.7175	0.1186	0.8094	1.9027	33.3976
3006	vis	1	153	0.8786	0.7400	21	0.6907	0.1165	0.6750	0.8657	10.9462
3006	sota	2	131	0.3897	0.7054	22	0.6412	0.1250	0.6175	7.2772	24.1074
3007	vis	0	79	0.0000	-0.4327	43	0.6340	0.1475	0.6525	10.9354	6.5514
3007	sota	1	102	0.0801	0.1980	23	0.7942	0.2091	0.6475	15.0027	15.3691
3007	son	2	105	0.2761	0.0243	24	0.8329	0.1677	0.7025	14.6522	23.9598
3008	sota	0	160	0.3488	1.6964	22	0.5107	0.2964	0.6475	9.5507	-1.2809
3008	vis	1	169	0.1862	2.2108	23	0.7143	0.2618	0.6125	11.7014	5.9366
3008	son	2	172	0.1957	2.2944	22	0.8104	0.1490	0.6700	33.5223	22.1243
3009	son	0	88	1.1364	-1.7104	23	0.7371	0.3495	0.6825	6.2913	21.6806
3009	sota	1	91	0.0000	-0.0460	28	0.5877	0.4415	0.4094	8.3066	24.6564
3009	vis	2	103	2.2739	-2.7962	29	0.5002	0.5400	0.3389	10.9070	12.4721
3010	vis	0	134	1.0756	-0.1441	19	0.4906	0.0272	0.7688	4.1496	21.7883
3010	son	1	138	1.2353	-0.2355	24	0.7614	0.1005	0.6675	39.4596	37.5199
3010	sota	2	152	1.7524	-0.4977	24	0.3385	0.0095	0.6975	47.3434	16.3049
3011	sota	0	123	0.4558	0.3565	20	0.6705	0.0945	0.8375	1.2400	16.2471
3011	son	1	119	1.0676	-0.6165	22	0.8461	0.1340	0.6050	11.6043	15.4507
3011	vis	2	131	1.3943	-0.6805	22	0.6695	0.0590	0.6975	10.1567	14.8709
3012	son	0	112	0.4306	0.0367	25	0.7927	0.1641	0.5875	10.9116	23.2983
3012	vis	1	144	0.4082	1.0988	21	0.5868	0.1275	0.6925	14.7426	11.7831
3012	sota	2	139	0.4102	0.9351	24	0.5667	0.0773	0.8844	13.1204	37.7776
3013	vis	0	178	1.3532	0.8910	37	0.4253	0.4907	0.3653	61.6671	44.5775
3013	sota	1	195	0.2315	2.9863	32	0.4458	0.2871	0.5667	55.6832	49.7013
3013	son	2	186	0.4648	2.3743	33	0.5940	0.4534	0.5750	43.7661	30.9470
3014	sota	0	111	0.1609	0.3766	24	0.5681	0.3650	0.7813	8.2950	29.2365
3014	vis	1	131	2.2496	-1.8603	28	0.6306	0.4932	0.5896	32.5599	13.5370
3014	son	2	120	1.8069	-1.6042	23	0.7404	0.3858	0.7625	14.3790	19.4509
3015	son	0	102	0.2629	-0.0542	24	0.7925	0.2527	0.6775	9.1967	24.4676
3015	sota	1	131	0.1336	1.0588	21	0.7887	0.1180	0.7688	13.8878	17.7249
3015	vis	2	104	0.3292	-0.0812	23	0.6233	0.1429	0.7844	6.6971	15.7148
3016	vis	0	126	1.0347	-0.3455	20	0.6788	0.3010	0.8031	6.7969	2.6382
3016	son	1	124	0.5463	0.2639	23	0.8017	0.2650	0.8094	14.8542	26.4264
3016	sota	2	137	0.9665	0.1031	24	0.6048	0.4931	0.7594	9.7516	12.4269
3017	sota	0	74	3.0326	-4.7775	25	0.2238	0.2050	0.4768	0.2090	15.0443
3017	son	1	68	0.9969	-2.1625	21	0.8300	0.2168	0.6719	7.8874	13.7659
3017	vis	2	56	2.4502	-4.5541	52	0.2144	0.3956	0.1227	63.4877	53.7995
3018	son	0	95	1.8428	-2.4594	30	0.5541	0.1365	0.6969	3.2125	68.1053
3018	vis	1	113	0.1974	0.3907	23	0.5540	0.0239	0.6775	12.6422	36.2145
3018	sota	2	104	0.2434	0.0372	22	0.5570	0.0340	0.7250	7.1441	12.8213

Appendix E.

Spectrograms

The following figures (Fig. E.1 - Fig. E.8) show spectrograms of the different sounds that are conveyed for each of the 6 machines, when the respective machine produces a unit. For most machines, only spectrograms of recordings during normal operation are shown. However, to demonstrate the effects of the different mappings, the spectrograms of two machines are contrasted with spectrograms during critical states. The input buffer of the "drop" machine has run low, and thus in addition to the volume of the sound being increased, noise is added to the sound. The output buffer of the "water jug" machine has run full, and thus the pitch of the respective sound is increased.

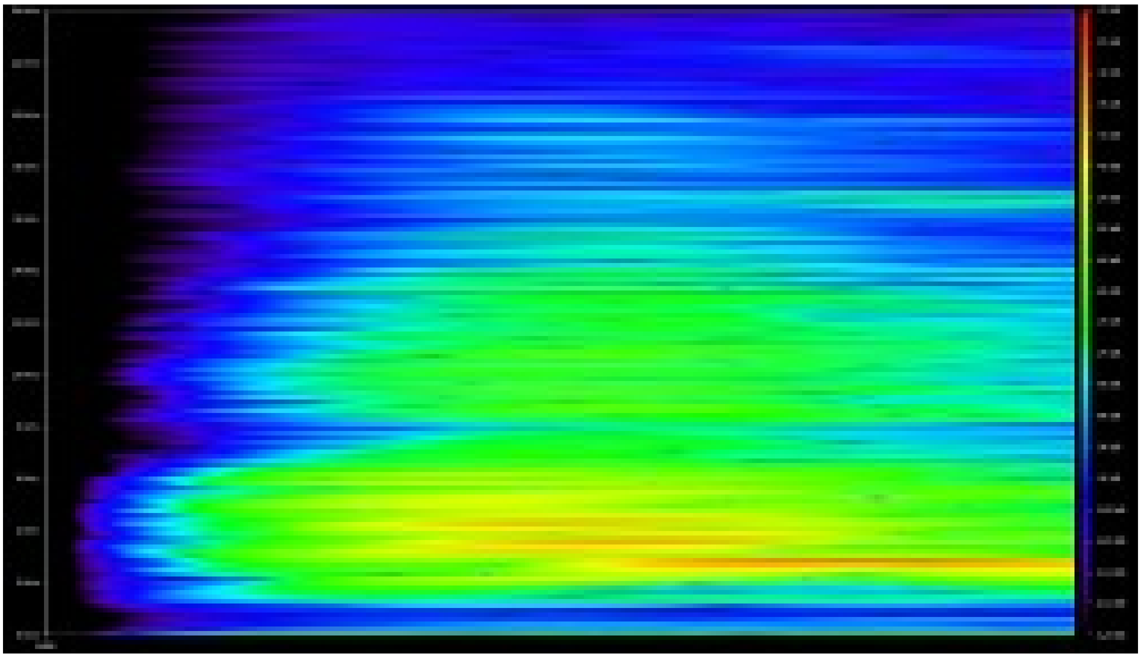


Figure E.1.: Bird machine in normal state.

Fig. E.9 shows the spectrogram of a short recording during a non-critical stage in process simulation, in which all 6 machines were active. Fig E.10 shows the spectrogram of a short recording during a stage in the process simulation in which one machine (the water drop-machine) has come to a halt due to its

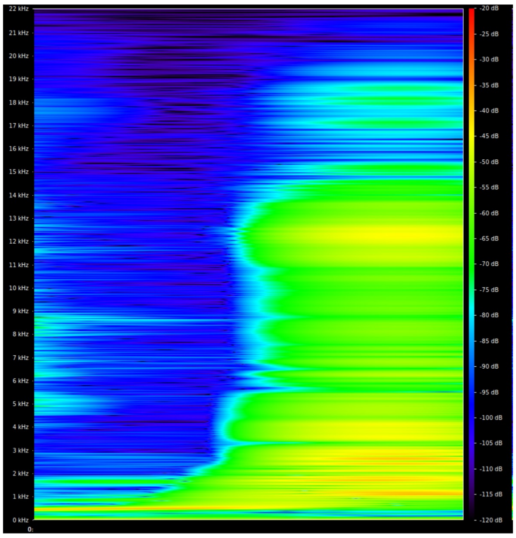


Figure E.2.: Dripping water machine in normal state.

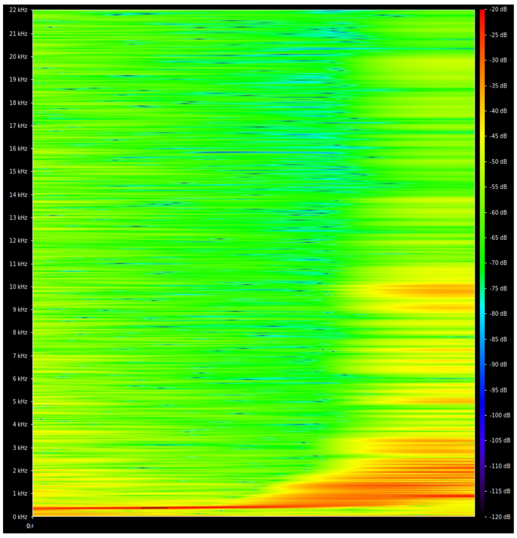


Figure E.3.: Dripping water machine in critical state (input buffer empty). The volume is increased and pink noise is added.

input buffer having run low, while the other 5 machines continue to function normally.

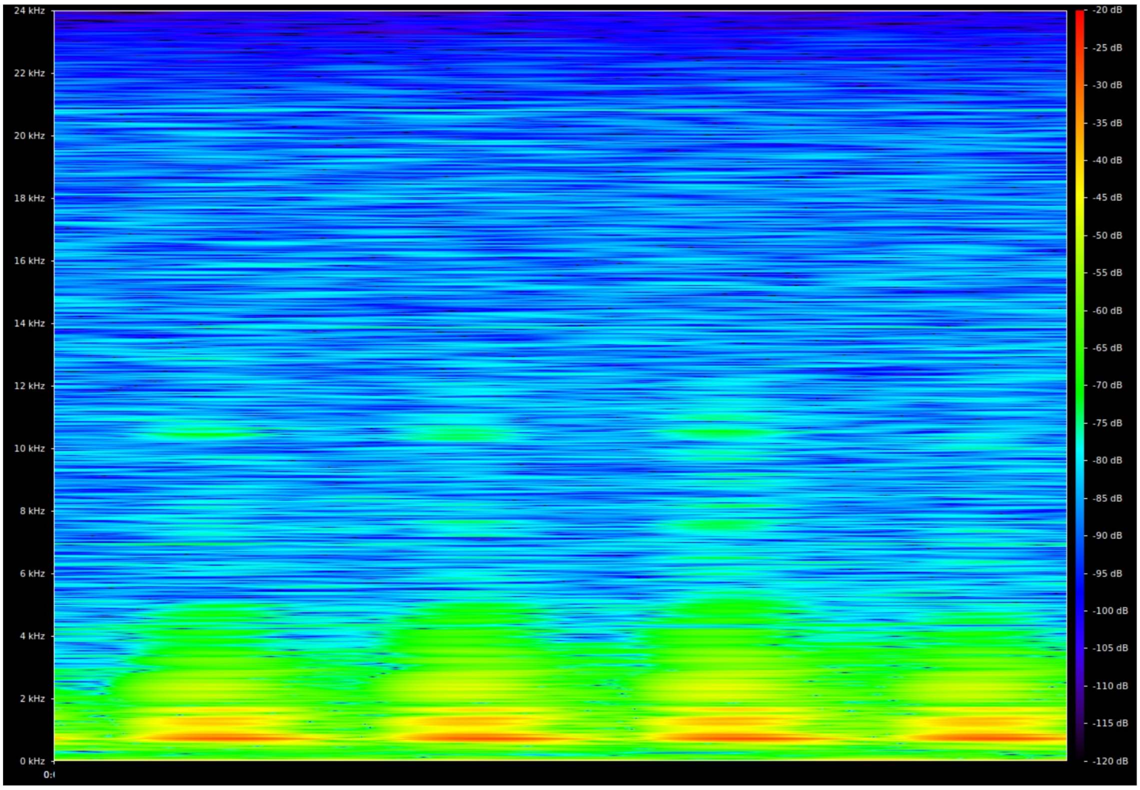


Figure E.4.: Woodpecker machine in normal state.

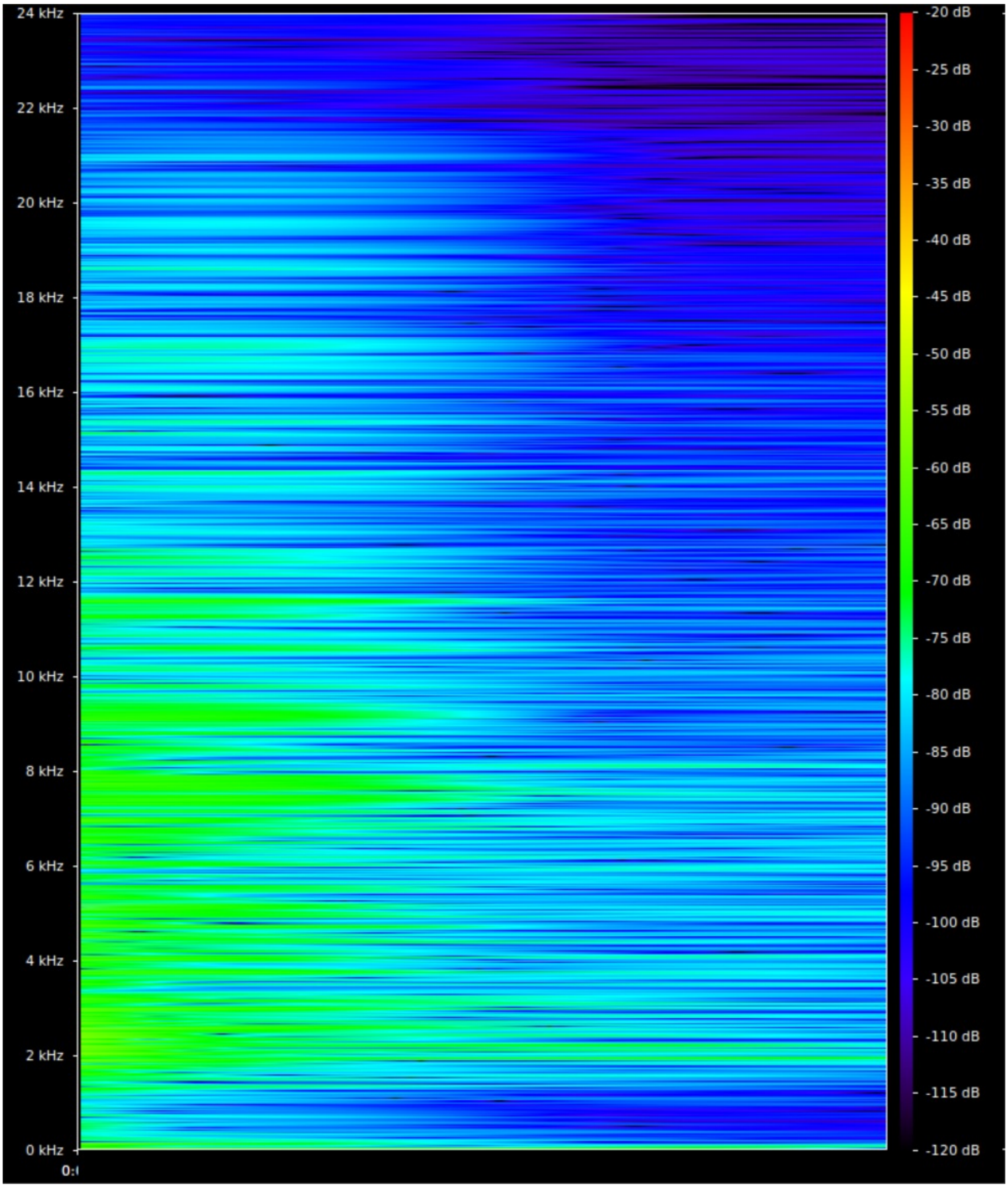


Figure E.5.: Breaking branch machine in normal state.

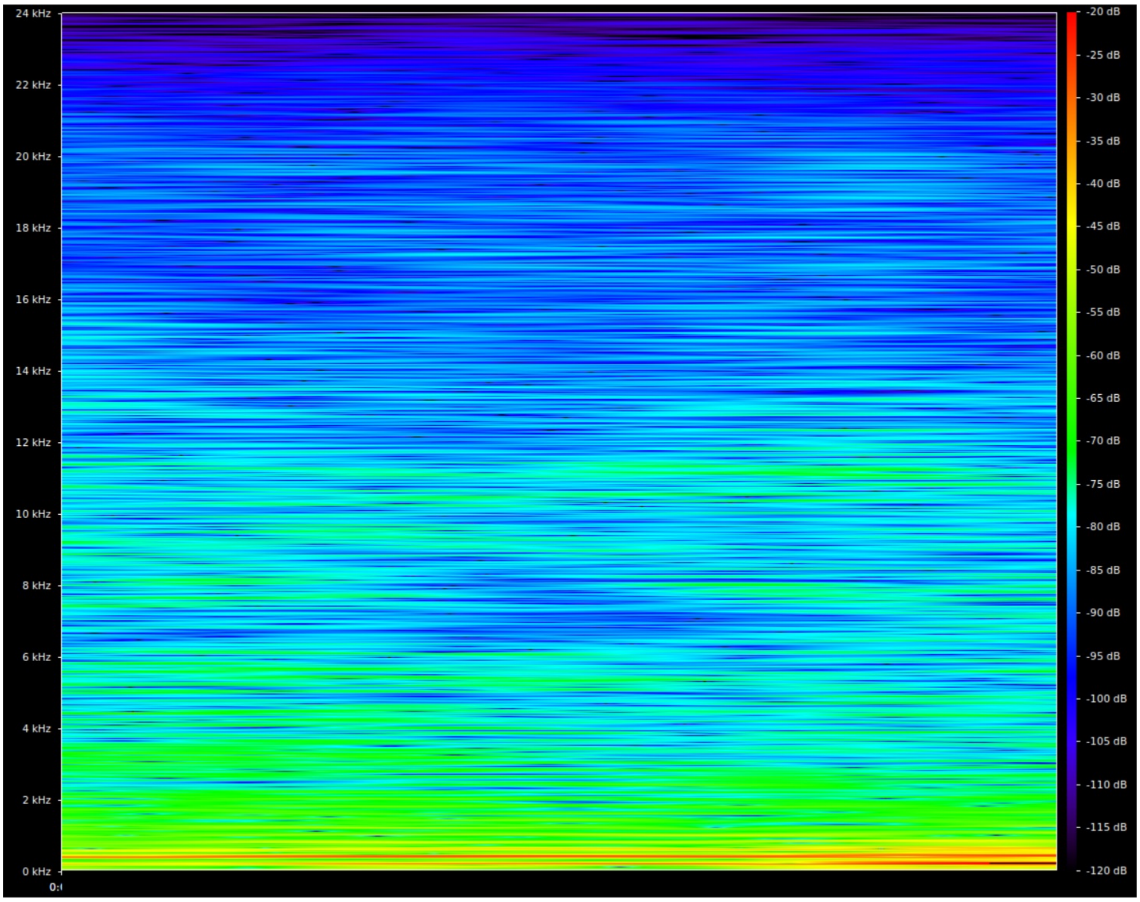


Figure E.6.: Bee machine in normal state.

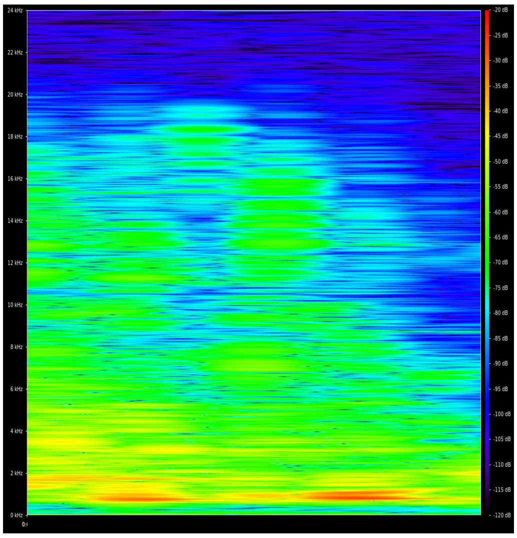


Figure E.7.: Water jug machine in normal state.

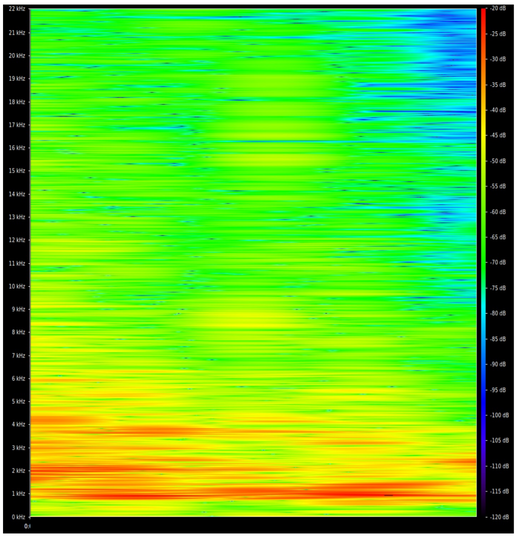


Figure E.8.: Water jug machine in critical state (output buffer full). Volume and pitch are increased.

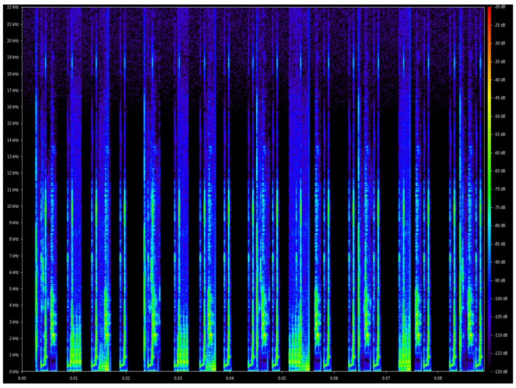


Figure E.9.: Spectrogram of short recording of C_{son} during normal operation.

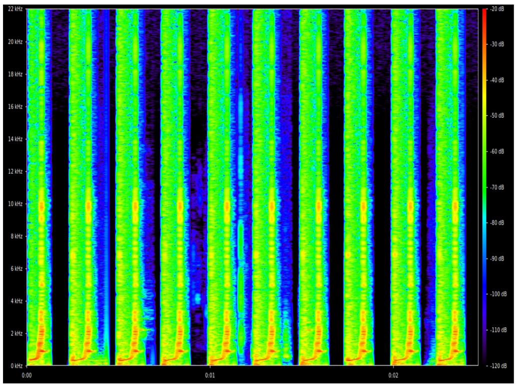


Figure E.10.: Spectrogram of short recording of C_{son} with one machine in critical state (input buffer of dripping machine empty). The sound of the respective machine is being repeated in short intervals, with increased volume and pink noise added.