Multimodal User Interfaces in IPS²

Ulrike Schmuntzsch and Matthias Rötting

Technische Universität Berlin, Department of Psychology and Ergonomics,
Chair of Human-Machine Systems, Franklinstr. 28-29,
10587 Berlin, Germany
{ulrike.schmuntzsch,matthias.roetting}@mms.tu-berlin.de

Abstract. The field of Industrial Product-Service Systems (IPS²) faces various challenges. Goal of the recently started project is the multimodal design of interaction specific warnings and instructions in IPS² for human operators with relation to their interactions. This approach should help to prevent mistakes of human operators whilst interaction and it should account for the support and optimization of work process in the heterogeneous area of IPS². Here we discuss how such a project can be realized step-by-step.

Keywords: Multimodal user interfaces, interaction specific warnings, user generated instructions, IPS².

1 Introduction

Regarding industrial production a global trend is the growing specialization and complexity within supply chains. The automation level increases over the time and the pressure to improve productivity grows in many areas. Procedures become increasingly multifaceted and require more expert knowledge. Especially the globalised field of Industrial Product-Service Systems (IPS²) is faced with these challenges. IPS² are characterized by integrated and reciprocally determined design, development, provision and utilization of products and services. This also includes the possibility to substitute partial aspects of services and products [1] during the design and provision phases of IPS².

These highly automated production systems can be seen as socio-technical systems involving both human factors and mechanical components. Human factors can be described as a set of human-specific physical, cognitive, or social skills and abilities such as attention, information processing, or dealing with stress [2]. Thus, within socio-technical systems such as IPS² the interplay between human operators as well as between technical components and the human-machine interaction is important. A variety of different requirements put high demands on human operators within their working environment. The dynamic and complex structure requires flexible operations and adaptability. Furthermore there is a need for broad understanding as well as special system knowledge. For these purposes, it is important to provide concepts that enable human operators to handle the complexity and dynamic as well as to optimize human reliability and performance. At least the developed systems should on the one hand reduce human error and adapt to it.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2011, LNCS 6777, pp. 347–356, 2011.

Most current research and development work focuses mainly on the improvement of technical components, neglecting the human factor within the socio-technical system. As a result 50-80% of the reported incidents are classified as *Cause Human* [3, 4].

For this reason, approaches are needed which allow an automatic and predictive detection of human errors and malfunctions. The major goal behind this approach is an adaptive and user-centered support of human operators. It is particularly important to pay best care and attention to error situations because they are critical for usability for two reasons. Firstly, they stand for situations where the user is in trouble or even in danger and potentially will be unable to use the system to achieve the desired goals. Secondly, they present options for supporting the user to bring the tasks to a favorable conclusion and in the end to understand the system better. That is why warnings are considered as an essential part of a general interface design strategy of guidance for users. It is important to make sure that the error messages are integrated, coordinated and consistent across one or more applications [5].

In this article an approach how to design a multimodal user interface for IPS² that outputs interaction specific warnings and user generated instructions will be presented. It begins with a brief introduction to the theoretical background of the topic. Chapter three deals with the approach for a multimodal user interface in IPS². Based on the theoretical background and the described approach the future considerations will be briefly outlined in chapter four. The article concludes with a discussion in chapter five and a short summary in chapter six.

2 Theoretical Background

Beside the special context of IPS², two aspects seem to be most important concerning the project definition – the idea of a multimodal user interface and the interaction specific warnings.

2.1 Multimodal User Interfaces

Multimodal user interfaces are characterized by two or more combined user input and system output modes which are featured in a coordinate manner. This kind of interfaces intends to identify naturally occurring human language and behavior, which include at least one recognition-based technology. In order to support intelligent adaptation to the user, task and context, information on the user's interface and the surrounding environment will be tracked and included from multiple sensors. Furthermore information can be transmitted to the user in different modes. Well-known examples are the presentation of speech or sound and their registration on the acoustical side. Regarding the visual side, pictures, text or videos are often used as system output. Likewise gestures, facial expressions and eye movements can be analyzed. Haptic feedback is often transmitted by pressure or vibrations and registered by force or manual operations [6]. Figure 1 provides a brief overview of the different input and output modes in human-machine systems.

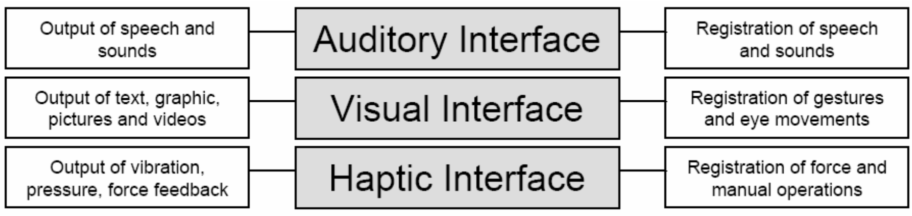

Fig. 1. The three perception modes on which interface designers mainly focus

Several advantages are offered by the approach using multimodal interfaces. For example, it allows different users to interact with technical system in a way which best suits their individual skills and abilities. Also, human factors such as the native language status, the skill or the age of the human operator can be taken into consideration. Furthermore a continuous accommodation to changing working conditions and contexts is possible. For example, if there is a high noise pollution, it is feasible to switch from the acoustic mode to the visual or the haptic one, or to receive these feedback modalities as back up. Moreover, in case of temporary disability, such as a human operator is not able to use an input mode for a period of time, it is also possible to use another one in order to perform the task. Thus, temporary disabilities must not be a factor anymore why assemblers who are in other respects healthy cannot execute their work [7].

Additionally, these interfaces fulfill users' needs to interact multimodally because they are already familiar to communicate in such a way during their natural interpersonal relationships. Regarding the error handling, various authors acknowledge that the multimodal design is superior in error avoidance and successful recovery from errors [8].

2.2 Warnings in Human-Machine Systems

Warnings are an essential part of a general interface design strategy of user guidance. Difficulties in understanding the warnings and instructions sometimes encountered when they do not correspond with the way a human operator usually perform their tasks. This indicates an obvious information gap in transitioning from error message to supporting the user in performing the corrective action. Moreover it is a sign of a gap between the system model designers have and the users´ mental model. The last one can be described as a representation of the surrounding world, the relations between its individual components as well as the intuitive perception of humans about their own activities and their consequences [9].

Overall, warnings and instructions are a sensitive issue. Since errors occur because of incorrect understanding, inadvertent slips, or a lack of knowledge, users are likely to feel uncertain or confused as well as inadequate. Furthermore, users are scared by an imperious tone, so that it makes more difficult to correct the problem. Thus, warnings have to be specific and constructive as well as formulated positively and in an appropriate physical format [10]. Moreover the dialogues should be simple and natural in order to speak the user's language. Information objects and operation must be presented in a sequence which corresponds with the way a user will accomplish the task mostly. Therefore it is necessary to discover mappings between machine information and the user's mental model of that information.

This requires carrying out a so called *Usability Engineering Process*. That involves a structured approach with empirical proven methods. In the following a common way for realization will be introduced.

2.3 Usability Engineering Lifecycle

The *Usability Engineering Lifecycle* by Deborah Mayhew consists of the following three phases:

1. Requirements Analysis
2. Design/ Testing/ Development and
3. Installation [11].

Each phase consists of a series of tasks. It begins with the requirements analysis where, for example, user profiles and task analyses have to be carried out. Thus, it is intended to describe specific characteristics of the potential users, for example the level of job experience, expected frequency use. The developed user profiles will influence the design decisions according to the user interface and also help to identify the major user categories. The tasks analyses are used to study work-flows and current user tasks in order to get an understanding of the underlying user goals. The gathered information will be used to set the usability goals and to guide the user interface design process. Additionally, the platform capabilities and constraints as well as the general design principles have to be considered within the requirement analysis. As a result usability goals and design style guides are available for the second phase. Therefore it is necessary to survey user profiles and task analyses carefully and with best attention.

Based on the established usability goals and the developed style guides, the design, testing and development of the user interface is carried out in the second phase. This is regarded as an iterative process in which the development runs from conceptual models and mockups to increasingly sophisticated prototypes. At the end, a detailed user interface will be tested. Here it is important to make sure that the established usability goals are met and that all system functionalities are addressed. Otherwise it does not make any sense to install the half-finished user interface in human operators´ work environment because it would only be counterproductive.

Phase three includes the installation and the spot survey of human operators in the daily working routine. At this point, various problems may emerge so that a re-engineering process is necessary.

Concluding, it is really important to pay best care and attention to the first phase where user profiles and task analyses are performed. The more time and effort is invested here, the more it is then saved later on. Figure 2 shows an overview of the complete *Usability Engineering Lifecycle*.

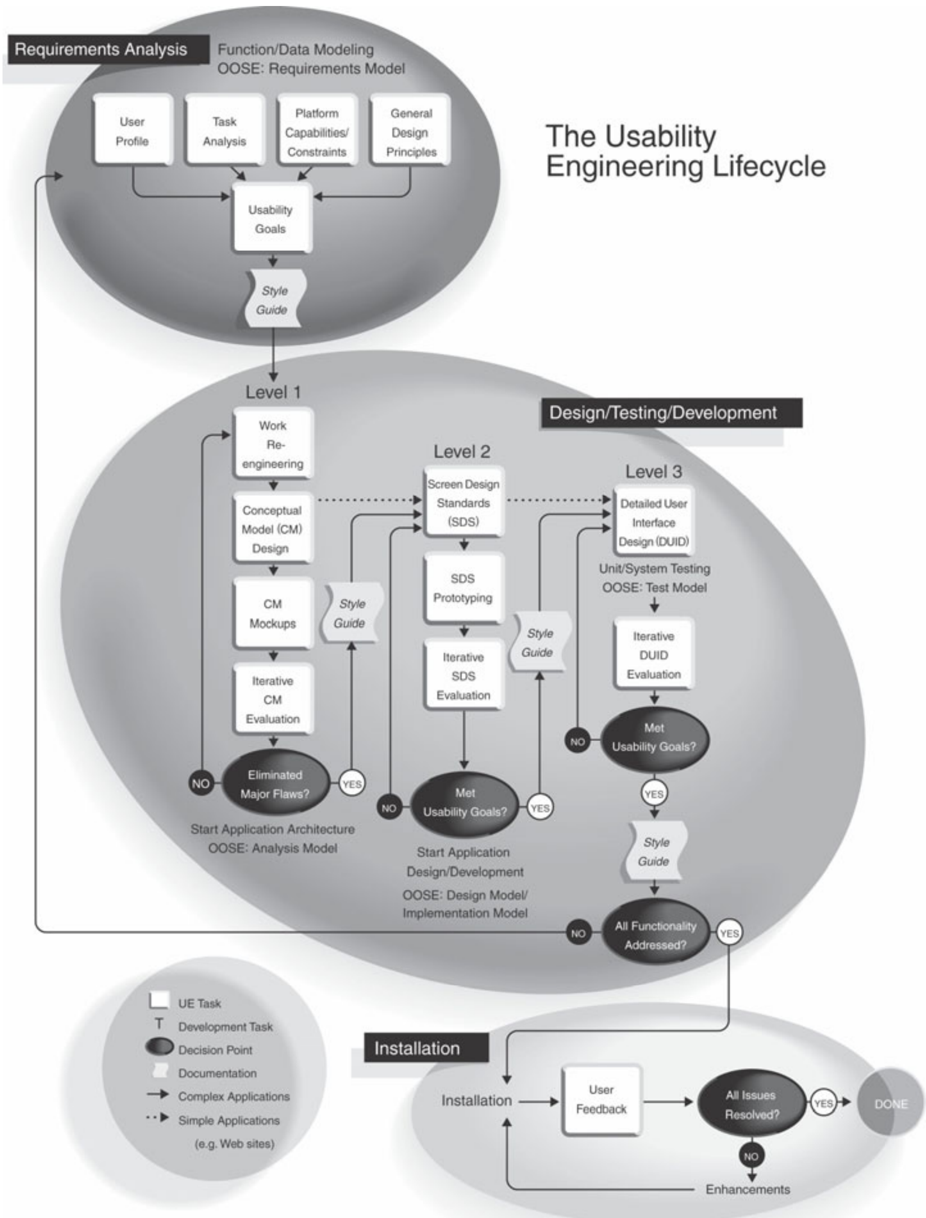

Fig. 2. The *Usability Engineering Lifecycle* by Deborah Mayhew. It consists of three phases – requirements analysis, development / design / testing and installation. Each phase consists of a series of tasks [12].

3 Approach for a Multimodal User Interface

In accordance with the heterogeneous and dynamic context of IPS² and the given particular importance of error situations, the remit of the project is to design a multimodal user interface that outputs interaction specific warnings and user generated instructions. The mission of this project is to take the human factors more into consideration within such a socio-technical system like IPS². Consequently, it is intended to make an optimal use of the human abilities and skills. Weaknesses and limitations should be compensated by the technical system.

The major goals behind this mission are to prevent mistakes by human operators whilst interacting and to support as well as to optimize work processes. Because of the heterogeneity of users and contexts both individual and contextual assistance is regarded as particularly useful. Based on the automatic detection and evaluation of human interaction, multimodal instructions and warnings will be generated and then presented contextually to the human operator.

For theses goals, we are striving to develop a concept which includes the following three aspects:

1. contextual and multimodal assistance
2. instructions
3. expert platform

First of all there is a need to provide contextual and multimodal assistance for human operators who are going to make mistakes while operating. Instructions about how to do better also have to be featured. And third an expert platform is required, e.g. in order to collect all previous successful solutions to fall back on it in case of problems. The integrated combination of these three aspects is necessary to support the human operator, so that human abilities and skills can be used in an optimal way.

The realization means in detail that a decision has to be made whether the user or the system state or both should be influenced. If the current situation is classified as less critical and solvable then it is intended to support the human operator in task performance. But if there is a safety-relevant situation, the machine state should be changed, e.g. to stop the engine and to provide information to the human operator [13]. After deciding who is getting regulated or changed, it has to be determined how complex and in which granularity the future action should be implemented. Regarding automation systems it is quite common to distinguish between different types of support functions, like to inform or to warn the user as well as to suggest one or more action alternatives or even to implement a suitable action by the machine itself [14]. Thereafter an action will be initiated which can mean a regulation on human side and/or a change of the technical system state. These operations lead to a new state of the socio-technical system which affects again both human operator and machine.

To realize this concept, detailed information about products and services as well as their reciprocal effects are necessary. Furthermore an adequate knowledge of the system state has to be provided in order to initiate specific instructions, e.g. to open / close a valve or to speed up the flow rate. Especially, information about the detected and classified deviation of human operators from the predicted action sequences will be required. For this purpose, cognitive user models are developed in order to simulate the optimum user behavior and to predict the future behavior [15]. All these information are provided by other subprojects. The interdependencies between the different subprojects and their provided information are shown in figure 3.

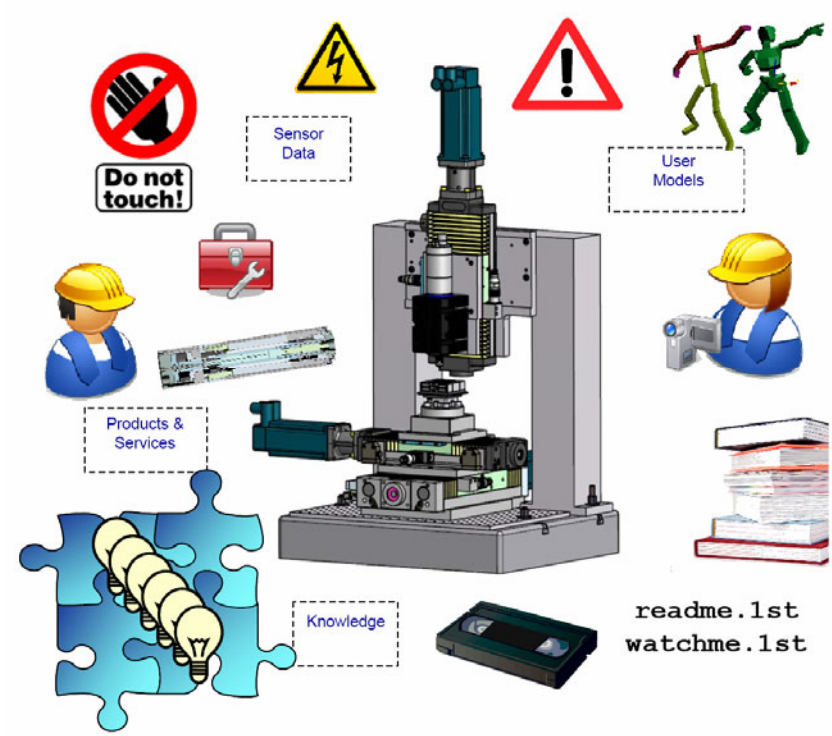

Fig. 3. A demonstrator at a micro production scenario in IPS² is shown. The aim is to develop a multimodal user interface that provides interaction specific warnings and user-generated instructions. Detailed sensor data and cognitive user models are required as well as information about products and services.

Regarding the project goals it is intended to accomplish a *Usability Engineering Process* according to the proposed concept by Deborah Mayhew.

4 Future Considerations

Considering both *Usability Engineering Lifecycle* and the described project goals, preparatory steps are taken in order to develop a system which provides contextual and multimodal assistance to human operators, who are going to make mistakes during interaction. First of all the user interface requirements have to be clearly defined, e.g. by applying ethnographical observations of potential users and their work environment. Furthermore major problems in the execution of different work steps have to be identified through task analysis. Therefore action oriented methods such as the Hierarchical Task Analysis (HTA) or cognitive task analysis like the Critical Action and Decision Evaluation Technique (CADET) can be used [16].

After getting a feeling for the daily working routine, the everyday speech and the problems human operators can face, a concept for contextual and multimodal assistance can be worked out in exemplary fault scenarios. At the beginning the developed interaction-specific warnings can be tested in paper-pencil versions or by using the card sorting method or a paired similarity rating. It is well-known from usability research experiments that the vote-winning terms enable users to get a better understanding of the system. For example, they can generalize their knowledge more often to correctly use it in new ways [17]. After making sure that user's mental model of the system matches with the ideas of the warning concept, it should be implemented and tested in scenarios with a prototype. The warnings can then be optimized and tested again.

5 Discussion

To design a multimodal user interface that outputs interaction specific warnings and user-generated instructions as well as to implement it within the heterogeneous area of IPS² it takes time. The heterogeneity of users and contexts dominates the work environment. That is why the aspect of multimodality and sensitivity for contexts as well as for individual user preferences are of such importance. But this heterogeneity and individuality is also a great challenge for developing a multimodal user interface for IPS². Therefore it is necessary to focus only on a few exemplary user contexts or fault scenarios and to work these out well.

Moreover, it has to be pointed out that the current behavior of the human operator as well as the estimation of the situation criticality and the presentation of the interaction specific warnings have to be transmitted in real-time.

Another open question will be what kind of processes and operations are supportable at all. Firstly, it is potentially feasible to present interaction specific warnings for each sequence of actions which is well-structured and previously defined. But every procedure that is executed spontaneously by the human operator will be difficult to deal with, because no cognitive user models or machine learning algorithms patterns could be previously defined. Thus, it is not possible to find a corresponding action in the database and so no appropriate warning or instruction can be displayed.

Regarding the acceptance of the proposed concept in real work environments it has be clarified how the human operator will interact with the developed system and how the digital data storage will be accepted.

Finally, it has to be questioned critically to what extent the developed technology will be used for. Major goals are to provide technical support to human operators in order to protect man and machine against risks as well as to optimize circuitous and unnecessary work processes. But there is also the risk that the continuous data storage a feeling of being monitored as well as time pressure will raise by the human operator due to. This might be a great challenge concerning to the acceptance of the technology. Another issue raised by the data storage will be the unresolved question of data security. These and other challenges have to be considered.

6 Summary

The written article describes an approach how to develop a multimodal user interface that provides interaction specific warnings and instructions in the heterogeneous field of IPS². This special industrial sector is determined by a dynamic and complex structure in which the design, provision and utilization of products and services are integrated and reciprocally.

The theoretical background of this project is the design of multimodal user interfaces as well as its current inventions. Therefore the particular conditions and challenges in the field of IPS², e.g. regarding the micro production, should be carefully analyzed and considered in the planning phase.

It seems most important that best care and attention should be paid to the implementation of a *Usability Engineering Process* in order to develop an appropriate and usable multimodal interface. Accordingly, in the near future it is intended to get an idea of the daily working routine and the special problem the human operator can be faced with. Based on these consolidated findings the project work will be put forward step-by-step.

Finally, the designed multimodal user interface leads to a more user-centered approach in IPS² which contributes to prevent mistakes of human operators and to support as well as to optimize work processes.

The TRR 29 B4 project aims to provide a method to integrate human factors into IPS² as a core context in its interaction techniques.

Acknowledgments. We thank the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) for funding this research within the Transregional Collaborative Research Project TRR 29 on Industrial Product-Service Systems – dynamic interdependencies of products and services in production area.

References

1. Meier, H., Uhlmann, E., Kortmann, D.: Hybride Leistungsbündel – Nutzenorientiertes Produktverständnis durch interferierende Sach- und Dienstleistungen. In: Wt Werkstattstechnik Online, 95, Jahrgang, pp. 528–532. Springer-VDI-Verlag, Heidelberg (2005)
2. Wickens, C.D., Gordon-Becker, S.E., Liu, Y., Lee, J.D.: Introduction to Human Factors Engineering – International Edition, 2nd edn. Prentice Hall, New Jersey (2003)
3. Major Accident Hazards Bureau (MAHB): Major Accident Reporting System (MARS), `http://mahbsrv.jrc.it/mars/Default.html` (October 01, 2007)
4. Working Group on Chemical Accidents: Preparation of a work-shop on human factors in chemical accidents and incidents. 5. In: Proceeding at the 16th Meeting of the Working Group on Chemical Accidents, Varese (2006)
5. Reason, J.: Human Error, 2nd edn. Cambridge University Press, Cambridge (2003)

6. Hedicke, V.: Multimodalität in Mensch-Maschine-Schnittstellen. In: Timpe, K.-P., Jürgensohn, T., Kolrep, H. (eds.) Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation, 5th edn., pp. 203–232. Symposion Publishing GmbH, Düsseldorf (2002)
7. Oviatt, S.: Multimodal Interfaces. In: Jacko, J., Sears, A. (eds.) Handbook of Human-Computer Interaction. Lawrence Erlbaum, New Yersey (2002)
8. Oviatt, S., Bernard, J., Levow, G.: Linguistic adaptation during error resolution with spoken and multimodal system. Language and Speech, Special Issue on Prosody and Conversation 41(3-4), 415–438 (1999)
9. Norman, D.A., Draper, S.: User centered system design: New perspectives on human-computer interaction. Lawrence Erlbaum Associates, Mahwah (1986)
10. Shneiderman, B., Plaisant, C.: Designing the User Interface – Strategies for Effective Human-Computer Interaction, 5th edn. Addison-Wesley, Boston (2010)
11. Mayhew, D.: The Usability Engineering Lifecycle – A Practitioner's Handbook for User Interface Design. Academic Press, San Francisco (1999)
12. Mayhew, D.: The Usability Engineering Lifecycle, `http://www.isrc.umbc.edu/HCIHandbook/figures/47-01.jpg` (January 31, 2011)
13. Dzaack, J., Höge, B., Rötting, M.: User-Centric and Contextual support in IPS^2. In: CIRP IPS^2 Conference, pp. 13–19 (2010)
14. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics 30, 286–297 (2000)
15. Beckmann, M., Dzaack, J.: Incorporating emotion data and cognitive models in IPS^2 (in press, 2011)
16. Embrey, D.: Task Analysis Techniques. Human Reliability Associates Ltd. (2000)
17. Nielsen, J.: Usability Engineering. Academic Press, San Francisco (1993)