Towards a Cognitive Load Aware Multimodal Dialogue Framework for the Automotive Domain

Robert Neßelrath
German Research Center
for Artifical Intelligence (DFKI)
Saarbrücken, Germany
Email: robert.nesselrath@dfki.de

***Abstract*—A significant aim in developing multimodal HCI for the automotive domain is to keep driver's distraction low. However, the measurement of the cognitive load is difficult and inaccurate but an approach to predict the effect of dialogue and presentation strategies on this is promising. In this paper we discuss cognitive load in theory and related work, and identify dialogue system components that play a role for monitoring and reducing driver's distraction. Subsequently we introduce a dialogue system framework architecture that supports cognitive load prediction and situation-dependent decision making & manipulation of the HCI.**

***Keywords*—HCI, multimodal dialogue system, cognitive workload prediction, cognitive load measurement, automotive domain, development toolkit, system adaptation, multimodal integration**

I. Introduction

Due to the enhancement of on-board electronics in modern cars during recent years, the amount of information the driver receives has increased to a large extent. Traditional car displays and controls like speedometer, light, wiper settings or radio add to the information load that is produced by the actual traffic and environmental context. Nowadays, many modern cars offer much additional information, services and assistance systems for navigation, "infotainment", entertainment and comfort. It is to be expected that new technologies like car-to-x [1] will complement the information load with status reports about actual traffic and street situation. However, although it is generally accepted that this information is beneficial to increase safety and the driver's comfort, it cannot be denied that this flood of cognitive stimuli harbours the risk of distracting the driver from his primary task, namely to steer the car.

Driving-related activities can be classified into three levels [2]: The first level is the maneuvering of the car, e.g. steering and operating the pedals. The second is maintaining safety while driving, e.g. using windshield wipers, control of lights and turning signals. The third is the control of comfort, infotainment and entertainment functions. A lower priority level must not critically influence or disturb the tasks of a higher priority level. The US Department of Transportation [3] has announced its first set of guidelines for preventing driver distraction by prohibiting a lot of infotainment functions while the car is moving. Due to this upcoming restriction automobile manufacturers face the challenge to develop user interfaces that reduce the effect on cognitive load. This can be achieved by using different interface modalities or adapting the provided dialogue strategies in order to reach a certain goal. Unfortunately there exist only a few patterns and guidelines that support the HCI development process or give an a-priori prediction of the influence of dialogue and information presentation strategies on driver's workload and the distraction from his primary task.

In this paper we explore models and strategies for supporting development and evaluation of cognitive load-aware multimodal user interfaces. For this purpose, a multimodal dialogue system and an associated development toolkit are developed that allow the rapid and flexible creation of new dialogue applications. A great focus is therefore placed on a carefully considered model-based approach and a modular platform architecture. Models and architecture are defined with respect to our strategies for cognitive-load evaluation and prediction and are planned to be open to support situation adaptive applications in the future.

The paper is structured as follows. First we provide an overview of cognitive load theory and working memory models in cognitive science. By considering related work from HCI research we identify relevant dialogue system modules that affect the cognitive load. We introduce a multimodal dialogue platform that allows the rapid development of multimodal applications. Finally, we explain how relevant modules will provide cognitive load awareness in collaboration with cognitive load measurement and prediction modules.

II. Cognitive Load

A. Cognitive Load Theory & Working Memory Model

In psychology, cognitive load theory addresses the cognitive effort required when learning new tasks. The theory maintains that it is easier to acquire new knowledge and expertise if the kind of learning instruction keeps the cognitive load, and therefore the demand on a user's working memory, low [4] [5]. The theory differentiates between three types of cognitive load: *intrinsic load*, *germane load*, and *extraneous load*. The intrinsic load results from an interaction between the amount and type of the material being learned and the expertise of the learner. Extraneous load relates to the manner in which the material being learned is presented. The germane load is needed for processing the learned content and organize them into new schemata or activating existing ones. The three

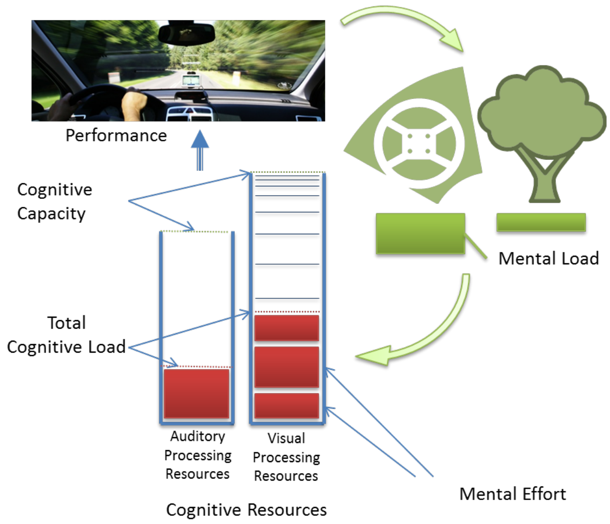

Fig. 1. Interplay between mental load, mental effort, performance and cognitive load. Mental load is imposed by stimuli and tasks. The mental effort is the actual allocated amount of cognitive load, that is individual for every user and distributed over different resources. The sum of the mental effort for all stimuli and tasks should not exceed the cognitive capacity, since it directly influences the task performance.

types are additive; together they build the overall load that should not exceed the cognitive capacity limit [6]. Paas and Van Merriënboer [7] describe assessment factors on cognitive load. The *mental load* is imposed by the task or environmental demands and is constant for a given task in a given environment, independent of a particular user's characteristics. The mental capacity actually allocated is represented by the *mental effort*. It is the outcome of the interaction between the task and the subject's characteristics. Thus, this represents the actual cognitive load on the individual. The quality of task solution is a third measure, the *performance*. It is influenced by the suspected mental load, the effectively invested mental effort and the individual prior knowledge and experience of the subject. Figure 1 depicts the simplified interplay between the above mentioned assessment factors.

Current theories for working memory are based on models which consist of multiple independent processors associated with different modes. Baddeley [8] [9] describes the two independent components *visio-spatial sketchpad* and *phonological loop* that are coordinated by a central executive module. The first processes visual input and spatial information, the second stores auditory-verbal information. The *four-dimensional multiple resource model* [10] divides resources into four categories/dimensions, postulating that there is a greater interference between two tasks when they share the resources of one category.

To sum up, the influence of tasks and cognitive stimuli on the cognitive load is dependent on various factors. These are the task difficulty, the individual experience of the user and the distribution of load among different working memory resources. Finally, the individual subject can have an active influence by ignoring factors of no importance and focusing on a specific task.

B. Measuring Effects of Cognitive Load

Several approaches have been used in psychology and HCI research to measure the actual amount of cognitive load. Generally methods can be classified in four categories.

1) Subjective Measures: A traditional way to assess the subjective workload of a user is *introspection*. The results are acquired by a questionnaire e.g. with the NASA Task Load Index (NASA-TLX) [11]. Because this method is an intrusive procedure and would add an additional task to the cognitive load it can only be done after the experiment. Beside other scales also in-depth interviews should help to gain more detailed information. Nonetheless, such measures cannot be used for real-time assessment.

2) Physiological Measures: One possibility for real-time assessment is to use physiological measures based on the assumption that the subject's cognitive stress is reflected in the human physiology [12]. Physiological indicators that have been used in previous research are heart rate, brain activity, galvanic skin response and eye activity (e.g. blinking or saccadic eye movements). Unfortunately, physiological sensors are often integrated in cumbersome equipment and it must be guaranteed that the method is non-intrusive before it is used in tests or real systems.

3) Performance Measures: Supposing that the performance of the task solution is influenced by the cognitive load, conclusions about the latter can be drawn from performance measures. Several concepts are used to estimate the performance. One is to evaluate the task processing requirements by considering the amount of time required, error rate and/or type of errors. Additionally, the response or reaction time to an stimulus event provides information about the actual cognitive load. One example for this is the Lane Change Test [13], that predicts the level of user distraction by measuring the reaction time of the driver to commands to change lane.

4) Behavioural Measures: Under high cognitive load users tend to change their interaction behaviour. [14] define *response-based behavioural features* as those that can be extracted from any user activity that is predominantly related to deliberate/voluntary task completion, for example, eye-gaze tracking, mouse pointing and clicking, keyboard usage, use of application, digital pen input, gesture input or any other kind of interactive input used to issue commands to the system. Characteristics of speech, such as pitch, prosody, speech rate and speech energy, can change under high cognitive load. Further features in speech which may indicate cognitive stress are high level of disfluencies, fillers, breaks or mispronunciations. Oviatt et al. [15] found out that users prefer multimodal, rather than unimodal, interaction when being confronted with new dialogue context or complex situations. Once users become more comfortable with the system, the percentage of unimodal interaction raises again. Another valuable finding is that, under stress, the informational load of utterances reduces to the minimal amount needed to achieve the communicative goal. For instance, [16] showed that use of anaphora in discourse optimize the semantic information processing in working memory.

III. Cognitive Load Prediction

Cognitive scientists try to explain the impact of task complexity and stimuli on the cognitive load by developing theories and computational models that calculate and predict the cognitive interference between different tasks, e.g. the computational multiple resource model [10].

The goal of Human-Computer Interaction researchers is not to explain human cognition in detail. In fact their research question is how presentation and interaction design affect the cognitive load of a user, especially in scenarios in which he controls safety-critical systems like flying an aeroplane, crisis management or steering a vehicle. Some projects treat this question and test strategies for manipulating the cognitive load with changes in interaction design for a multimodal system [17] [18]. [19] analyzed the impact of presentation features like font size and contrast on glance time for a visual display and [20] predicts presentation complexity on the basis of presentation layouts. Prediction approaches have also been researched for other modalities like speech [21].

Results from these studies can be used to find patterns and propose guidelines that help to develop interfaces with a low effect on the cognitive load. Since not every application designer will have adequate experience to apply these in practice, a system that predicts the complexity of an interaction design and supports the application developer in his work will provide a valuable benefit.

IV. The Cognitive Load-Aware Dialogue System

This project is composed of two interrelated topics. The first is the development of a multimodal dialogue system that supports functionalities that were examined in previous research projects at our institute [22]. The system will support multimodal and context fusion, discourse processing and multimodal fission. One focus is a carefully designed semantic model based approach that will, together with a good toolkit, ease the development process for application engineers [23]. The second topic is to encapsulate functionalities in clearly defined and independent modules that support the prediction of the cognitive load. Findings from related work introduced in the previous sections have helped us to identify four components that have a potential influence on it.

(A) *Dialogue Manager* - The strategy for solution of the task has influence on the cognitive load of the user. Different strategies are conceivable collecting the necessary information for solving a task, e.g. to reserve a cinema seat. One approach is to collect all information at once: A GUI modality would provide one form to collect all values of required parameters, to use speech dialogue the system would allow more complex and content-rich utterances. Another approach is to collect the needed information step by step by asking the user in a question-answer-based speech dialogue or with several GUI windows.

Our model for dialogue specification is a combination of flow- and state-charts that uses semantic dialogue acts as defined in [24]. The concept of abstract dialogue-acts keeps presentation management independent from the dialogue manager. Thus, the decision on which modalities will be used and how modalities present the content are made separately in the presentation planner.

(B) *Presentation Planner & Multimodal Fission* - Related cognitive research showed that multimodality has a great effect on the cognitive load [15] [18]. In the previous section we discussed that the realization of unimodal presentation is also relevant for user attention. Hence the presentation planner module has an impact on cognitive load by deciding which modalities are used and how presentation is realized.

(C) *Multimodal Fusion and Context Resolution* - This component can optimize load on working memory by using context information in order to resolve referring expressions, e. g. anaphora. Thus, according to [16], the activation cost for semantic content would be reduced.

Figure 2 shows our concept for cognitive load awareness in a multimodal dialogue system. It it inspired by [25], which introduces a method for a cognitive load-aware presentation manager. The three above-mentioned components mainly define and generate the human-machine interface (HMI) of the dialogue system that is part of the driving context. This context directly affects the cognitive load of the driver and may have influence on his driving performance.

One supported component of the dialogue platform will be a prediction module for cognitive load. This module will have access to input data from the driving context, the driving performance and also possible techniques for measurement of the actual cognitive load of the user. The module and its algorithms will be adaptable for different use cases, theories and measurement methods. Thus, the system will be able to support on the one hand more pragmatic heuristic ap-

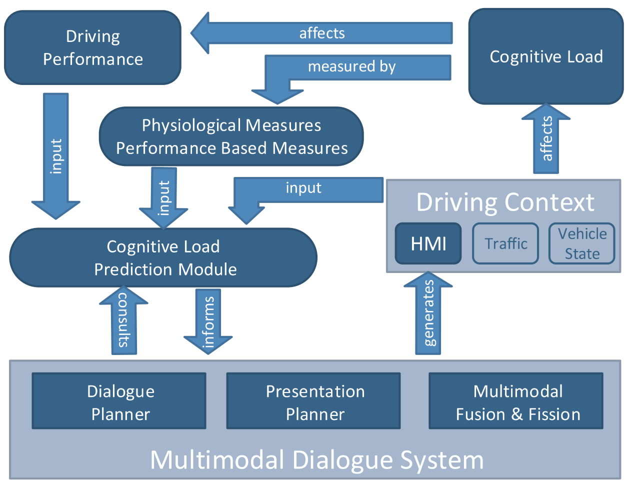

Fig. 2. Relevant dialogue system components have a direct influence on the cognitive load and the driving performance. These can be measured and estimated in order to provide situation adapted dialogue and presentation strategies.

proaches for use in live-applications and on the other hand the evaluation of more complex models from cognitive science. The components of the system relevant to the cognitive load will cooperate with the prediction module in order to plan situation-aware behaviour of the HMI. The architecture will also support future plans to build situation-adaptive systems that change their behaviour during runtime with respect to the actual cognitive load of the driver.

V. Conclusion

The paper presents the research topic of my PhD project, which is to develop a cognitive load-aware multimodal dialogue system. By citing cognitive science theory and related work from human-computer interaction research we identified the following dialogue system features that have an influence on the cognitive load of a user:
Multimodality, *presentation planning*, *dialogue planning* and *supporting context resolution and referential expressions*
With regard to these aspects, we designed a multimodal dialogue system architecture that will support the rapid development of multimodal applications and is open to provide the following functionalities relevant for cognitive load:

(a) *Support for application developers* - During the design process, dialogue platform tools can advise the developer with cognitive load prediction for dialogue and presentation strategies.
(b) *Support for cognitive science* - The modular architecture approach allows the easy adaptation or replacement of components for cognitive load measurement and prediction. With an adequate development toolkit, the validation of theories and models from cognitive science with live experiments can thus be improved.
(c) *Support for situation adaptive systems* - A future goal is to build situation-aware systems that adapt their communication behaviour with respect to the current context and cognitive load.

Acknowledgements

This work was funded by the German Ministry of Education and Research (project SiAM: Situation-Adaptive Multimodal Interaction for Innovative Mobility Concepts for the Future; grant number 01IW11004).

References

[1] H. Stubing, M. Bechler, D. Heussner, T. May, I. Radusch, H. Rechner, and P. Vogel, "simTD: a car-to-X system architecture for field operational tests [topics in automotive networking]," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 148 –154, May 2010.

[2] C. Müller and G. Weinberg, "Multimodal input in the car, today and tomorrow," *IEEE MultiMedia*, vol. 18, no. 1, pp. 98 –103, jan. 2011.

[3] Department of Transportation. (2012, Feb.) Visual-manual nhtsa driver distraction guidelines for in-vehicle electronic devices. [Online]. Available: http://www.webcitation.org/6Fzc9xMO2

[4] P. Chandler and J. Sweller, "Cognitive load theory and the format of instruction," in *Cognition and Instruction*, L. E. Associates, Ed. Taylor & Francis, 1991, pp. 292–332.

[5] J. Sweller, "Cognitive load during problem solving: Effects on learning." *Cognitive Science*, vol. 12, no. 2, pp. 257–285, 1988.

[6] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. V. Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Educational Psychologist*, vol. 38, no. 1, pp. 63–71, 2003.

[7] F. G. Paas and J. J. Merriënboer, "Instructional control of cognitive load in the training of complex cognitive tasks," *Educational Psychology Review*, vol. 6, no. 4, pp. 351–371, 1994.

[8] A. Baddeley, "Working memory: looking back and looking forward," *Nature Reviews Neuroscience*, vol. 4, pp. 829–837, 2003.

[9] A. Baddeley, "Working memory: theories, models, and controversies," *Annual review of psychology*, vol. 63, pp. 1–29, 2012.

[10] C. D. Wickens, "Multiple resources and performance prediction," *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, pp. 159–177, 2002.

[11] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Human mental workload*, vol. 1, no. 3, pp. 139–183, 1988.

[12] A. E. Kramer, *Physiological Metrics of Mental Workload: A Review of Recent Progress*. Defense Technical Information Center, 1990.

[13] C. Endres, R. Math, and D. Braun, "Simulator-based evaluation on the impact of visual complexity and speed on driver's cognitive load," in *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)*, Portsmouth, New Hampshire, USA, October 2012, pp. 30–31.

[14] F. Chen, N. Ruiz, E. Choi, J. Epps, M. A. Khawaja, R. Taib, B. Yin, and Y. Wang, "Multimodal behavior and interaction as indicators of cognitive load," *ACM Trans. Interact. Intell. Syst.*, vol. 2, no. 4, pp. 22:1–22:36, Jan. 2013.

[15] S. Oviatt, R. Coulston, and R. Lunsford, "When do we interact multimodally?: cognitive load and multimodal communication patterns," in *Proceedings of the 6th international conference on Multimodal interfaces*, ser. ICMI '04. New York, NY, USA: ACM, 2004, pp. 129–136.

[16] A. Almor, "Noun-phrase anaphora and focus: The informational load hypothesis," *Psychological Review*, vol. 106(4), pp. 748 – 765, October 1999.

[17] A. Niculescu, Y. Cao, and A. Nijholt, "Manipulating stress and cognitive load in conversational interactions with a multimodal system for crisis management support," in *Development of Multimodal Interfaces: Active Listening and Synchrony*, ser. Lecture Notes in Computer Science 5967, A. Esposito, N. Campbell, C. Vogel, A. Hussain, and A. Nijholt, Eds., Heidelberg, 2010, pp. 134–147.

[18] S. Oviatt, R. Coulston, S. Tomko, B. Xiao, R. Lunsford, M. Wesson, and L. Carmichael, "Toward a theory of organized multimodal integration patterns during human-computer interaction," in *Proceedings of the 5th international conference on Multimodal interfaces*, ser. ICMI '03. New York, NY, USA: ACM, 2003, pp. 44–51.

[19] D. Imbeau, W. W. Wierwille, and Y. Beauchamp, "Age, display design and driving performance," *Automotive ergonomics*, pp. 339–357, 1993.

[20] C. Endres, M. Feld, and C. Müller, "A layout-based estimation of presentation complexity," in *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)*, Portsmouth, New Hampshire, USA, October 2012, pp. 19–20.

[21] V. Demberg and A. Sayeed, "Linguistic cognitive load: implications for automotive UIs," in *Adjunct Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI 2011)*, 2011.

[22] N. Pfleger and J. Schehl, "Development of advanced dialog systems with PATE," in *Proceedings of Interspeech 2006—ICSLP: 9th International Conference on Spoken Language Processing, Pittsburgh, PA, USA*, 2006, pp. 1778–1781.

[23] R. Nesselrath and D. Porta, "Rapid development of multimodal dialogue applications with semantic models," in *Proceedings of the 7th IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems (KRPD-11). Twenty-Second International Joint Conference On Artificial Intelligence (IJCAI -11)*, Barcelona, Spain, July 2011.

[24] ISO, "Language resource management, semantic annotation framework, part 2: Dialogue acts," International Organization for Standardization, Tech. Rep. ISO 24617-2, 2012.

[25] C. Endres, "Real-time assessment of driver cognitive load as a prerequisite for the situation-aware presentation toolkit PresTK," in *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2012)*, Portsmouth, New Hampshire, USA, October 2012, pp. 76–79.