ABSTRACT

Besides travel related information driver related information plays an increasing role while driving. In order to process this information and be able to drive safely at the same time, an information management is necessary in the vehicle. Speech interfaces offer a good solution since verbal interactions use resources which are not needed when driving. However, in practice experiments show that verbal interactions may bind attention. It is crucially important that verbal interactions are designed in such a way that for the driver the task "driving" remains in the focus of attention.

KEYWORDS

Divided attention, human-machine-interaction, speech interface, secondary tasks while driving

INTRODUCTION AND BACKGROUND

Sources of information while driving

More and more information is made available to the driver which is partially related to travel or vehicle, additionally to other areas of interest. The travel related information affects decisions while driving: The tachometer is needed to comply with speed limits and supports the adjustment to environmental condition like the road condition. Revolution speed is important for gear switching and the visual presentation becomes the more important the less the engine noises can be perceived acoustically. Fuel gage leads to short term adjustments of the route. On board computers provide additional information useful for driving. For example, at low temperatures the driver can reduce the speed because slippery roads are to be expected. This travel related information is extended by driver information systems and advanced driver assistance systems. Driver information systems like the navigation system include information about traffic jams into route planning and give additional information which may lead to route choice changes (e.g. interesting places in the surrounding, leisure-time facilities etc.). Advanced driver assistance systems support different aspects of the driving task. For example, the system Adaptive Cruise Control (ACC) takes over speed control and ensures safe distances toward preceding vehicles. If no other cars are present a selected speed is driven (cruise control function). With this system, on the one hand the driver is relieved of the longitudinal control in certain situations (like driving on the highway). On the other hand this
system provides additional visual (e.g. system active, distance towards preceding car) and acoustic (e.g. a beep if a driver action becomes necessary) information.

Moreover, information not related to the vehicle or the trip becomes widely available in the vehicle. This concerns entertainment function (e.g. radio, CD, MP3) and additional interests of the driver (e.g. telephone calls, internet etc.). Especially the use of handheld cellular phones during driving has increased during the last years, for example, in Michigan from 2001 to 2005 from 2.7% to 5.8% [1]. The information is offered on different channels. Apart from the acoustic presentation of music and speech (telephone, reading out e-mails by text-to-speech systems) much information (e.g. texts and diagrams from the Internet) is offered visually. Moreover, many of these systems require operating buttons, keys or other controllers.

Since perception, information processing and manual control is also needed for driving safely, the question arises which information can additionally be processed by the driver and how to avoid interferences with the driving task. In order to answer this question the demands exerted on the driver in order to drive safely are presented in an overview in the next section.

Requirements of the Driving task

During a trip the driver continuously plans and executes driving manoeuvres including the longitudinal and lateral stabilization. To accomplish this different resources are needed (e.g., [2-4]). First, different stages of the action are distinguished. The driver perceives information from the environment (perception), decides whether and which actions are required (cognition) and responds (responding). Secondly, different modalities are used in perception. Different resources are assumed for the visual and auditory modality. Third, in both modalities different codes or kinds of information can be transmitted. Spatial and verbal resources are distinguished here. Finally, for the responses either manual or vocal responses are possible which again represent different resources.

For example, the driver monitors the position of the car within the lane. If larger deviations occur, a correcting action is executed via the steering wheel. In this case, spatial information is processed coming from the visual modality. This results in a manual reaction. The same holds true for keeping a certain speed and distance towards preceding cars. Overall, in car driving the use of the visual modality, spatial codes and manual reactions dominate.

The multiple resources model includes two assumptions: (1) Different resources exist which can be accessed independently. (2) Each of these resources may be shared but is limited. These assumptions implicate that two tasks interfere stronger, if their execution requires the same resources. On the other hand, two tasks should be executable at the same time if they stress different resources. During driving mainly visual information is required, spatial codes are processed and a manual reaction takes place. Thus, any action that processes verbal information which is perceived with the auditory channel and which encompasses vocal reactions should be easy to do while driving. In contrary, the visual information presented by a navigation system including spatial information and requiring manual responses should disturb driving as the same resources are needed.

From these considerations it is derived to not stress the driver visually and increase using the auditory channel. On the one hand these considerations appear highly plausible. On the other hand some experimental findings contradict the multiple resources theory (for a discussion,
In the following section selected experimental findings from telephoning while driving are discussed to find out under which circumstances auditory perception, verbal codes and vocal responses may lead to negative consequences.

**SPEECH AND DRIVING**

**Selected effects of telephoning while driving**

In order to analyse the structure of effects of dialling a number in comparison to having a conversation on the phone a review was conducted including 19 studies [7-25]. Different aspects of driving behaviour were examined in these studies. The review summarizes these results with regard to lane keeping, speed and distance regulation and reaction to special stimuli. Ten studies were done in a driving simulator, five in real traffic and two on test tracks. In two additional studies tests were conducted in a laboratory to examine certain tasks similar to driving. The studies used different speech tasks to induce heavy cognitive load of the driver. Figure 2 gives the results.

Dialling substantially impairs the lateral control of the vehicle. 43% of the studies found an impaired lane keeping. In 50% of the studies, leaving the lane occurred more often. Sometimes the drivers drove slower (20% of the studies). The distance towards preceding cars did not change. If responses to external stimuli were required dialling did not impair the reaction time but more errors occurred (33% of the studies). When having a conversation on the phone, lane keeping was impaired in 21% of the studies. However, other 21% found an improvement of lane keeping. Leaving the lane occurred less often during conversation than while dialling (25% of the studies). In 56% of the studies driver went more slowly while talking. However, in 60% of the studies smaller distances towards preceding cars were found. Finally, 63% of the studies showed prolonged reaction times and more errors (40% of the studies).

![Figure 2: Effects of dialling and having a conversation on the phone](image-url)

Overall dialling clearly impairs lane keeping while the longitudinal control changes only insignificantly. This may be explained by the short duration of this activity. Thus, significant effects on longitudinal control might be expected if these tasks took longer. Similarly, reaction time to sudden events is not changed but the number of errors increases. In summary,
in accordance with the multiple resources model dialling disturbs the driving task since same resources are used.

When having a conversation on the phone the effects in the area of lateral control are smaller. The effects in the area of longitudinal control show that the drivers reduce speed. This may be a compensatory response to make the driving task easier. However, the smaller distances towards preceding cars combined with slowed down reactions and more frequent errors may lead to dangerous situations during car follow. Comparable effects are found in newer studies not included in this review (e.g. [1, 26-35], to mention just a few of the many studies). This interpretation is supported by accident studies demonstrating an increased accident risk when telephoning [36-40]. This negative effect of phoning is hard to explain with the multiple resources model. It cannot be explained by drivers holding the phone as in 10 of the 19 studies a hands-free phone was used. In a study where the effects of handsfree and handheld phones were compared, both modes lead to an increased workload [26]. Additionally, while drivers compensated talking on a handheld phone by reducing speed, this was not found in handsfree phone conversations.

On the one hand the results of this short review support the hypothesis that perceiving additional visual information and handling the phone while dialling substantially impairs driving. According to the multiple resources model it should be possible to avoid these interferences by using auditory information and speech. Unfortunately, this hypothesis is not supported by the review as having a conversation was shown to impair driving even when handsfree phones are used. This conclusion is supported by other studies showing that conversations themselves decrease performance while driving (e.g., [32, 33]). Other factors not included in the multiple resources model have to be considered to explain this. The following section examines under which circumstances negative effects of verbal interactions arise.

The influence of complexity and emotional quality of verbal interactions

Studies examining the influence of passengers on driving show that driving with passengers decreases accident risk as compared to driving alone (e.g., [41,42]). This result points to the interesting fact that under certain circumstances verbal interaction during driving can even increase safety. In a set of experiments in our laboratory the effect of different kinds of interactions were examined in a simple driving simulation (for the complete description, see [43]). A simplified representation of a car had to be kept on a curvy road by means of a video game steering wheel, gas pedal and brake. Additionally, at some points of barriers suddenly appeared in front of the car. In order to avoid a ‘crash’ drivers had to break fast and hard. These parts of the driving task were termed ‘freeway driving’. Within certain visually defined areas these barriers appeared very frequently. A crash could only be avoided if the speed had been reduced when entering this area. These parts were termed ‘city driving’.

In a control condition the subjects drove with a quiet passenger on the right front seat. Three different interaction conditions were induced as role-plays. In the condition ‘small talk’ subjects talked about inconsequential, simple topics without large emotional or cognitive involvement. The condition ‘argument’ induced a negative emotional atmosphere where both participants tried to win the upper hand in a controversy. Finally, in the ‘complex’ condition subjects invented a story together by alternatively producing a sentence. In this condition the analysis was also carried out separately for periods of speaking and listening. For the analysis,
the number of crashes was counted when the drivers were not able to avoid the barrier. From this, a relative accident risk was calculated where the number of accidents with a quiet front seat passenger served as reference, i.e. the relative risk in this condition was set to "1". Figure 3 shows the relative risks in the different conditions.

Figure 3: Relative risks for different conversations on a simulated freeway and city

During city driving both small talk and argument clearly increase accident risk. The strongest effect is found for the complex interaction. However, the increased accident risk in this condition is limited to speaking and not found while listening. At first glance a completely different picture emerges for freeway driving. Here only the argument condition leads to an increased accident risk. In the complex interaction the accident risk is even decreased and particularly when speaking. The explanation for these differences is found when speed is considered during freeway driving (s. Figure 4).

Figure 4: Average speed in the different conversation situations

The speed while driving with a quiet passenger was set to 100%. The speed in the other conditions is given as a percentage in relationship to that speed. Speed is reduced during small talk and the complex interaction and particularly when speaking during the complex interaction. In contrary, speed is hardly reduced in the argument condition. The speed reduction in the other conditions can be understood as efforts of the driver to compensate the
increased cognitive load caused by the interaction. Drivers know that talking reduces their attention for the driving task and that they cannot react as effectively and drive as safely as without talking. To counteract this, drivers reduce their speed which represents an effective means to provide more time for any necessary reactions. During city driving the speed is already reduced so much that a further reduction of the speed is no longer possible (and thus not shown in the Figure). Thus, in city driving the negative effects of the interactions become visible as an increase of the accident risk. The compensatory reaction of speed reduction on the freeway is not done in the argument condition resulting in an increased accident risk in this situation. This condition was described by the drivers as the most emotional and involving situation. Additionally, the aggressions resulting from the conversation work against a relaxed driving.

To summarize: (1) Speaking impairs driving while listening appears unproblematic. (2) Impairment due to speaking is found when the interaction is either complex (see also [28, 33]) or emotional. (3) Drivers try to compensate for the negative effects. (4) However, this is not the case if the interaction is emotional or if compensation by reducing speed or increasing distances towards preceding cars is not possible due to driving conditions. The results of the study cited above comparing handsfree and handheld phones [26] make an additional point: The drivers compensate only if they think that they are distracted as in the handheld phone condition. Thus, the positive effect of keeping the hands free while phoning may be neutralised as drivers do not reduce their speed when phoning with their hands free.

Overall, introducing a verbal interaction between driver and system (speech input and acoustic output) may be a good means to avoid interferences with driving. However, additional constraints have to be considered. The verbal interaction between drivers and the technical system must be arranged in such a way that the driver does not become emotionally involved, that the conversation does not become too difficult and that the driver is able to compensate by reducing the difficulty of the driving task. One important aspect with regard to this is the quality of speech recognition and of the speech used to interact with the driver. Degraded speech occurs quite often in the moving car (e.g. [44]) and requires additional resources from the driver to understand what is said ([44]). From this point of view it is essential that in cars only high quality speech output is used. Additionally, if recognition of speech is bad, this will probably involve the driver more strongly in the conversation which again leads to negative effects. Thus, the first measure to ensure that negative effects of verbal man-machine-interaction are prevented is to ensure a good quality of speech output by the systems and a high recognition rate. In the future, an appropriate support of advanced driver assistance systems could be used to support the compensatory efforts of the drivers. For example, one could use the initially described ACC to reduce speed and keep larger distances when the driver interacts verbally with the vehicle.

Avoiding emotional interactions seems to be easy at first glance since driver-vehicle interactions usually do not comprise emotional contents. However, each interaction has a social aspect where even a technical system is regarded as some kind of partner to whom certain characteristics are attributed (e.g., energetic, active, dull etc.). Moreover, the partner himself is evaluated in categories such as like/dislike. For these evaluations the nonverbal aspects of an interaction are especially important, i.e. how the interaction is performed. This is demonstrated in the next section.
**Influence of Nonverbal aspects of speaking**

Information about the speaker and the kind of interaction may be gathered independently of the content by analysing the nonverbal aspects of speaking. To this aim, Krüger introduced the method of the Speech Chronemics to examine the time structure of speaking on different hierarchical levels [45]. These time factors are especially important in dialogues. To demonstrate this, the results of an experiment from our laboratory with 6 pairs of male partners are presented. They performed 3 different interaction scenarios as role-plays [45]. A small talk situation served as a control condition where subjects talked about inconsequential topics. In the condition ‘intense positive interaction’ subjects were supposed to talk about a personal and intense situation. Both partners were instructed to open up to each other, to listen closely and to try to openly talk about one’s feelings. In the condition ‘argument’ a controversial topic was selected and each of the partners was instructed to emerge as the winner of the argument. During the interactions, the duration of four dialogue events was measured. ‘Undisturbed speech’ referred to time periods where only one partner was talking. If the partner interrupted and both spoke this was termed ‘double talk’. If one partner stopped and then started again, this pause was called ‘isolated pause’. When the other partner started after the pause, this was termed ‘switching pause’. Figure 5 shows the mean duration of the events in the three dialogue conditions.

![Figure 5: Mean duration of dialogue parameters in different kinds of conversation](image)

In the intense interaction as compared to small talk undisturbed speech as well as switching pauses were substantially longer. The other parameters did not differ between these two interactions. Speakers may continue speaking for longer time in the intense communication. The listener waits for a longer time when the speaker has stopped before he himself begins to talk. The atmosphere of this interaction to open up to the other one and let him talk is reflected very well in this pattern. In the argument condition, undisturbed speech, isolated pauses and switching pauses become shorter while the duration of double talk increases. In this situation, the speaker is interrupted frequently so that he is not allowed to speak undisturbed for a long time. When he stops speaking, the partner begins to talk very fast. Again, the atmosphere of this kind of interaction is described very well with the nonverbal dialogue parameters indicating that both of the partners want to win in this situation and try to get the upper hand.
Measuring these very basic indicators of the temporal structure of a dialogue allows one to describe the atmosphere and the intention of the dialogue quite well. It should also be possible to use this knowledge to actively influence the atmosphere of a dialogue by manipulating these basic temporal structures. One could use this in the context of a man-machine interaction to create a certain atmosphere and impression of the machine as a dialogue partner. This was examined in a first pilot study using a simulated advanced driving assistance function (ADAS) which interacted verbally with the driver who manoeuvred the car within a simulated city by voice commands (‘next right’ etc.). The ADAS acknowledged the command and drove the vehicle automatically. The acknowledgement took place either quite fast or quite slowly [46]. N = 30 test drivers compared these two versions of the ADAS. After each trip, on the one hand the functionality of the system was evaluated. On the other hand the test drivers evaluated the system with regard to the impression which it made on the driver. Figure 6 shows the results.

![Figure 6: Subjective evaluation of two systems differing in reaction time](image)

At the left part of the figure the answers of the drivers indicated that the different reaction speed of the system is very obvious to the drivers as the system with short pauses is described as being faster. It is interesting to note that the faster system is also perceived as the better system. At the right, different characteristics of the system are given. The different reaction speed does not affect perceiving the system as likeable, active, being in a good mood or extraverted. However, the system with a slow reaction appears less efficient. Thus, a simple variation of reaction speed is sufficient to create a certain impression of the system. The system which answers quite fast is regarded as being more efficient and overall as the better system.

This section demonstrates that basic nonverbal aspects of an interaction are essentially important for the emotional quality of an interaction. By manipulating the temporal structure of the verbal man-machine-interaction a certain impression of the system and a certain quality of the dialogue can be obtained. As shown above, the negative effects of speaking depend in part on the emotional quality of an interaction. Thus, more emphasis should be put on designing the interaction in a way which does not lead to an overly deep involvement and distraction from the driving task.
SPEECH AND DRIVING - SOLUTION AND PROBLEM

The starting point of the paper was the discussion about possible problems resulting from the increasing amount of information available to the drivers. Perception, information processing and reactions are required from the driver to cope with this additional information and this may interfere with the driving task. Resources of the driver are limited, i.e. when the same resources are needed for driving and additional tasks interference may result which can impair driving even if this additional information is meant to support the driver. From the multiple resources model it can be derived that the presentation of verbal information which is perceived via the auditory channel and where verbal reactions (speech input to the system) is used offers a way out since other mental resources are used than in driving. However, a review on the effect of telephoning and experimental studies on speaking while driving showed that this is not always true. If the interaction is complex or the partners become involved emotionally interferences arise and driving is impaired despite using different resources. These effects are particularly pronounced when the driver is speaking and less when he is listening. This is probably also true for a bad quality of speech output and recognition which both could lead to a stronger involvement in the interaction and distract the driver. Additional studies on nonverbal aspects of speaking show that emotions do not only result from a certain content but may also be provoked by the manner of speaking (e.g., volume, pitch) and temporal aspects of the interaction. Manipulating even simple temporal factors in a dialogue affects the evaluation of the system and its effectiveness.

Thus, speech commands or verbal man-machine interaction offer a solution to the increasing amount of information available for drivers and may enable safe driving. However, the following aspects have to be considered:

1. Verbal operation must be simple. Complex content and difficult interactions involve the driver cognitively and may distract from driving. The system may provide complex content (listening is not as problematic as speaking) but the answers of the driver and the speech commands must be simple.
2. The quality of the speech output of the system and the recognition rates for speech have to be high in order to avoid larger workload for the driver and increasing involvement in the dialogue.
3. Moreover, nonverbal aspects of speech of the system have to be designed in a way that a positive evaluation of the system by driver results.
4. The dialogue behaviour (e.g. interruption, switching times etc.) of the system should be designed in way that a neutral, small talk like interaction results. Anger or other negative emotions have to be avoided.
5. Advanced driver assistance system could be used to support the driver to actively adopt a compensatory behaviour (speech reduction, larger distances towards preceding cars) so that driving remains safe.

Thus, on the one hand restrictions for verbal man-machine interaction have to be introduced in the vehicle. On the other hand new possibilities result, since nonverbal characteristics can be quite easily manipulated in a technical system. In certain situations e.g. during a monotonous night travel during which the danger of falling asleep is the main problem, an emotionally involving, activating discussion with a system could work quite well. The adjustment of speech behaviour with regard to the driving situation in a way like a passenger does could work well to support safe driving even further. In this sense speech implying a well adapted verbal man-machine interaction may really be the solution for the problem of introducing additional information in the vehicle.
ACKNOWLEDGEMENTS

The work of this paper was conducted within the project „Strategies for driver assistance“, which is funded by the Ministry of Economics, Labor and Transport of Niedersachsen.

REFERENCES


