Kognitions- und
Ingenieurpsychologie

TECHNISCHE UNIVERSITÄT
CAROLO-WILHELMINA
ZU BRAUNSCHWEIG

IN-CAR DISTRACTION STUDY

FINAL REPORT

Braunschweig, 9.7.2008

Prof. Dr. Mark Vollrath
Dipl.-Psych. Jannette Maciej
Dipl.-Psych. Ute Niederée
Gaußstraße 23
38106 Braunschweig
Tel: +49-531-391-3644
Fax: +49-531-391-8181
Email: mark.vollrath@tu-braunschweig.de

Contents

1. Background and overview

The objective of the study, commissioned by Nuance, was to examine the benefits of speech operation in comparison with manual operation of various in-vehicle information systems. For this purpose, a simple but valid driving simulation test (the Lane Change Task) was carried out at the HMI laboratory of the Technical University of Brunswick, Germany. Thirty subjects operated different systems while driving the simulator, once manually and then again using voice commands. The varying levels of distraction were examined by recording driving behaviour, eye movement, subjective assessments by the driver and the evaluation by the study manager.

The main results of the analyses showed:

- that manual operation of the different systems resulted in greater distraction. This was demonstrated by the increased number of times drivers looked away from the road, poor lane keeping, delays in vehicle guidance and negative subjective assessments.
- for all systems, speech operation significantly reduced the number of times drivers looked away from the road (in comparison with manual operation), resulting in a substantial improvement in driving behaviour, which was also reflected in drivers' positive assessments.
- that good design of the voice human-machine dialogue reduced distraction still further, to the extent that in simple driving situations at least there was no perceptible difference in comparison to driving without operating any of the in-vehicle information systems.

2. Method and implementation

2.1. The Lane Change Task (LCT)

The Lane Change Task (LCT, Mattes, 2003; Kuhn, 2005) is a simple driving simulation carried out on a PC with a joystick steering wheel (see illustration 1). The test subject drives on a three-lane road, where signs (see middle of illustration 1) continuously indicate various lane changes to be carried out as quickly as possible. The test measures how well the driver follows a specified ideal path. To measure and compare the level of distraction caused by various secondary tasks, each task is carried out while driving (see illustration 1, left, with a navigation device). Typically, increased distraction resulted in a corresponding decrease in driving performance.

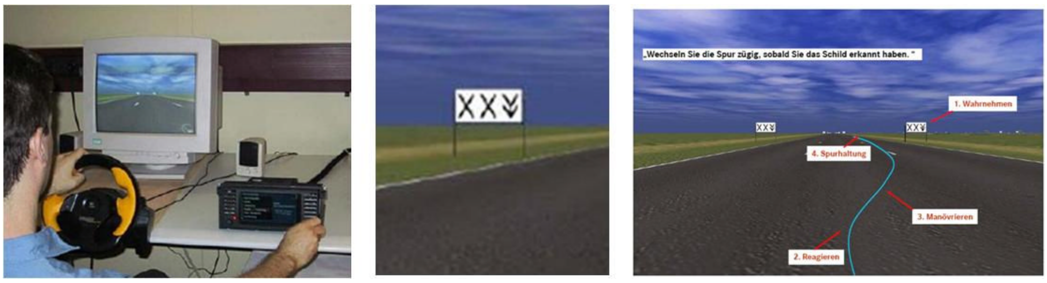

Illustration 1: The Lane ChangeTask (Mattes, 2003).

This task was developed by vehicle manufacturers and is currently advocated as an international standard (ISO/DIS 26022) to record standardised distraction effects.

Each test includes 18 random lane changes to ensure that the driver cannot predict which lane he or she will need to move to. A constant speed of 60 km/h is maintained to allow the driver to concentrate on lane keeping and lane changes. It takes around three minutes to complete a test.

2.2. Study plan, systems and flow

In the first instance, each test subject had to practice at least six circuits of the LCT. Our own investigations had shown that this ensures a stable level of performance. At the start of each test session, drivers completed a circuit without any in-vehicle systems to serve as a baseline (control condition). Drivers then operated each system once manually and then again using voice commands, based on a random sequence to control time effects. Drivers first practised a specific task (e.g. address entry) while stationary until they had understood it and could carry it out confidently. They then completed two circuits of the Lane Change Task, the first to practice operating the systems while driving and the second to carry out the evaluation.

To ensure continuous system operation, the study manager announced each new task verbally as soon as the driver had completed the previous one. Tasks were prepared in the form of lists, arranged in a random sequence. Each task was carried out both manually and using speech commands:

1) Music selection (“Audio”): the study manager specified an artist, an album or a title which the test driver then had to select.
2) Telephone (“Phone”): the study manager stated a name which the driver then had to select and call from the stored phone book.
3) Navigation – Address (“Navigation – Address multiple confirmation”): the study manager asked drivers to enter an address (town, street, house number). The device used (Falk N240L) required confirmation of recognition after each input step (town, street, house number).
4) In comparison, as an additional condition, another device (Navigon 8110) which required confirmation just once, after entry of the full address, was introduced (“Navigation – Address single confirmation”)[1].
5) Navigation – Points of Interest (“Navigation – POI”): the study manager stated various points of interest which drivers then needed to locate in certain towns.

Since the main objective was to assess the effects of the different input methods, there was no particular emphasis on whether the entries made by drivers were correct. The objective was more to assess whether drivers were able to operate the device as continuously as possible during the circuit, to obtain meaningful conclusions about the effects of the different types of operation.

An iPod was used for the manual audio task and an experimental Nuance prototype for voice operation (which has a similar UI to Ford SYNC). The Bury CC9060 hands-free car kit was used for phone operation and the above-mentioned Falk and Navigon devices for navigation.

The devices (iPod, hands-free car kit, mobile phone, navigation systems, computer with media player) were provided by Nuance. They were positioned beside the steering wheel as shown in illustration 1 and then operated during the LCT test.

---

[1] In this case, no separate condition with manual operation was introduced, instead manual operation of the first device (Falk N240L) was used as a comparison. Manual operation of the two devices was so similar that no different results were to be expected.

2.3. Measurements

Performance in the LCT was determined as the deviation between the actual path driven by the test subjects and a defined ideal path (see the blue line in comparison with the green line in illustration 2). This includes both reaction errors, where the driver responds too late to a prompt to change lane or completely overlooks a sign, as well as uncertainty in lane keeping, described in the literature as typical distraction effects. This deviation is calculated as a mean value (Mean Deviation). In addition to calculating the total deviation over the entire circuit, it is useful to distinguish the lane keeping phases from the lane change phases, as shown in illustration 2.

In the literature, the "standard deviation of lateral position" measurement is also used for lane keeping tasks. This is calculated as the standard deviation in a driver's position on the road. It describes the extent to which a driver "weaves" within a lane, i.e. the quality of his or her lane keeping.

For lane changes, as well as the deviation another indicator felt to be relevant was the reaction time, i.e. the time between the appearance of the sign and the first measurable movement of the steering wheel.

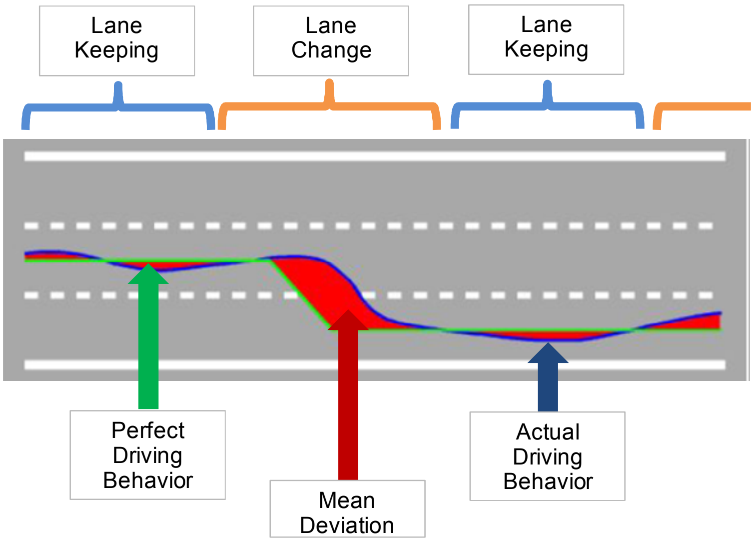

Illustration 2: Measurement of driving behaviour. The green line indicates the ideal path, the blue line shows an example of a test driver's path. The deviation between the blue line and the green line (shown in red) describes the extent of the distraction. A further distinction can be made between the lane change and lane keeping phases.

This results in the following measurements to describe driving behaviour:

- Mean Deviation in metres over the entire circuit
- Mean Deviation in metres during the lane keeping and lane change phases
- Standard Deviation of lateral position in metres during lane keeping
- Reaction time in seconds to a lane change prompt

In addition, to interpret the values more clearly, the mean value of the baseline circuit (without in-vehicle systems) was set to 100% each time and related to the mean values of the experimental conditions to allow all results to be shown as percentages in the presentation.

For gaze behaviour, trained observers divided gazes into gazes at the road and at the secondary task. For each circuit, the number of gazes, their average duration (in seconds) and the total time the driver's eyes were diverted from the road (as % of the driving time) were calculated.

After each condition, the test subjects answered a few short questions about the perceived effort, how well they rated their driving, the handling of the task and the ease-of-operation. Illustration 3 shows the possible responses for how well drivers felt they had driven. This two-step process (a broad verbal description followed by more detailed differentiation with "-", "0" and "+") provided very reliable values.

How well did you drive?

very bad			bad			medium			well			very well		
-	0	+	-	0	+	-	0	+	-	0	+	-	0	+
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Illustration 3: Example of a question in the subjective assessment.

Finally, after each circuit, the study manager evaluated each test driver's performance and perceived stress regarding the primary "Driving" .

2.4. Test subjects

The test set consisted of 30 test drivers aged between 19 and 59. 16 men and 14 women were assessed. All drivers held a driving licence. This wide age distribution ensured that the results could be transferred to different age brackets and were applicable to both sexes. However, a more detailed analysis, for example of younger versus older drivers, was not possible. This would have required further tests based on a targeted selection of individuals.

3. Results

The parameters described above are presented below. The presentation also includes percentage conversions of the values (see above), as well as comparisons of individual conditions. Since they do not provide any additional information in terms of content, no detailed description is provided in this report.

3.1. Summary of the main results

Illustration 4 shows as the main result the mean deviation over the entire circuit. The statistical evaluation[2] shows that for all systems voice operation resulted in significantly lower deviations than manual operation. Manual operation resulted in deviation values (around or more than 1.2 metres) described as critical. The benefits of voice entry were greatest for audio, followed by navigation address entry (multiple and single confirmation). For phone and entry of POIs, the benefits obtained from voice entry (lower deviation) were least.

For audio, phone and single confirmation address entry, voice operation resulted in a statistically significant higher deviation than driving without any in-vehicle systems. However, the deviation

---

[2] For the four tasks and the two operating methods, a two-factorial variance analysis was calculated for dependent random samples (within-subjects design). In addition, for each task a single-factorial variance analysis was used for dependent random samples to test whether there was any difference between manual operation, voice operation and the control circuit without any in-vehicle systems. A paired comparison was carried out after the global test. For all effects described as "significant" $p < 0.05$.

(around 1.0 metre) for speech operation of the three systems fell within a range that would not be classified as significantly distracting.

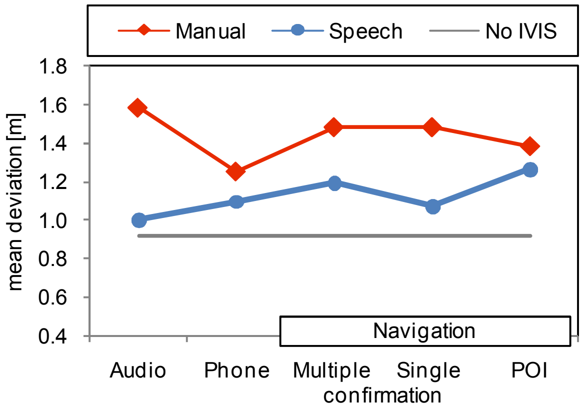

Illustration 4: Mean deviation over the entire circuit. For the five tasks (audio, phone, navigation with multiple and single address confirmation, navigation with POI), the value for manual operation is shown in red and voice operation in blue. The grey line shows the mean deviation for driving without any in-vehicle information systems.

This level of distraction during driving is determined mainly based on how often the driver needs to look away from the road, as shown in illustration 5. In the control journey without any in-vehicle information systems, the test subjects kept their eyes on the road the whole time (0% visual distraction). For manual operation, drivers looked away from the road for between 30 and 40% of driving time. Voice operation dramatically reduced the duration of visual distraction. For audio, drivers hardly needed to divert their gaze from the road. For phone and single confirmation address entry, just very short gazes away from the road were required.

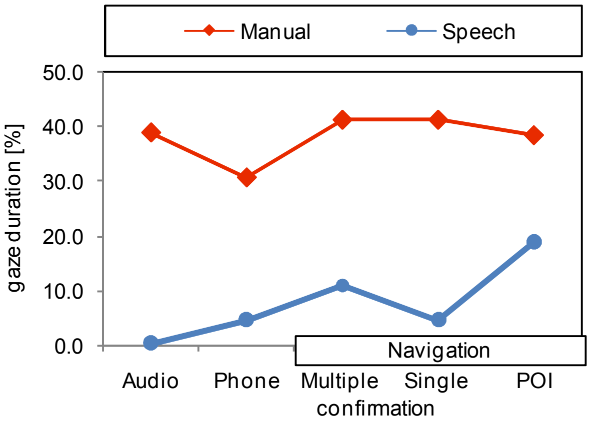

Illustration 5: Percentage of time the driver's gaze is diverted from the road over the entire circuit for the different conditions. The control condition is not entered since in this case test drivers kept their eyes on the road the entire time (0%).

An examination of the number and mean duration of gazes supports this interpretation (see illustration 6). During a circuit of around three minutes, manual operation required between 60 and 70 gazes at the secondary task. Drivers looked away from the road for an average of 0.7 to 0.9 seconds. Speech operation reduced both the number and mean duration, giving drivers a considerably higher chance of taking in relevant information from their environment.

According to illustration 6, the low gaze duration for audio and the lower percentage of deviations for phone and single confirmation address entry are primarily due to the fact that fewer gazes are required, clearly explained by the different control methods. In other words,

- the more the dialogue can be completely controlled by speech and the less visual information is required, the greater the increase in safety achieved through voice control.

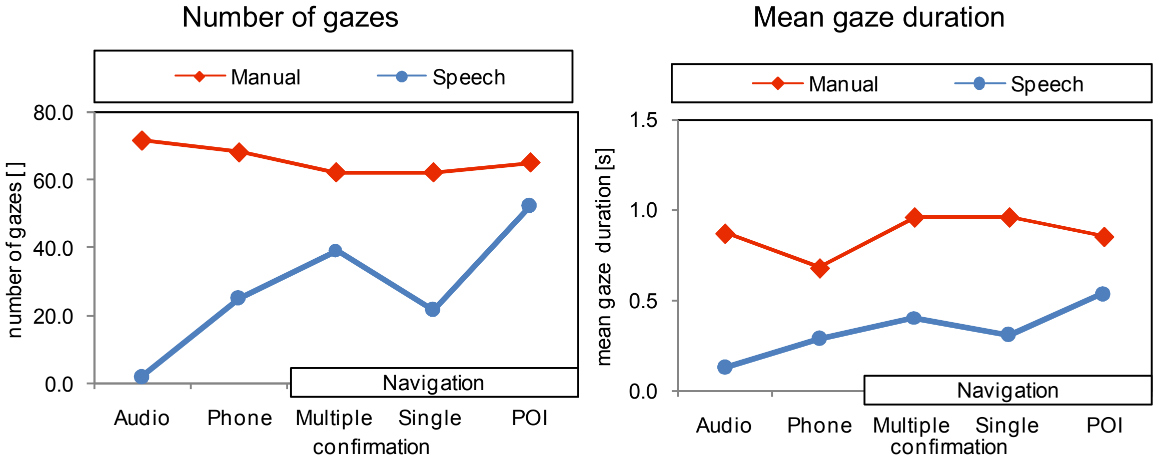

Illustration 6: Number (left) and mean duration (right) of gazes at the different tasks.

Drivers were also very clearly aware of these distractions, as shown in illustration 7, left. For voice operation, they assessed driving performance as "good" (10-12) or "average" (7-9). For manual operation, values tended to be within the "bad" range (4-6). This trend was confirmed by the study manager's assessment, illustration 7, right, although it was obviously more difficult for this individual to judge performance from outside as the effects were less noticeable.

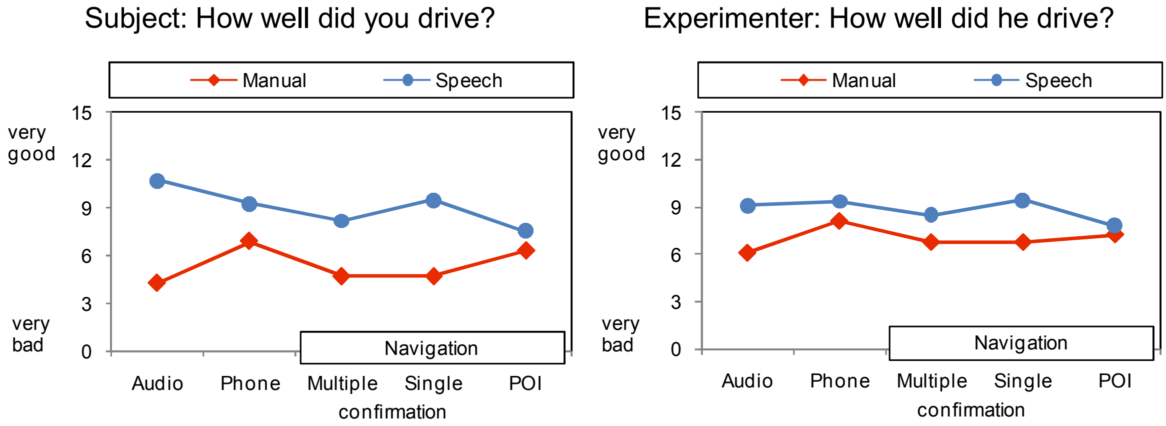

Illustration 7: Assessment of perceived driving performance ("How well did you drive?") by the driver (left) and the study manager (right).

Illustration 8, left, clearly shows that manual operation was more stressful for the driver, with this effect being particularly apparent for music selection. For speech operation, this effort was rated as "low" to "medium" compared with "high" for manual operation. This is mainly the result of feelings of distraction, as shown in the right-hand section of the illustration. Manual operation was also rated subjectively as very distracting. For voice operation, this fell to "medium" or "low".

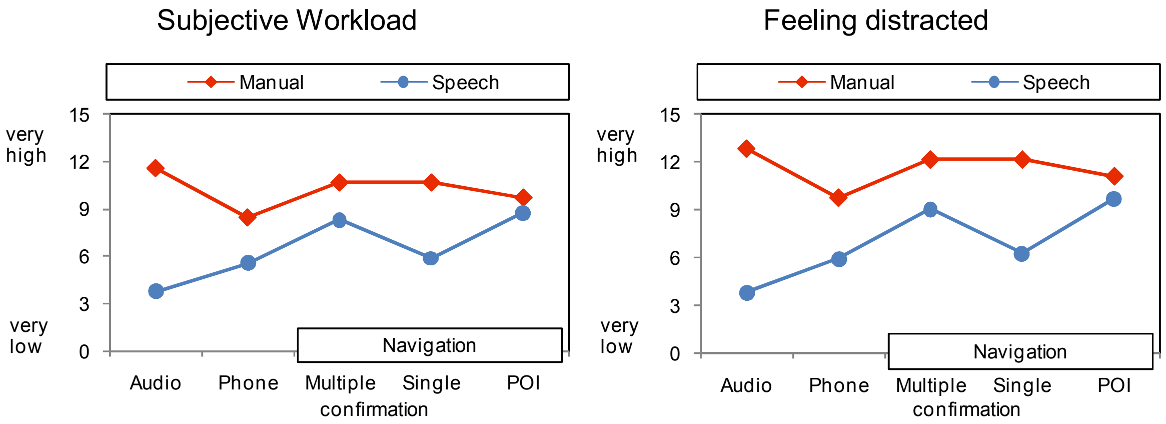

Illustration 8: Assessment of the workload and feelings of distraction for the different conditions.

Summarizing these results, some very clear effects are identified:

- For all devices and tasks, manual operation resulted in a significant fall in driving performance.
- This effect depended on the number of gazes required: the more often a driver needs to look at a display and the longer the gaze duration, the poorer the driving performance.
- Voice operation dramatically reduced gaze duration, resulting in a performance that was in some cases comparable to driving without any in-vehicle information systems.
- The more speech operation allows visual displays to be eliminated, and rely solely on the speech dialogue, the less the distraction effect.
- From the driver's point of view, manual operation requires more effort and is more distracting than voice operation. Accordingly, drivers also rated their driving performance as poorer.
- The benefits of speech operation are clear from the driver's point of view.

This allows us to draw some clear conclusions:

- In-vehicle information systems should be controlled by voice.
- Voice operation should be optimised to eliminate as far as possible the need to refer to visual displays.
- This objective increase in safety was also felt subjectively by drivers and assessed positively, producing a clear customer benefit.

3.2. Additional assessments of driving behaviour

3.2.1. During lane keeping

Illustration 9 shows the mean deviation and the standard deviation of the lateral position during lane keeping phases. A very similar picture appears in both cases. For all tasks there was very little difference between voice operation and driving without in-vehicle information systems for the standard deviation of lateral position. The standard deviation of lateral position is used in the literature as a measurement of lane keeping quality. Manual operation resulted in a very significant decrease in lane keeping quality, in particular for navigation (address entry and POI).

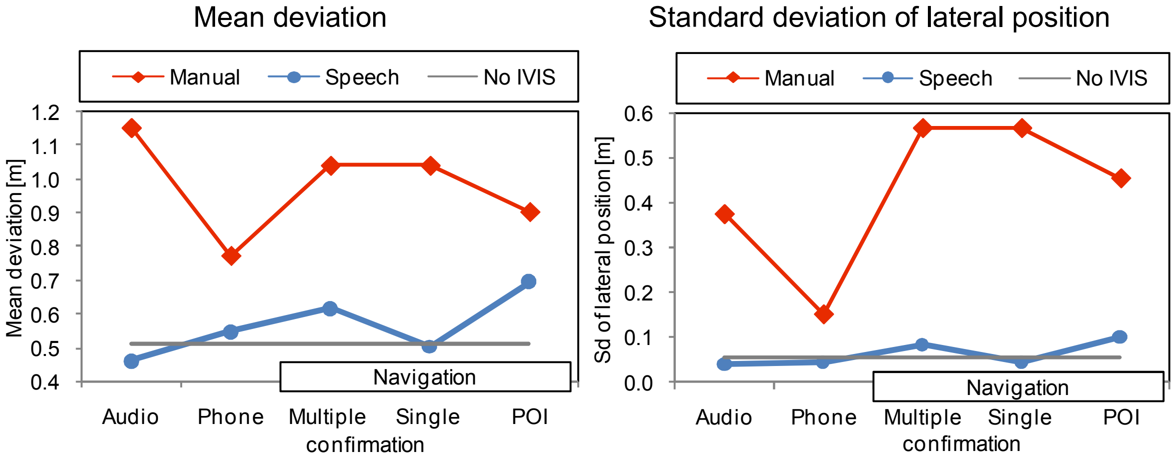

Illustration 9: Mean deviation and standard deviation of lateral position in lane keeping phases for the different tasks.

This lane keeping phase involved a very simple driving situation: on a straight road the only requirement for the driver is to stay in lane. In this case, automatic, intuitive control is sufficient to ensure good performance. However, visual information about your position on the road is required to exercise this control. Voice operation supports this and brings about the good performance shown here.

3.2.2. During lane changes

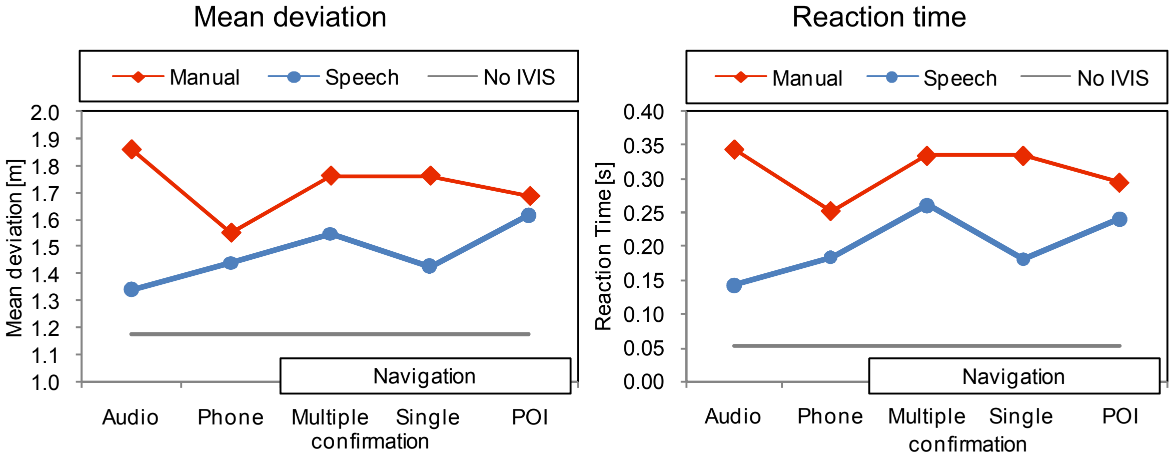

Illustration 10: Mean deviation and reaction time in lane change phases for the different tasks.

Lane changes place additional demands on the driver. He or she must recognise the signs and plan the appropriate action (steer right or left; cross one or two lanes). For this more difficult task, the distraction caused by speech operation was greater. Although less visual effort was required from the driver, the mental effort and concentration on operating the system meant that reactions were slower (see illustration 10, right) and the mean deviation was higher. Overall, voice operation shows in a significant advantage in the mean deviation and reaction time. However, the difference is not as great as in the simpler driving task.

3.2.3. Simple and difficult driving tasks - conclusion

In particular, voice operation reduces visual effort for the driver. In simple driving situations, where the driver automatically corrects his or her position on the road, voice operation brings in a clear advantage in terms of safety. Voice operation also reduces mental effort, i.e. the mental involvement required by the system. However, this effect is less strong, as shown by the lane change assessments. Menu design is particularly important for more difficult driving tasks, as encountered frequently in real traffic. Voice operation is essential to minimise these effects, but this is not enough: the concentration required by a menu requesting frequent driver confirmations involves the driver in system operation to an extent that affects driving performance. In other words,

- Voice operation is important in more complex traffic situations to relieve the strain on the driver.
- In complex traffic situations, the speech dialogue must be designed to minimise driver distraction.

4. Discussion

The study very clearly shows that manual operation of in-vehicle information systems results in a significant decrease in driving performance, even in relatively simple driving situations (lane keeping and lane change). This is mainly a result of the need for the driver to look away from the road. Drivers are aware of this effect, leading to increased strain which cannot be adequately compensated.

Voice operation can relieve the driver and help him or her to keep their eyes on the road, and therefore cope more effectively with driving situations. Voice-operated systems that do not use speech throughout the entire process and require the driver to check certain information visually (navigation with multiple confirmation and POI) are problematic. The main requirement for the design of the HMI of in-vehicle information systems is therefore:

- Voice operation of in-vehicle information systems is desirable from the point of view of safety and acceptance.

In difficult driving situations (lane change) even the best voice-operated system in this study (audio) showed a decrease in performance compared to driving without operating any in-vehicle systems. Obviously concentrating on a system can still be distracting even if the driver can keep his or her eyes on the road the whole time. In my view, this represents the limits of system design. However, this level of distraction is comparable to chatting with a vehicle passenger or being deep in thought. In such cases we should not look to make changes in the design of the information systems, but consider other possibilities (e.g. driver assistance systems).