A Study on the Design of Voice Navigation of Car Navigation System

Chih-Fu Wu, Wan-Fu Huang, and Tung-Chen Wu

Graduate School of Industrial Design, Tatung University
40 JhongShan North Road, 3rd Section Taipei 10452, Taiwan
huangsm@tsh.ttu.edu.tw

Abstract. This study tries to find the designing blind spots of the voice prompt function in the current car navigation systems and make improvement suggestions. The experimental plan was implemented through videotape analysis of the voice-prompt mode, referring to Urban Road Classification Regulations and the questionnaire survey results. Driving simulation tests were conducted with 15 experimental subjects, 13 road combinations, and 3 running speeds, and different prompt modes which were run synchronously were also included. Compared with the present mode (prompt time is determined by distance.), the newly-designed mode (prompt time is determined by running speed.) significantly improved driving performance and reduced mental workload. When driving on a main artery with fast lanes and slow lanes, adding a changing-lane prompt with a clear sound to the system can help increasing the driving accuracy rate.

Keywords: navigation systems, voice prompt function, driving accuracy rate.

1 Introduction

Following the development of science and technology, car navigation systems have increasingly been used by a lot of drivers. Car navigation systems provides drivers with information about how to get from one place to another in a turn-by-turn format such as distance to the turn, the name of the street to turn onto and turn direction [1]. However, whether a car navigation system is pre-installed on the dashboard or setup on the windshield of the driver, the driver inevitably needs to move the eyesight from the road ahead to a 3.5 to 8-inch LCD display of the automobile navigation system. Such distraction from the road is one of the main causes of traffic danger [2,3]. Nowadays, car navigation systems provide information through not only monitor displays but also voice prompt messages, so as to reduce the time drivers spend on monitors and reduce the danger of driving. Moreover, certain drawbacks remain in voice prompt function of car navigation systems, which prevent drivers from using them as the only way to receive navigation information. However, noises on the road, chattering voices of the passengers or sounds made by other vehicles, as well as music, radio and other factors may all interfere with navigation information. If voice playbacks are overlapping, it may confuse drivers and cause accidents. Unclear

J.A. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2009, LNCS 5612, pp. 141–150, 2009.

navigation information may cause accidents, especially when drivers need to switch lanes from one multi-lane road section to another, or to enter special terrains such as tunnels or overpasses. Due to the limitation of information processing, attention needs to be allocated when a person is multitasking, and his/her mental workload is increased as a result.

Attention resources of drivers can be categorized into visual resources, operative resources, mental-workload resources, acoustic resources and information ranking for analytical purposes [4]. Drivers depend largely on the visual modality for driving-related information [5]. While mental workload varies along with changes in acoustics and operation, several studies claim that if a message is instant, simple and to be reacted immediately after being received, vocal prompt is suggested; if a message is complicated, lengthy and not to be reacted immediately after being received, visual prompt is preferable [6]. Some studies suggest that names of road are not necessary to be provided by voice prompt in navigation system, because these are not easily comprehended by drivers immediately [7].

Given that driving is metal-requisite, the timing of information to be provided should be limited: if the information is provided too early, drivers may forget such information; if the information is provided too late, drivers may not have enough time to execute relevant tasks. Although mobile navigation systems have been increasingly popular, issues with regard to navigation have appeared as well. Foreign or domestic studies in this area, however, are rather scarce. This study, as a result, aims first at pointing out the issues regarding comprehension of displayed information and voice prompt information from currently available car navigation systems through questionnaires to drivers. Further, with driving simulation tests, driving patterns of drivers passing intersections of different kinds of roads are recorded and compared with the results of questionnaires. Finally, through the combination of variables such as monitor displays, driving reminders and timing of voice prompt, further designs and verification are conducted for the study to be applied in the design of voice prompt systems.

2 A Survey of Currently Car Navigation Systems

The study started from investigating products on the market. Three types of car navigation system from four large manufacturers were selected, and installed to proper positions above the wind outlet of air-conditioner in a real car. After setting the route, the car was driven on the same way in Taipei city and the playing mode was recorded in video taping for three runs. From the content of the video, the syntax of driving reminders and broadcasting timing with matching road classification principle were analyzed for designing questionnaire. Questionnaires are used in the early analysis so as to understand issues regarding the provision of information to drivers from currently available automobile navigation systems. They are described below.

- Sample Setting: Three most representational car navigation systems (Mio CT720, PaPaGo R12, TomTom ONE) were selected from the products available in the market. Two CCDs were used to simultaneously record the status on the road ahead and the graphic displayed and voice played by the navigation systems. In total, three rounds of tests were conducted. Grammar analysis is employed based on the abovementioned samples. After the analysis, Mio-Tech's product is the most complicated one, with driving information provided before a decision point in a pattern of a intersection in front + b please drive close to X + c please turn + d enter Y + e leave Y + f drive Z, in which "a" and "f" are broadcasted before every intersection, "b" is broadcasted when driving straight is required, "c" is broadcasted when a turn needs to be made at an intersection, "d" and "e" mean entering or leaving special roads, such as overpasses or tunnels, are required, "X" means turning left or right, "Y" means special terrains, such as overpasses and tunnels, and "Z" means the name of the name to drive on. The information provided by PaPaGo is less complicated than that by Mio-Tech but in a similar pattern. The difference is that it does not provide names of the roads to drive on (namely, without the item, f drive Z). TomTom uses the simplest way to broadcast, which only presents one simple sentence such as "PLEASE TURN LEFT" or "PLEASE DRIVE STRAIGHT".

- Categorization of and Regulation on Downtown Roads: Based on the analysis of voice prompt functions of the abovementioned mobile navigation systems, navigation scripts do not adapt in accordance with different navigation contents for different roads. Though additional contents are made for special roads such as overpasses and tunnels, they are too limited. In order to fully understand the issues regarding the provision of driving information with mobile navigation systems, roads need to be fully understood. In accordance with Urban Road Classification Regulations relevant regulations in Taiwan [8], according to the property of the road service such as the speed limit, traffic lane number, lane width, traffic control devices, etc. condition, downtown roads may be categorized into expressway, major road, sub road and service road. Combination of roads and questionnaires are used to understand issues regarding comprehensibility of voice prompt of mobile navigation systems.
- A survey of questionnaires interview: The information from the navigation system was presented on paper in the questionnaires with voice prompts being played. As for the graphic information, 13 representative graphic images were captured during the actual uses of the current navigation systems. Through cluster analysis with Urban Road Classification Regulations, with road names changed to prevent the effect from subjects' memories, they were then re-made into the testing images for the questionnaires with CoreDraw. There are 12 questions in 3 categories including awareness, UI preference, and habit. Only those who had used a navigation system before need to answer the habit-related questions. The investigation was carried out for forty-three subjects (20-55 years old, *M*=28.4) who owned driving licenses and had driving experiences. Fifteen of forty-three subjects had driven an automobile with a car navigation system.

From the above analysis, we can summarize the finding as follows: (1) Products of Mio and PaPaGo have better comprehensibility compare to that of TomTom for those who haven't used car navigation systems to get more detail the drive information. (2)On the roads with fast lanes and slow lanes, starts at about 20~30 meters from intersections, there might be double solid white line as "keep in lane" lane marking to divide the car flows toward the same direction for the sake of safety. If a right-turn needs to be made at this kind of intersections, drivers must switch to the slow lane on the outside before passing where the double solid white line starts. (3) The voice prompts in the current products are usually played 500 meters before critical intersections. The same content is played four times. The last time begins about 20~30 meters before intersections are also present graphically in Fig. 1. But drivers usually care only about the last prompt and ignore the first three. Therefore there would be actions like braking, slowing down, etc. right after the last prompt.

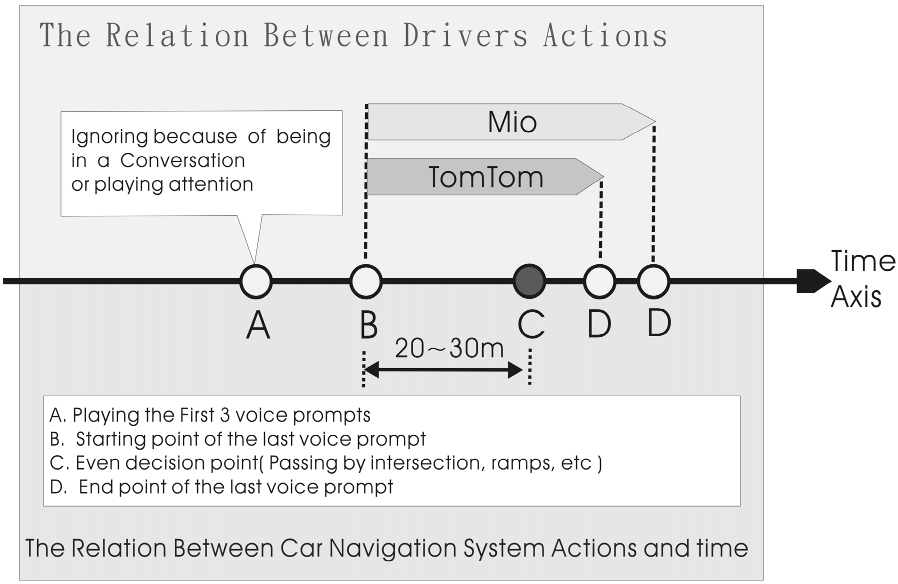

Fig. 1. Some voice prompt models of the currently available products (on expressway)

3 Driving Simulation Test

The key point in the design of human-machine interface is to discuss the interaction relationship among practical users, objects and environment. As to this study, the real experience in car driving was the best in the sense of reality. However, due to lots of factors interfering the road driving and experimental variables not easy to control, simulation driving was carried out to assure the consistency of test conditions in the experiment. Virtual reality scenes were constructed by thirteen combinations of turns in the road that created from the roads classification and regulation on downtown roads.

3.1 Participants

A total of 15 graduate students of Tatung University at the age of 20 or more (20-46 years, M=26, SD=3.2) and were paid to participate in the experiment. All participants have valid driving licenses and driving experience. All participants have normal or corrected visual acuity.

3.2 Apparatus

In the present study, the simulated vehicle cab, a 1990 FORD TELSTAR car, included all the normal automotive displays and controls found in an automatic vehicle. The simulator used two Pentium IV PCs to control the simulator, scenario respectively and one Notebook to show visual in-car navigation system. The scenario graphics were projected onto a projector screen located about 2 m in front of the driver's seat to produce 60 *40 degree field of view. The steering, throttle and brake inputs were connected to the PC that controlled the simulator software. The car navigation system is set up on in right-front of the driver, the height and location of which is around the ventilator of the air-conditioning system as the visual display. For a driver of normal height, the in-vehicle display was 15degrees below his/her straight-ahead plane and 20 degrees to the right. A speaker in front of the passenger seat provided auditory information in the form of a digitized human female voice with a speech rate of ~150 words/min and sound effects.

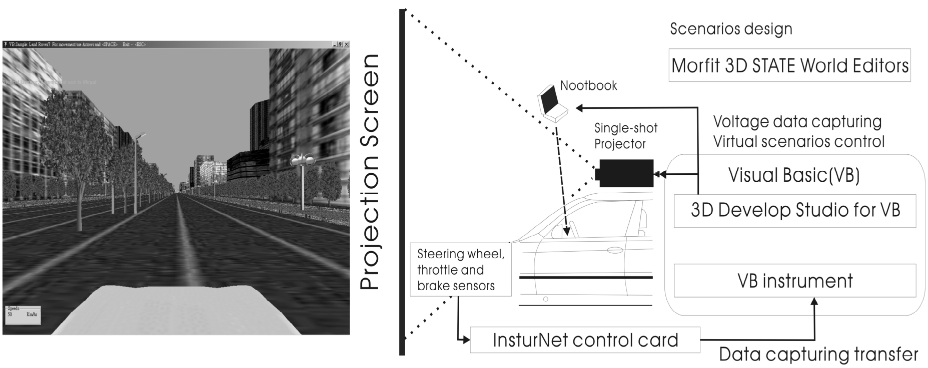

Fig. 2. One scene of driving simulation & Structure of virtualization control software/hardware

GW InstruNet Data Retrieving System is used for transformation and retrieval of data. Data is retrieved and input into Excel and SPSS for further determination. For analyses of the broadcasting timing of navigation systems, speakers are connected with the GW Instruments data retrieving box at Vin+ and Vin-, and hence voltage values vary as sounds are played. 3DSTATE World Editors developed by Morfit was used to generate the virtual environment. Scene data can be recalled by using Visual Basic together with the animation engine of 3D Developer Studio for Visual. Structure of Virtualization control software/hardware is depicted in Fig. 2.

3.3 Experimental Designs

Two different tests were assessed. One was driving simulation tests of currently available car navigation system with modifying switch-lane reminder (hereafter referred to as Test1), and the other was driving simulation tests for newly-design evaluation car navigation system (hereafter referred to as Test2).

Test1 was a 3 *2*2*2 mixed-factors model that compared results by **speed on the driven road** (three levels: 70 km/hr in expressway , 50 km/hr in major road, as well as 20 km/hr in service road), **way of voice prompt** (two levels: complicated broadcasting patterns: Mio-Tech, simple broadcasting patterns: TomTom), **monitor display**(two levels: on, off), **reminders of switching lanes** (two levels: on, off).When the last factor is set to be "off", the mode being tested is the currently available voice-prompt mode; when it's set to be "on", the mode being tested is a new trial. The way it works is that a clear and short sound will be played before a voice prompt to remind divers paying attention to the voice content which is about to be played. The effect is like that of the broadcasts to look for somebody in the railroad stations.

The evaluation on the newly-designed mode (prompt time is determined by running speed instead of prompt time is determined by distance.) was performed in Test2. In this voice-prompt mode, the reaction distance should be [speed (km/h)*2.5second+30] meters. The number of tests was reduced because of the experience from Test1. Only the factors "speed on the driven road" was kept the same as planned. The numbers of levels were reduced to 1 for the two factors "monitor display (one level: on)" and "way of voice prompt (one level: simple broadcasting patterns: TomTom)". But the number of levels for the factor "reminders of switching lanes" was increased to 3 (three levels: display, sound, voice). The display mode shows a red arrow icon as a guide to switching lanes (to the left or right lane) on the navigation screen. The sound mode is the same as the "on" mode in Test1. The voice mode has voice prompts for switching lanes (to the left or right lane).

Dependent variables were based on both objective and subjective measures. Whether it would be successful or not while turning into scheduled road would be regarded as the object evaluation criterion. Subjective measures were obtained using a modified five-point scale NASA TLX workload assessment.

3.4 Procedure

First, a ten-minute explanation is provided with regard to the test purpose and procedure and to instruct the subjects on how to manipulate in the driving simulation environments with the wheel, accelerator and brake pedal (those were connected to the serial port of the personal computers). Then the subject entered the test phase where he will receive the 36 navigation scenarios in pre-planned different sequences but were counterbalanced to prevent any learning/order effect. A short break was taken if necessary after each scenario trial. For each navigation test, all subjects were required to drive through the designated routes to reach an identical destination. All participants were told to arrive the destination in accordance with the driving information provided by available mobile navigation models, to drive along the route as accurate as possible, and to maintain the car moving in a controllable speed. Chart of changes in voltage are drawn with Excel based on data retrieved, which are analyzed to determine whether subjects have driven in accordance with the rules of the test. Finally, subject filled out a NASA TLX subjective evaluation questionnaire to describe his/her subjective feelings toward different experimental conditions. It took ~120 min for participants to complete the present study.

4 Results

Data collect for present study were analyzed by inferential statistics of ANOVAs. The LSD treatment contrast tests were used for *post-hoc* comparison.

Table 1. Test1 Resuls -1 (Currently Available Products)

Driving Speeds	Broadcasting Patterns	Driving Correctness(%)			
		Monitor Display On		Monitor Display Off	
70km/h Expressway	Mio	13.3%	26.7%	0.0%	0%
	TomTom	40.0%		0.0%	
50km/h Major Road	Mio	93.3%	96.7%	26.7%	36.7%
	TomTom	100%		46.7%	
20km/h Service Road	Mio	100%	100%	93.3%	90%
	TomTom	100%		86.7%	

Table 1 shows that only 26.7% of drivers can correctly leave the expressway by only listening to the final navigation information 20~30 meters before the event decision point. On major roads, 96.7% of drivers can make turns correctly with the monitor on, and 36.7% with it off. On service roads, whether monitors are on or off, the correctness rates are more than 90%. As TomTom's navigation contents are shorter and require less time for broadcasting within the same distance, they tend to be fully listened to by drivers compared with those of Mio-Tech's.

Table 2. Test1 Resuls -2 (Switch-Lane Reminding with sound effect)

Driving Speeds	Switch-Lane Reminding Sound	Driving Correctness	Chi-Square Inspection		
			Value	DOF	Sig.
70km/h Expressway	On	78.3%	1.534	1	0.215
	Off	68.3%			
50km/h Major Road	On	83.3%	9.076	1	0.003**
	Off	58.3%			
20km/h Service Road	On	80.0%	0.519	1	0.472
	Off	85.0%			

From Table 2, it shows that when driving on major roads with median strips, driving correctness rate is 83.3% with switch-lane reminders and 58.3% without such, while "p" is less than 0.05 under the chi-square test. Table 2 shows descriptive statistics and a chi-square inspection sheet of driving with different broadcasting patterns. There is neither significant difference between these two broadcasting patterns, nor between whether the monitor display is on or not.

Two major disadvantage of currently available car navigation systems based on the conclusion from early analysis: (1) The timing of final voice prompt on function cannot be changed based on the car speed. As a result, when the driving speed is rather high, navigation information may not be fully broadcasted, and driving action may not be made in time by drivers. (2) It would result in the same problem if drivers don't fully understand the switch-lane prompt from the navigation system when driving. But when using reminding sounds on some sections of the roads, as they may be misunderstood as "PLEASE MARK A TURN", when no switch-lane reminder is provided, driving correctness is higher, nevertheless.

Table 3. Test2 Results (Switch-Lane Reminding with Different Ways)

Driving Speeds	Switch-Lane Reminding	Driving Correctness	Chi-Square Inspection Value	DOF	Sig.
70km/h Expressway	Sound	93.3%	5.850	2	0.05*
	Voice	100%			
	Display	73.3%			
50km/h Major Road	Sound	86.7%	10.833	2	0.004**
	Voice	100%			
	Display	53.3%			
20km/h Service Road	Sound	73.3%	2.218	2	0.345
	Voice	80.0%			
	Display	93.3%			

Table 3 shows descriptive statistic data of correctness on three kinds of tested roads. Correctness rates on the three kinds of roads are generally the same, while different driving reminders contribute to changes in correctness rate. Based on the significance derived with chi-square tests, on expressways and major roads, three different kinds of reminders affect driving correctness significantly, while there is no significant difference on service roads.

Subjective measure Subjective grading items include quality of information provision, suitability of broadcasting timing, necessity of monitor display and comprehensibility of information. Through single-variable analyses of variance, it shows that broadcasting patterns designed by this study are better than those of the currently available products and can reduce drivers' mental workload.

5 Discussion and Conclusions

During driving and navigating, drivers have to monitor the car by searching the environmental information and shifting attention from one information source to another [10,11], and drivers depend largely on the visual modality for driving-related information [5]. According to the multiple resource theory, in a heavily loaded visual display environment, an auditory display will improve time-sharing performance [12]. When driving conditions and information are complicated, drivers may have more difficulty in filtering and remembering useful information presented by an auditory display because of the memory interference problem [13]. Similar results were found in the present study.

Although voice information aims at reducing the occasions in which drivers move their eyesight from the road, deficiencies in voice navigation functions may nevertheless become a safety concern. However, drivers may have difficulty in paying attention to the auditory display all the time. The timing of the appearance of prompt messages is a keypoint. Four of these findings are worth summarizing:

- Prompting Issues of Voice Navigation Systems: Insufficient prompted information and unsuitable timing of prompting result in misunderstanding of voice navigation information. Unclear information, such as not reminding drivers to switch lanes, leads to driving actions not being accomplished when event decision points, such as turning and switching lanes, are passed. In terms of prompting timing, drivers usually neglect the first three prompted contents before the decision, while the final information is not prompted based on driving speeds and may not be heard by drivers in time.
- Switch-Lane Reminder : With regard to reminding sounds for switching lanes: (1) On major roads, such sounds can effectively improve driving correctness ($p<0.05$), with voice reminders being the best one (100%), sound reminders the second (93.3%) and display reminders the third (73.3%); (2) On express ways, although there is no significant difference, driving correctness is still improved (from 68.3% to 78.3%), with voice reminders being the best one (100%), sound reminders the second (86.7%) and display reminders the third (53.3%); (3) On service roads, there is no significant difference, with display reminders being the best one (93.3%), voice reminders the second (80.0%) and sound reminders the third (73.3%).
- Timing of prompt: According to the voice prompt pattern in the tests conducted, the distance before event decision point should be [speed × 2.5 seconds] meters for the broadcasting to be finished. According to the study [9,14], 2.5 seconds may be further reduced to 1.93 seconds due to reaction time in acoustic transmission is shorter than that in visual one. As switching lines is prohibited within 30 meters from intersections in downtown areas, for major roads and expressways, the reaction distance should be [speed × 2.5 seconds + 30] meters.
- Further Suggestions: A proper amount of navigation information and its optimization have not been discussed in this study. In terms of tests conducted, interfering events and other vehicles are not simulated for the driving environment.

Acknowledgements

This research was supported by the National Science Council under contract number NSC 97-2221-E-036-032-MY2, Taiwan(R.O.C).

References

1. Perez, W.A., Mast, T.M.: Human factors and advanced traveller information systems (ATIS). In: Proceedings of the Human Factors Society 36th Annual Meeting, pp. 1073–1077. Human Factors Society, Santa Monica (1992)
2. Wierwille, W.W.: Development of an initial model relating driving in-vehicle visual demands to accident rate. In: Third Annual Mid-Atlantic Human Factors Conference Proceedings. Virginia Polytechnic Institute and State University, Blacksburg (1995)

3. French, R.L.: In-vehicle navigation-status and safety impacts. Technical Papers from ITE's 1990,1989, and 1988 conference (1990)
4. Wierwille, W.W., Hulse, M.C., Fischer, T.I., Dingus, T.A.: Visual adaptation of the driver to high-demand driving situation while navigating with an in-car navigation system. In: Vision in Vehicles III, pp. 79–87. Elsevier Press, Amsterdam (1991)
5. Lansdown, T.C.: Visual allocation and the availability of driver information. In: Gothengater, T., Carbonell, E. (eds.) Traffic & Transport Psychology: Theory and Application, pp. 215–223. Pergamon Press, Amsterdam (1997)
6. Sanders, M.S., McCormick, E.J.: Human Factors In Engineering and Design, 7th edn. McGraw Hill Press, Singapore (1993)
7. Rasker, P.C., Post, W.M., Schraagen, J.M.C.: Effects of two types of intra-team feedback on developing a shared mental model in command & control teams. Ergonomics 43, 1167–1189 (2000)
8. Construction and planning Adency Ministry of the Interior(CPAMI) Information, `http://www.cpami.gov.tw/web/index.php`
9. McGehee, D.V., Mazzae, E.N., Scoot, G.H.: Baldin: Driver Reaction Time in Crash Avoidance Research:Validation of a Driving Simulator Study on a Test Track. In: Proceedings of the IEA 2000/HFES 2000 Conference, vol. 3 (2002)
10. Dewar, R.E.: In-vehicle information and driveroverload. International Journal of Vehicle Design 9, 557–564 (1988)
11. Dingus, T.A., Hulse, M.C.: Some human factors design issues and recommendations for automobile navigation information systems. TransportationResearch 1, 119–131 (1993)
12. Wickens, C.D., Sandry, D., Vidulich, M.: Compatibility and resource competition between modalities of input, central processing, and output. Human Factors 25, 227–248 (1983)
13. Liu, Y.-C.: Comparative study of the eVects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. Ergonomics 44(4), 425–442 (2001)
14. Oglesby, C.H., Hicks, R.G.: Highway Engineering. John Wiley & Sons, Chichester (1982)