The Use of Modality in In-vehicle Information Presentation: A Brief Overview

Yujia Cao
Human Media Interaction
University of Twente
y.cao@utwente.nl

Mariët Theune
Human Media Interaction
University of Twente
m.theune@utwente.nl

ABSTRACT

In-vehicle information systems (IVIS) are multimodal presentation systems that are designed to assist drivers. The challenge of IVIS information presentation is to deliver information effectively while minimizing the interference with driving. Modality is a presentation factor that has been known to influence the performance of IVIS. In this paper, we attempt to provide an overview on the advantages and disadvantages of three types of modalities (visual, auditory and tactile) when used in the IVIS context, as well as the benefit and cost of combining multiple modalities. Findings presented in this work can be used as guidelines when making design choices in a specified system context.

ACM Classification Keywords

H.5.2 Information interfaces and presentation: User Interfaces, User-centered design

INTRODUCTION

In-vehicle information systems (IVIS) are intended to provide drivers with information such as route instructions, traffic conditions, hazard warnings, vehicle monitoring warnings and so on. IVIS are a subtype of multimodal information presentation systems designed for driver assistance. The driving context brings specific challenges to in-vehicle information presentation. In a moving vehicle, the primary task of the driver is to watch the road and control the vehicle. The perception and comprehension of IVIS information[1] can be considered as a secondary task that needs to be performed concurrently with the driving task. On the one hand, IVIS information is meant to assist driving and improve safety, thus needs to be effectively delivered to the driver. On the other hand, it imposes attentional cognitive demand and may distract the driver from driving. Since distraction has been identified as a major cause of car accidents [24], IVIS information could potentially have a negative influence on safety as well. In sum, the challenge of IVIS information presentation is to provide information effectively while minimizing the additional cognitive demand on, and distraction of, the driver. This paper addresses the design of IVIS messages from a modality perspective. Research findings regarding visual, auditory and tactile modalities are briefly summarized, with a focus on the former two. Note that the intention here is not to point out one generally-valid best choice of modality for IVIS, but to provide an overview on the advantages and disadvantages of various modalities when used in the IVIS context. This knowledge is relevant when making design choices for a system in a specified context.

[1]The expression 'IVIS information' or 'IVIS message' refers to information or messages presented by IVIS to drivers.

MIAA 2010, February 7, 2010, Hong Kong, China.

VISUAL MODALITIES

Driving is a highly visual task, requiring almost continuous visual attention while the vehicle is in motion [26]. This causes a drawback for visual IVIS presentations – they impose extra load on the visual perception channel and usually "drag" the eyes away from the road (except when the information is presented on the windshield). According to the multiple resource model from Wickens [47], two visual tasks can not be performed in parallel due to the perceptional resource competition in the visual channel. This means drivers need to switch their visual attention between the outside driving environment and the in-car display. Another drawback of visual presentation is a lack of salience. The onset of a visual presentation might be overlooked if the driver happens to be looking somewhere else. That is to say, the information delivery is likely to be delayed or even unattended. The advantage of visual presentations is their self-paced feature [37]. They allow the driver to inspect them at his/her own pace, e.g. at once, step by step or selectively. They can also be read multiple times, which is particularly beneficial in case certain details need to be kept in mind for later recall.

AUDITORY MODALITIES

When used in a driving environment, auditory modalities have the major advantage of consuming a separate perceptional resource from driving. Based on the multiple resource model [47], perceiving auditory IVIS information can be time-shared with driving, which means drivers can watch the traffic and listen to the message at the same time. Being omnidirectional is another advantage of auditory modalities. They can be picked up from all directions, independent from where the driver is facing. However, auditory modalities (especially speech) have been found to have a 'preemption effect' on driving [48], which means they can pull attention away and temporally suppress the primary driving task. The

reason is twofold. First, attention is promptly directed to an auditory signal upon the onset of its presentation [39], and this intrinsic alerting characteristic can easily grab attention from an on-going visual task, such as driving. Second, auditory information is transient and force-paced. To capture the full message (for speech in particular), drivers need to continuously attend to it during its presentation, which might impair the processing of driving-related visual inputs.

VISUAL VS. SPEECH PRESENTATION

The choice between presenting information visually or orally has been well investigated in IVIS studies. Theoretically, there is an apparent conflict between the resource competition view and the attention preemption view. The former addresses a perceptional aspect and supports a choice for speech for a more efficient use of perceptional resources. The latter, however, addresses an attentional aspect and supports a choice for visual presentation for less impairment of driving. Empirically, both auditory benefits and visual benefits have been found by a number of studies.

The auditory benefits were commonly demonstrated with a navigation-assisted driving scenario, in which drivers followed the instructions from a navigation device [4, 11, 25, 28, 40, 46]. These studies showed that when navigation instructions (e.g. turn notification) were presented aurally compared to visually, drivers reacted faster, made fewer errors and showed better driving performance in terms of speed and steering control. In addition, auditory benefits were also found in a couple of studies that used other types of secondary tasks, such as a warning detection task [33], a letter detection task [18] and an information searching task [34]. These secondary tasks interfered less with driving when relevant information was provided aurally compared to visually.

The auditory preemption effect (or visual benefits) were also found by a number of studies [17, 29, 32]. Concurrently with driving, drivers in [29] were asked to listen to/look at statements and reason whether they were true or false; drivers in [17] had to listen to/look at phone numbers of different length and recall them; drivers in [32] were asked to listen to/look at road sign information and react if necessary. In these studies, auditory presentations interfered with driving more than their visual alternatives, causing greater variations in lane position, speed control and headway distance. However, concerning the secondary tasks, auditory presentations often led to faster reactions and better performance. In contrast, visual modalities showed an advantage of perceptional flexibility. Drivers could choose to attend to the visual display at a suitable/safe moment, or take multiple steps to read a message and return to driving in between.

Neither theories nor empirical findings revealed a winner between the two types of modalities. In fact, this is a high-dimensional choice, which means many factors play a role and the final choice cannot be made without evaluating these factors in a specified design context. After all, the choice should be made to let the potential benefit outweigh the potential damage. Based on theories and empirical findings, the following factors were identified to guide this selection:

- **The relevance to driving (or the priority) of the message** [21, 37]
  When the IVIS information is relevant to driving, it usually has a high priority and requires a timely perception. Speech presentation is preferred in this case, because it is beneficial to have attention preempted to information that pertains to the driving task and is intended to support the driving task. Visual presentations lack salience, thus critical messages are likely to be overlooked. In contrast, when the information is not driving-related (e.g. a weather forecast), it has a low priority in the driving context. Visual presentation is more suitable in this case, because the driver needs to be able to temporarily ignore it if he/she has to concentrate on driving at that moment.
- **The spatial location of the visual display** [37]
  In case of visual presentation, drivers must divide their focal visual attention between the driving environment (outside the windshield) and the in-car display. The larger the distance between display and windshield, the greater the (cognitive) cost of dividing one's attention between the two. Therefore, the advantage of visual modalities can be more pronounced when head-up displays are used than when head-down displays are used.
- **The length of speech** [5, 38]
  Due to their transient characteristics, when speech messages get longer and/or more complex, they keep drivers' attention longer and impose a higher load on working memory. This can in turn enhance the auditory preemption effect. Besides, the full meaning of speech may not become clear until the end of the message. This makes long speech inappropriate for urgent warning messages that demand an immediate response. In short, speech has stronger advantages over visual presentations when it can be kept short and precise.
- **The mental workload** [18, 25, 37, 40]
  An increase in mental workload, due to either a more demanding driving task or a more complex secondary task, may increase the benefit of auditory presentation in relation to visual presentation. When driving is more demanding, the division of visual attention between the road and the in-car display becomes potentially more dangerous. When the secondary task is relatively complex, drivers simply have to spend more time looking at the in-car display in case of a visual presentation. Speech is superior in both cases, because of its "eyes-free" feature.
- **The environment condition** [8, 31]
  Properties of the environment in- and outside the car at the moment of presentation can also affect the utility of a certain modality. For example, speech messages might be less effective if the driver is at the same time talking to a passenger, getting a phone call, listening to the radio or playing loud music. Visual messages might be less effective on a sunny day when strong light makes it hard to see what is on the display.
- **The information type** [2, 41]
  This factor addresses the expressive power of modalities, because one modality can be naturally better than another in presenting a certain type of information. For example, speech is better at presenting instructions, commands and abstract information (e.g. logic, relations). Short speech

is good for warnings and alarms. Regarding visual modalities, text is suitable for quantitative values (e.g. distance, speed, road numbers); icons are effective to indicate physical objects (e.g. gas stations along the highway) and directions (e.g. left/right turns); maps are good for locations and spatial information.

AUDITORY ICONS

Auditory icons refer to familiar environmental sounds that imitate real-world events (originally defined in [10]). They inherit common characteristics of auditory modalities, such as the high salience. Besides, they also have their own unique features that are beneficial to IVIS information presentation, and are worth mentioning in separation. In IVIS, auditory icons are typically used as warning signals. For example, presenting a car horn sound or a screeching car tire sound can warn drivers of an impending collision [12, 14]. First, auditory icons are language independent and culture independent. Second, when well chosen (the more intuitive, the better [42]), they inherently covey the meaning of the events that they are meant to signify [22, 38]. In other words, the meaning of the events is immediately clear to the driver. This feature explains the common empirical findings that drivers reacted significantly faster to auditory icon warnings than to speech warnings [1, 12]. However, the use of auditory icons also has limitations [38]. There is evidence that the fast reactions may be accompanied by an increase in inappropriate responses [3, 12]. This is because drivers may react before they have properly evaluated the situation to know what the most appropriate response would be. Moreover, auditory icons are likely to be considered unpleasant, due to inappropriate loudness or high pitch [30].

TACTILE MODALITIES

Compared to visual and auditory modalities, the use of tactile output is rather new in IVIS information presentation. However, existing findings have already shown positive promises. Two tactile modalities have been investigated so far: force pulse (given by pedals [19] or steering wheel [43]) and vibration (given by seat [27, 45], steering wheel [23, 43] or additional equipments attached to the driver [14, 16]). Regarding information type, tactile modalities have been typically used as alerts and directional cues. For example, they have been applied to warn drivers of a rapidly approaching vehicle [15, 16], a sudden deceleration of the lead vehicle [13, 14, 27], or a lane departure [9, 43]. In most cases, more than one tactors were applied at different locations in order to add another dimension to the presentation – the direction drivers should look in for the event. For example, two tactors were used in [16], one at the front side and one at the back side of the driver's torso. Only one tactor was activated at a time, indicating a vehicle was approaching from either the front or the back. In addition, the 'left versus right' tactile cues have also been applied to indicate turning directions [45, 44] and lane change directions [23].

A commonly obtained finding is that tactile signals induced significantly faster reactions to the presented events in comparison with either the absence of tactile signals [13, 14] or the auditory/visual presentation alternatives [9, 15, 43, 44, 45]. This is mostly because tactile signals are highly salient and can almost always draw attention immediately. Besides, tactile modalities also have other advantages: 1) they don't compete with driving for visual perceptional resources; 2) their effectiveness is not influenced by the lighting condition, driving noise, radio or conversation; 3) they are private to the driver and do not bother the passengers. Apart from various advantages, tactile modalities have two limitations. First, they are limited in their expressive power, meaning that they are only suitable for a few types of information, such as time-critical alerts and directions. Second, tactile stimuli are likely to induce annoyance and physical discomfort. To minimize this negative effect, the duration and intensity of the signal should be carefully chosen (several suggestions can be found in [20]).

MODALITY COMBINATIONS

As each single modality has its own advantages and disadvantages, a question that naturally arises is: can a combined use of multiple modality types bring advantages over using each single type alone? In general, multimodal information presentation has been widely applied in intelligent user interfaces. Commonly found benefits of multimodality include an enhanced robustness of communication due to redundant or complementary use of modalities, and an increased bandwidth of information transfer [35, 36].

The multimodality benefit has also been found in IVIS information presentation. A couple of studies investigated the combined use of visual and auditory modalities [6, 7, 28]. The presented IVIS information include navigational messages [28], vehicle monitoring messages [28], headway distances and local danger warnings [6]. Drivers were required to perform secondary tasks based on the IVIS messages. Results showed that the combination of modalities allowed better performance in both driving and secondary tasks, compared to using either single modality alone. The combination was also the most preferred modality variant by the drivers. Moreover, several other studies revealed that it could be particularly beneficial to combine tactile presentation with auditory presentation [14] or with visual presentation [23, 45]. These combinations were shown to induce faster reactions and impose lower cognitive load than using either single modality alone.

The risk of using multiple modalities is to induce additional costs in terms of perception load, interface management and monitoring demand [36]. Therefore, designers need to make sure that the combination enhances human cognitive and physical ability and is compatible with user preference, context and system functionality [35]. In other words, additional modalities should be added to the system only if they improve efficiency, satisfaction, or other aspects of performance for a given user and context.

CONCLUSIONS

The choice of modality should be carefully made for IVIS, because it can certainly influence the system's performance in diver assistance. To determine the optimum choice of modality for a specific presentation task, designers need to

make an overall optimization over a number of aspects, such as the driving demand, the environmental condition in- and outside the car, the priority of the message, the type of information to be conveyed, the type of task induced by the presentation, and the driver's physical and cognitive condition. This optimization process can be guided by the existing findings on the advantages or disadvantages of various modalities associated with various aspects, such as the ones presented in this overview. Ideally, choices of modality should be made dynamically during driving, so that they adapt to real-time changes in the driving environment.

REFERENCES

1. S. Belz, G. Robinson, and J. Casali. A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors*, 41(4):608–618, 1999.

2. N. Bernsen. Multimodality in language and speech systems - from theory to design support tool. *Multimodality in Language and Speech Systems*, pages 93–148, 2002.

3. J. Bliss and S. Acton. Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics*, 34(6):499–509, 2003.

4. G. Burnett and S. Joyner. An assessment of moving map and symbol-based route guidance systems. *Ergonomics and Safety of Intelligent Driver Interfaces*, pages 115–137, 1997.

5. Y. Cao, S. Castronovo, A. Mahr, and C. Müller. On timing and modality choice with local danger warnings for drivers. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 75–78, 2009.

6. Y. Cao, A. Mahr, S. Castronovo, M. Theune, C. Stahl, and C. Müller. Local danger warnings for drivers: The effect of modality and level of assistance on driver reaction. In *International Conference on Intelligent User Interfaces (IUI'10)*. ACM, 2010.

7. T. Dingus, D. McGehee, N. Manakkal, S. Jahns, C. Carney, and J. Hankey. Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human Factors*, 39(2):216–229, 1997.

8. J. Edworthy and E. Hellier. Complex nonverbal auditory signals and speech warnings. *Handbook of warnings*, pages 199–220, 2006.

9. M. El Jaafari, J.-F. Forzy, J. Navarro, F. Mars, and J.-M. Hoc. User acceptance and effectiveness of warning and motor priming assistance devices in car driving. In *Proceedings of the Humanist Conference*, pages 311–320, 2008.

10. W. Gaver. Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2(2):167–177, 1986.

11. K. Gish, L. Staplin, J. Stewart, and M. Perel. Sensory and cognitive factors affecting automotive head-up display effectiveness. *Journal of the Transportation Research Board*, 1694:10–19, 1999.

12. R. Graham. Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics*, 42(9):1233–1248, 1999.

13. C. Ho, N. Reed, and C. Spence. Assessing the effectiveness of intuitive vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis and Prevention*, 38(5):988–996, 2006.

14. C. Ho, N. Reed, and C. Spence. Multisensory in-car warning signals for collision avoidance. *Human Factors*, 49(6):1107–1114, 2007.

15. C. Ho, C. Spence, and H. Tan. Warning signals go multisensory. In *Proceedings of HCI International*, volume 9, pages 1–10, 2005.

16. C. Ho, H. Tan, and C. Spence. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Psychology and Behaviour*, 8(6):397–412, 2005.

17. W. Horrey and C. Wickens. Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors*, 46(4):611–624, 2004.

18. J. Hurwitz and D. Wheatley. Using driver performance measures to estimate workload. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, volume 46, pages 1804–1808. Human Factors and Ergonomics Society, 2002.

19. W. Janssen and H. Thomas. In-vehicle collision avoidance support under adverse visibility conditions. *Ergonomics and safety of intelligent driver interfaces*, pages 221–229, 1997.

20. L. Jones and N. Sarter. Tactile displays: Guidance for their design and application. *Human Factors*, 50(1):90–111, 2008.

21. C. Kaufmann, R. Risser, A. Geven, and R. Sefelin. Effects of simultaneous multi-modal warnings and traffic information on driver behaviour. In *Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems*, pages 33–42, 2008.

22. P. Keller and C. Stevens. Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, 10:3–12, 2004.

23. D. Kern, P. Marshall, E. Hornecker, Y. Rogers, and A. Schmidt. Enhancing navigation information with tactile output embedded into the steering wheel. In *Proceedings of the Seventh International Conference on Pervasive Computing*. Springer, 2009.

24. S. Klauer, T. Dingus, V. Neale, J. Sudweeks, and D. Ramsey. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic study data. Technical Report DOT-HS-810-594, US Department of Transportation: National Highway Traffic Safety Administration, 2006.

25. G. Labiale. In-car road information: Comparisons of auditory and visual presentations. In *Proceedings of the Human Factors and Ergonomics Society 34th Annual Meeting*, volume 34, pages 623–627. Human Factors and Ergonomics Society, 1990.

26. T. Lansdown. Visual allocation and the availability of driver information. *Traffic and Transport Psychology: Theory and Application*, pages 215–223, 1997.

27. J. Lee, J. Hoffman, and E. Hayes. Collision warning design to mitigate driver distraction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 65–72. ACM, 2004.

28. Y. Liu. Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, 44(4):425–442, 2001.

29. G. Matthews, T. Sparkes, and H. Bygrave. Attentional overload, stress, and simulated driving performance. *Human Performance*, 9(1):77–101, 1996.

30. J. McKeown and S. Isherwood. Mapping candidate within-wehicle auditory displays to their referents. *Human Factors*, 49:417–428, 2007.

31. N. Merat, A. Jamson, and U. Leeds. Multisensory signal detection: How does driving and ivis management affect performance? In *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessemnt, Training and Vehicle Design*, pages 351–357, 2007.

32. M. Mollenhauer, J. Lee, K. Cho, M. Hulse, and T. Dingus. The effects of sensory modality and information priority on in-vehicle signing and information systems. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, volume 38, pages 1072–1076. Human Factors and Ergonomics Society, 1994.

33. T. Poitschke, F. Laquai, and G. Rigoll. Guiding a driver's visual attention using graphical and auditory animations. In *Proceedings of the 8th International Conference on Engineering Psychology and Cognitive Ergonomics: Held as Part of HCI International 2009*, pages 424–433. Springer, 2009.

34. T. Ranney, J. Harbluk, and Y. Noy. Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors*, 47(2):439–454, 2005.

35. L. Reeves, J. Lai, J. Larson, S. Oviatt, T. Balaji, S. Buisine, P. Collings, P. Cohen, B. Kraal, J. Martin, et al. Guidelines for multimodal user interface design. *Communications of the ACM*, 47(1):5769, 2004.

36. N. Sarter. Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, 36(5):439–445, 2006.

37. B. Seppelt and C. Wickens. In-vehicle tasks: Effects of modality, driving relevance, and redundancy. Technical Report AHFD-03-16 & GM-03-2, Aviation Human Factors Division at University of Illinois & General Motors Corporation, 2003.

38. C. Spence and C. Ho. Multisensory warning signals for event perception and safe driving. *Theoretical Issues in Ergonomics Science*, 9(6):523–554, 2008.

39. C. Spence and D. J. Audiovisual links in attention: implications for interface design. *Engineering Psychology and Cognitive Ergonomics*, 2:185–192, 1997.

40. R. Srinivasan and P. Jovanis. Effect of selected in-vehicle route guidance systems on driver reaction times. *Human Factors*, 39(2):200–215, 1997.

41. K. Stanney, S. Samman, L. Reeves, K. Hale, W. Buff, C. Bowers, B. Goldiez, D. Nicholson, and S. Lackey. A paradigm shift in interactive computing: Deriving multimodal design principles from behavioral and neurological foundations. *International Journal of Human-Computer Interaction*, 17(2):229–257, 2004.

42. K. Stephan, S. Smith, R. Martin, S. Parker, and K. McAnally. Learning and retention of associations between auditory icons and denotative referents: Implications for the design of auditory warnings. *Human Factors*, 48(2):288–299, 2006.

43. K. Suzuki and H. Jansson. An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *JSAE review*, 24(1):65–70, 2003.

44. J. Van Erp and H. Van Veen. Vibro-tactile information presentation in automobiles. In *Proceedings of Eurohaptics*, volume 2001, pages 99–104, 2001.

45. J. Van Erp and H. Van Veen. Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Psychology and Behaviour*, 7:247–256, 2004.

46. J. Walker, E. Alicandri, C. Sedney, and K. Roberts. In-vehicle navigation devices: Effects on the safety of driver performance. In *Proceedings of Vehicle Navigation and Information Systems Conference*, volume 2, pages 499–525, 1991.

47. C. Wickens. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2):159–177, 2002.

48. C. Wickens, S. Dixon, and B. Seppelt. Auditory preemption versus multiple resources: Who wins in interruption management. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 49, pages 463–467. Human Factors and Ergonomics Society, 2005.