

NAMIDA: Multiparty Conversation Based Driving Agents in Futuristic Vehicle

Nihan Karatas^(✉), Soshi Yoshikawa, P. Ravindra S. De Silva,
and Michio Okada

Interaction and Communication Design Lab,
Toyohashi University of Technology, Toyohashi 441-8580, Japan
{karatas,yoshikawa}@icd.cs.tut.ac.jp, {ravi,okada}@tut.jp
<http://icd.cs.tut.ac.jp/en/profile.html>

Abstract. We propose socially interactive NAMIDA (Navigational Multiparty based Intelligent Driving Agents) as three friendly interfaces those sit on the dashboard inside a car. NAMIDA interfaces conduct multiparty conversation within each other to provide outside information for driver by utilising the context aware interaction. In this research, we introduce the conceptual design of NAMIDA and discuss about the superiorities of multiparty communication in the aspects of decreasing cognitive workload and increasing sociability plus enjoyable driving experience through an experiment which includes mock driving tasks by getting support from 1-NAMIDA (single) and 3-NAMIDAS (multiple) consecutively. Finally we mention about the results of the experiment which depicts that using multiparty based communication (3-NAMIDAS) could reduce the workload of driver and induced more enjoyable driving experience by increasing interaction and social bonding inside the car.

Keywords: Multiparty conversation · Context aware interaction · Conversational workload · NAMIDA

1 Introduction

While driving, a driver needs to achieve a variety of information; e.g., sightseeing places, best restaurants, the condition of roads, etc., also demands to spend enjoyable time of driving. Plenty of people prefer to use appealing mobile devices to attain useful applications to acquire mentioned information and enjoyment. Nevertheless, mobile devices are not designed to fulfil the drivers' demands inside a car, and the usage of those devices can divert driver's attention, consequently leads to accidents. Regarding to those issues, recently In Vehicle Infotainment (IVI) systems provide multitask opportunities to derive an efficient and entertaining driving as safe as possible [1]. Unfortunately common problems were identified in these systems as a workload of control menus, audio and visual feedback (looking at display), route guidance problems (always associated with designer's suggestions), reliability, lack of joy and intelligence and to give suggestions to driver. Navdy [2] is designed and developed as a transparent head up display in order to eradicate mobile phone interaction inside the car while

providing navigation facility. By using Navdy drivers can control their smart phone through hand gestures and speech without getting in touch with the mobile device. In addition to reduce the workload and ensuring safety, considering the driver's social and emotional state is also crucial in driving. In this respect, Nissan has been implementing an assistive robot in order to present more human-like approach rather than other IVI systems [3]. Furthermore MIT created a friendly assistant AIDA which provides personalisation speciality inside the car [4]. AIDA can decide the relevant information while driving and express it at the most appropriate time with a suitable facial expression. AIDA communicates with the driver through speech coupled with expressive body movements.

All those systems use one to one communication (between a human and interface) which cannot reduce the workload sufficiently. Due to the fact that drivers still should be in an interaction with the above systems to obtain their demanded instructions while trying to focus on the road, having distraction is inevitable. It is possible to imagine a driving environment that requires less attention to be get informed and exposes social and emotional interaction with driver. At that point, applying multiparty conversation methods to an innovative driving agent interfaces in order to convey the information about outside of the car (e.g., sightseeing places, best restaurants, the condition of roads, etc.,) and providing sociability through this method can be considered.

Through multiparty conversation, the user doesn't have to participate into the conversation; instead of that user can be notified by the discussion among the other participants. Besides, whenever users want they can join into the conversation and can acquire additional information. Nakagawa [5] clarifies the advantages of multiparty conversation as (1) the conversation becomes more lively, (2) various interactive controls become possible (3) it is possible to expect the range of new applications of spoken dialog systems to widen. Furthermore, Mawari [6] has designed as an interactive social medium to boardcast information (e.g. news, etc.,). Mawari interface consists of three robots which conduct multiparty conversation to diminish the workload on the user.

In respect of reducing the workload and providing social and enjoyable environment inside a car we propose NAMIDA. NAMIDA has three intelligent social interfaces those fix on the dashboard of a futuristic vehicle and conduct multiparty conversation with respect to driver's suggestions by getting advantages of context aware interaction (e.g. suggested places in close around) facilities while providing enjoyable driving experience via establishing social bonding between the robots and the driver.

2 Concept of NAMIDA

NAMIDA interfaces conduct multiparty conversation such that the driver can obtain outside information (e.g., sightseeing places, best restaurants, etc.,) without participating in to the conversation. (Fig. 1) The multiparty conversation is fed by the outside information, which is the essence of the context-aware interaction to enlighten the driver about vicinity. Mentioned multiparty conversation includes asking questions, giving consistent answers and having discussions

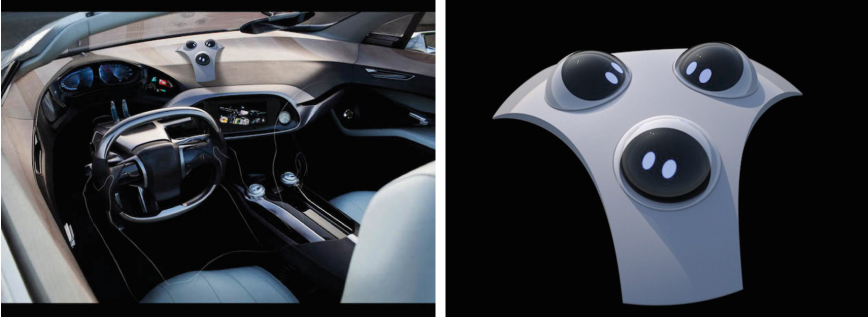


Fig. 1. The appearance of NAMIDA inside the car (*left*) and closer appearance of NAMIDA (*right*).

among the NAMIDA interfaces. While the driver having credible information from NAMIDA, he/she can make better decisions and starts to feel trust to NAMIDA. The social and emotional state based on trustfulness between two parties (driver and NAMIDA) brings more enjoyable and sociable environment inside the car. This pleasing engagement appears as a social bonding.

3 Design

We adapted to the minimal design mechanism as shown in Fig. 2. As an initial step of NAMIDA interfaces, we followed to implement an animation. The round-shape display of NAMIDA is for some facial expressions. Each NAMIDA has one degree of freedom for moving their head to right and left. All NAMIDAS sit on a common base which fits on the dashboard of the car. We used three different discernible colours (red, green, blue) for each NAMIDA's eyes and recognisable distinct roles in the conversation ((1)wise (2)ignorant (3)cooperative) with varied voices to be implied that each NAMIDA has a different character. When one of them starts to talk, the other two turn their heads toward to speaker to expose the talking NAMIDA. Moreover, we used Eye Tribe in order to track the driver's eye gaze to become aware of his/her attention into road empirically.

Each NAMIDA utters its own lines according to the prepared script for a designed route. We used the Wizard Voice (ATR-Promotions) as a voice synthesis engine and Unity [13] to develop a driving simulation which consists a route of suggestible spots for driver. While getting closer to the spots, NAMIDAS start to perform multiparty conversation to give information about the place (context-aware interaction).

4 Interactive Architecture

4.1 Multiparty Conversation

One of the advantages of a multiparty conversation approach is having different personalities for each individuals to possesses their different kind of knowledge [7]. Mutlu and his colleagues [8] has explored a conversation structure,



Fig. 2. Minimal and futuristic design of multiparty based NAMIDA.

the participants' roles and the methods of shifting the roles during a multiparty-conversation. Also Goffman [9] introduced the concept of "footing" that explains the participants' roles in a conversation. Also it describes the concept of shifting roles in understanding the social interaction. It is possible to define four main roles of participants in a multiparty conversation which are speaker, address, side participants, and bystanders [10, 11]. The side participant's role is waiting for the conversation to participate. However the bystander doesn't contribute to the conversation at the moment. These roles rotate/change during a conversation.

The proposed interface of NAMIDA is based on the above criteria in order to reduce the conversation workload due to shifting roles of participants during a conversation. The driver doesn't need to give an answer or respond to the conversation. Instead, another participant takes over the responsibility to continue the conversation. However there are times when the driver has to participate in the conversation. As long as the driver's role remains as a bystander, the NAMIDA interfaces interact with each other considering to the context aware interaction feature in a productive way so that, the driver can make a decision by listening the conversation. When the driver participates to the conversation then the roles are changed in to the side participant, speaker and addressee. The utterance generation of NAMIDA collaborates with symbolic display for eyes shape and basic body motions and the utterance generation architecture by utilizing the fillers, back-channel, turn-initial, etc., (Fig. 3).

4.2 Context-Aware Interaction

We designed a context-aware interaction system for NAMIDA. NAMIDA is capable of establishing interaction based on the location and it can capture those information within a certain km around from the car's current location and reveals it through a natural way in multiparty conversation. The utterance generation of NAMIDA based on modifying predetermined sentences with fillers,

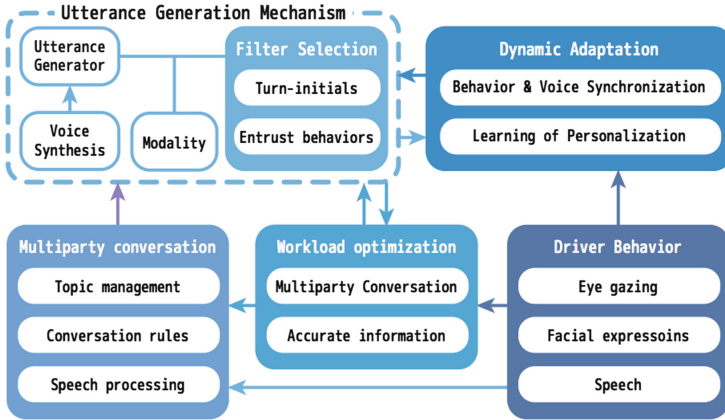


Fig. 3. Minimal and futuristic design of multiparty based NAMIDA.

back-channel, turn-initial, etc. to coordinate a productive conversation [12]. The content of the conversation (predetermined sentences) changes according to the surrounding locations. Each NAMIDA utters their roles (according to the prepared script) via a voice synthesizer to generate the conversation (in Japanese).

As an initial step, we developed a driving simulation that has several preferable destination spots to specify the locations statically for each driver. Through the context-aware interaction NAMIDAs acquaint those locations.

5 Experiment

Through our experiment, we intended to measure the workload, subjective impression (e.g. social bonding), effects of multiparty conversation, social interaction and attachment between the driver and the robots in two sessions (1-NAMIDA and 3-NAMIDA). In 1-NAMIDA interface, the participants listen to the one way conversation which consists of direct information. On the other hand, in 3-NAMIDAs interface, the participants listen to the multiparty conversation that involves asking questions, giving answers and having discussions. In both cases, the systems are using context-aware interaction. We divided our 14 participants (age range is in between 21 - 35; 3 female, 11 male) into half; while 7 participants had the experiment in the order of one NAMIDA case and three NAMIDA case, the other half had the experiment at first three NAMIDA case and then one NAMIDA case. Such a strategy is useful to acquire a counterbalance of the data, thereby reducing the effect of the sequence of trials on the results.

5.1 Experiment Set Up

In the experiments, each participant sit in a mock-in car environment and performed mock driving by watching the projected driving simulation on the wall (Fig. 4). The NAMIDA interface animations were displayed on a small screen

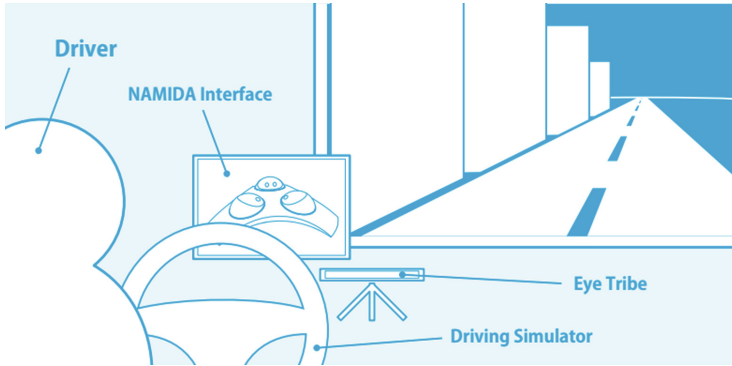


Fig. 4. The driver goes along the road while listening and understanding the contents of the nearby information through multiparty conversation.

that was placed left side of the dashboard. We arranged up two different sessions (one session for 1-NAMIDA interface, and one session for 3-NAMIDAS) for each participants. They charged to listen to the two distinct NAMIDA interfaces to wise up about the environment and remember the information to give right answers to the questions at the end of the each session. Each session took approximately 5 min. While in 1-NAMIDA session, the participants were charged to listen to the one NAMIDA interface in a one to one communication scenario; in 3-NAMIDAS session, the participants had to listen to the three NAMIDA interfaces in a multiparty conversation scenario. After each session the participants were required to evaluate 5 questions for workload and 6 questions for subjective impression which consists of social bonding. We used Driving Activity Load Index (DALI) as a subjective assessment tool in order to evaluate cognitive workload of the participants.

5.2 Driving Activity Load Index (DALI)

The DALI (Driving Activity Load Index) [14] is a revised version of the NASA-TLX [15] and adapted to the driving task. Mental workload is multidimensional and depends on the type of loading task. The basic principle of DALI is the same as the TLX. There is a scale rating procedure for six pre-defined factors, which are Effort of attention, Visual demand, Auditory demand, Temporal demand, Interference and Situational stress, (Table 1) followed by a weighting procedure in order to combine the six individual scales into a global score. However, in our study, we used five factors by excluding Interference factor, because this factor is most suitable when it is used in real driving environment.

5.3 Results

Both Workload and Interaction and Social Bonding questionnaires were scaled in a ranged of 1 - 5. For each question we applied pair-wise t-test to determine if the difference between the 1-NAMIDA and 3-NAMIDAS case is significant or not.

5.3.1 DALI Results: Table 1 shows that there is significant difference on Attention demand ($p=0.033<0.05$, significant) which shows that it is possible remember more content in the case of one to one communication (1-NAMIDA) due to its conveying method is directly. However, the significance in Visual ($p=0.009<0.01$, highly significant) and Auditory demands ($p=0.0449<0.05$, significant) reveal that driver needs to pay more attention to watch the road and listen to the conversation for comprehending the information in 1-NAMIDA case.

Accordingly, 3-NAMIDAs (multiparty conversation) requires less visual (watching the road) and auditory (listening to the conversation) effort to understand and remember the information (Fig. 5). The other dimensions of Temporal demand and Situation Stress have nonsignificant effect in either of the scenarios. The Global value of Dali yielded nonsignificant difference for 1-NAMIDA and 3-NAMIDA cases, because of the Temporal demand rated as higher in 3-NAMIDA case. The reason can be considered as the further conversation amount (more information) has been presented in 3-NAMIDA rather than 1-NAMIDA in the equal period of time (5 min).

Table 1. DALI factors and the mean differences for the DALI values in 1-NAMIDA and 3-NAMIDA. The endpoints for each factor is in rang 1 to 5 (1= Very Low, 2= Low, 3= Neutral, 4= High, 5= Very High)

Dimention	Endpoint	Question	Dali Values (Mean and Stand dev)		P-value
			1-NAMIDA	3-NAMIDA	
Attention demand	Low/High	How well do you remember the content of the conversations?	56.44 (8.55)	51 (28.34)	$p=0.033<0.05$ d.f.=13 significant
Visual demand	Low/High	Did you get tired of looking at road?	41.14 (18.72)	18.28 (9.75)	$p= 0.009 <0.05$ d.f.=13 significant
Auditory demand	Low/High	How much effort did you spend to understand the conversation?	34.28 (17.71)	17.5 (17.62)	$p= 0.0449 <0.05$ d.f.=13 significant
Temporal demand	Low/High	Did you feel time pressure to understand the conversation of the robot(s)?	19.00 (23.78)	29.81 (16.43)	$p=0.145>0.05$ d.f.=13 non-significant
Situation Stress	Low/High	Did you get stressed during driving?	26.66 (16.32)	25.71 (16.94)	$p=0.45992>0.05$ d.f.=13 non-significant
Global			35.50 (5.50)	28.46 (6.67)	$p=0.224>0.05$ d.f.=13 non-significant

5.3.2 Interaction and Social Bonding: The subjective impression questions which are based on interaction and social bonding (Table 2) indicates that there is significant differences on Q1, Q3 and Q6. The significance on Q1 implies that the multiparty conversation exhibits more human like approach rather than one to one communication. Also, the significance of Q3 in 3-NAMIDA case reveals to sense more animacy from multiparty conversation. Furthermore, the significance of Q6 depicts that the multiparty conversation exposes more natural way to convey the information.

On the other hand, the nonsignificant effect of Q2 can be interpret as the both system are adequate to convey the information. There is also nonsignificant difference for Q4 which means 1-NAMIDA and 3-NAMIDAS give almost the same feeling of being fellow(s) of the driver. Moreover the nonsignificant difference for Q5 indicates that both cases expose a high rated persuasiveness (Fig. 5) due to the well designed interface.

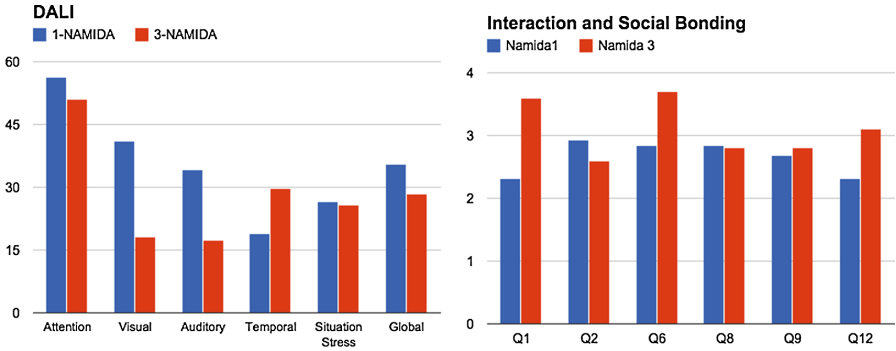


Fig. 5. Figure shows the result of Dali factors (*left*) and Interaction and Social Bonding (*right*) comparing 1-NAMIDA and 3-NAMIDAS

Table 2. Result of interaction and social bonding questionnaire indicates the significance state between 1-Namida and 3-Namida cases.

Code	Question	P-value	Result
Q1	Did you feel human likeness from the conversation?	$p=0.000828 < 0.05$ d.f.=13	significant
Q2	How often did you want to interact with the robot(s)?	$p=0.082372 > 0.05$ d.f.=13	non-significant
Q3	Do you feel that the robot(s) exhibited some animacy?	$p=0.000944 < 0.05$ d.f.=13	significant
Q4	Did you feel the robot(s) as friend(s)?	$p=0.241561 > 0.05$ d.f.=13	non-significant
Q5	Did you feel the robot(s) is/are persuasive?	$p=0.357635 > 0.05$ d.f.=13	non-significant
Q6	Did you feel the robot(s) conversation was spontaneous?	$p=0.003853 < 0.05$ d.f.=13	significant

5.4 Discussion and Conclusion

Overall, both cases (1-Namida and 3-Namida) there were nonsignificant differences in DALI factors of Temporal and Situation Stress demands. The reason is the both NAMIDA interfaces were capable of establishing a well-disposed social driving agents when they convey the information such that participants were not pressured to be concern about the timing and therefore experienced no fatigue or discouragement feeling. The significance in Attention demand reveals that one to one communication has better effect on remembering the content of the conversation. Yet, the highly significant difference on Virtual demand implies that driver exerts more visual effort to understand and remember the content. Consequently, we can deduce that 3-NAMIDAs can assist to avoid from visual distraction (lack of focus on road), because this approach requires less visual effort. Moreover, the significant difference on Auditory demand reveals that listening a multiparty conversation including asking questions, giving answers and discussions, requires less effort to understand the content.

According to the subjective impression questionnaire, 3-NAMIDAs interface presents the feeling of humanlike and spontaneous conversation more than the 1-NAMIDA does. This state infers that it is possible to create a social environment with multiparty based driving agents inside a car. The humanlike sense gives the feeling of the system (3-NAMIDAs) is far beyond of being just an instrument. The natural conversation manner of the multiparty conversation provides an enjoyable driving experience. In addition, the significantly highly rated animacy for 3-NAMIDAs implies that multiparty based interfaces expose more life-likeness which would be a core contribution to human-robot interaction research area. In our future work, we intend to implement a user-tracking mechanism in NAMIDA to use it in an interactive multiparty conversation to improve efficiency and sociability inside the vehicle.

Acknowledgement. Authors would like to express their gratitude to robotics designing group at Professor Michiteru Kitazaki and his Lab. This research has been supported by both Grant-in-Aid for scientific research of KIBAN-B (26280102) and Grant-in-Aid for scientific research for HOUGA (24650053) from the Japan Society for the Promotion of Science (JSPS).

References

1. Integrated Computer Solutions Inc, Qt-based Automotive Software Solutions (1999). <http://www.ics.com/services/ivi>
2. Navdy Inc., Feels Like Driving in Future (2013). <https://www.navdy.com/>
3. Nissan Motor Company Ltd., PIVO: An in-car robotic assistant, In Automotto (2005)
4. Williams, K., Breazeal, C.: Reducing driver task load and promoting sociability through an Affective Intelligent Driving Agent (AIDA). In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013, Part IV. LNCS, vol. 8120, pp. 619–626. Springer, Heidelberg (2013)

5. Todo, Y., Nishimura, R., Yamamoto, K., Nakagawa, S.: Development and evaluation of spoken dialog systems with one or two agents through two domains. In: Habernal, I. (ed.) TSD 2013. LNCS, vol. 8082, pp. 185–192. Springer, Heidelberg (2013)
6. Yoshiike, Y., De Silva, P.R.S., Okada, M.: MAWARI: a social interface to reduce the workload of the conversation. In: Mutlu, B., Bartneck, C., Ham, J., Evers, V., Kanda, T. (eds.) ICSR 2011. LNCS, vol. 7072, pp. 11–20. Springer, Heidelberg (2011)
7. Suzuki, N., Takeuchi, Y., Ishii, K., Okada, M.: Talking eye: autonomous creatures for augmented chatting. *Robot. Auton. Syst.* **31**, 171–184 (2000)
8. Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., Hagita, N.: Footing in human-robot conversations: how robots might shape participant roles using gaze cues. In: HRI 2009: Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, pp. 61–68. ACM, New York (2009)
9. Goffman, E.: Footing. *Semiotica* **25**, 1–29 (1979)
10. Rehmann, A.J.: Handbook of Human Performance Measures and Crew: Requirements for Flightdeck Research. Ohio, Springfield, Virginia, 22161 (1995)
11. Okada, M., Kurihara, S., Nakatsu, R.: Incremental elaboration in generating and interpreting spontaneous speech. In: Proceedings of 3rd International Conference on Spoken Language Processing, pp. 103–106 (1994)
12. Sakamoto, D., Hayashi, K., Kanda, T., Shiomi, M., Koizumi, S., Ishiguro, H., Ogasawara, T., Hagita, N.: Humanoid robots as a broadcasting communication medium in open public spaces. *I. J. Soc. Robot.* **1**(2), 157–169 (2009)
13. Unity - PC, Mac and Linux desktop game development (2013). <http://unity3d.com/>
14. Pausie, A.: A method to assess the driver mental workload: the driving activity load index (DALI). *Intel. Transp. Syst. IET* **2**(4), 315–322 (2008)
15. Hart, S.G., Stavenland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) *Human Mental Workload*, ch. 7, pp. 139–183. Elsevier, Amsterdam (1988)