The Proposal of Cognitive Support for Driver by Voice Guide Using Soliloquy Expression

Takuya Yamawaki[1(✉)], Takayoshi Kitamura[1], Tomoko Izumi[2], and Yoshio Nakatani[1]

[1] Ritsumeikan University, Kusatsu, Siga, Japan
is024lsh@ed.ritsumei.ac.jp
[2] Osaka Institute of Technology, Hirakata, Osaka, Japan

Abstract. In a car navigation system, a voice guidance system is equipped to ensure the safety of driving. On the other hand, a driver requires a certain period of time to understanding instructions from the voice guidance and may misjudge them. To solve these problems, we propose a new expression (SVN) process of the voice guidance based on drivers' soliloquy. When drivers confirm a point to turn or a distance to the point based on the voice guidance, they quite often make correspondence the expressions of the guidance to their own expressions. Such the expression in the brain of a driver is expected to be similar to the soliloquy type expression of the driver. We assume that the soliloquy expression of the guidance will be understandable easier than the conventional expression. We conducted the experiment on the driving simulator to verify our hypothesis. The results suggest that SVN decreases the time for understanding the instruction and the frequency of misjudgment.

Keywords: Car navigation system · Human-machine-interface
Personal cognitive support · Voice navigation

1 Introduction

Today, the major car navigation system has adapted a display monitor to guide the route. However, it is pointed out that the gaze to the display during driving degrades the stability of the driving operation and prevents the driver from detecting a dangerous object [1]. Therefore, it is necessary to study a new route guidance process not using the display. One of the new route guidance processes is that using voice guidance only. This new processology is considered to solve the safety problem because drivers need not to gaze the display. However, using only voice guidance will cause a new problem. Freundshuh et al. [2] pointed out that the problem was caused by the high abstraction degree of voice information. That is, the disadvantage of the voice guidance is its low amount of information and its accuracy. As a result, a driver may take a time to understand the instruction and identify a different object, such as a corner or a landmark, as the target one in the instruction. Therefore, it is necessary to verify the expression process of voice guidance assuming the use only of voice guidance.

2   Related Studies

Personalization is the concept of providing information according to the individual's ability, characteristics. Otani [3] indicated that personalized route guidance is effective in improving understandability, because individual differences exist in the environmental knowledge and the spatial cognitive ability. Therefore, by applying the concept of personalization to the voice guidance expression, it is possible to reduce the understanding time and improve the accuracy of decision. Audio information can be flexibly changed from the various points of views and have various expressions [4]. Therefore, it is considered that the voice guidance will be personalized in a relatively easy way. Kawai et al. [5] evaluated the personalized voice guidance based on a verbal navigation by participants. This evaluation revealed that there is a difference in preferred voice guidance expression depending on age in terms of the frequency of referring names of landmarks. However, in [5], they consider the situation where drivers use both of navigation from in-vehicle display and voice guidance. In the existing studies about the personalized voice guidance, the display guidance is also used, and there few studies focusing on using voice guidance only.

3   Proposal of New Voice Guidance Expression Process

In this research, we propose a new expression of voice guidance, called Soliloquy Voice Navigation (SVN), which uses soliloquy of a driver. We consider the situation where the navigation adopts the voice guidance only. In our work, we defined the "soliloquy of a driver during driving" as the monologue of the driver for confirmation by him/herself of a point of turning right or left. For example, "I wanna turn to the left a little further ahead…". The soliloquy during driving is caused when the divers confirms the routes or the target intersection to turn. Hence, it is considered that SVN, using the soliloquy is easy to understand for the driver.

4   Evaluation

4.1   Overviews

Our goal is to verify the hypothesis that the use of SVN is more effective than that of a general voice guidance (VN). To evaluate the efficiency, we focus on the time required for understanding the instruction content, called understanding time, and the accuracy of identifying the instruction, called the accuracy of decision. That is, we evaluate the understanding time and the accuracy of decision for SVN and VN in the experiment. To evaluate the efficiency of SVN, we compare the following evaluation indexes with VN:

1. The understanding time: It is the length of time from output of voice guidance to understanding of the instruction by a driver.
2. The accuracy of decision: It is the number of correctly decisions of the corner that the voice guidance instructed.

Furthermore, to reveal the suitable type of driver to SVN, we used the Driving Style Questionnaire (DSQ) which was proposed by Ishibashi et al. [3] and Workload Sensitivity Questionnaire (WSQ) which was proposed by Akamatsu et al. [4]. In this study, we do not focus on a process for construction of SVN. In the experiment, we made suitable SVN for each participant simply.

4.2   Participants

We focus on inexperienced drivers as participants. It is because that experienced drivers are accustomed to VN. Therefore, VN is a familiar voice guidance for experienced drivers, and then, such the experience will give some impact to evaluation results. In this research, the inexperienced driver is defined as a driver who driving frequency is less than once a month.

The details of the participants are shown in Table 1. The participants were selected from the university students and the graduate students majoring in the information science. All the participants were Japanese. We explained the contents in the informed consent, and all of the participants confirmed them in the writing. The previous research has reported that there are differences in cognitive ability during driving depending on gender [6]. On the other hand, other research pointed out that the differences between men and women is caused by the differences in their driving frequency [7]. In this experiment, since all the participants are inexperienced drivers, the gender of the participant does not affect the evaluation results. Figures 1 and 2 show the average scores about the results for DSQ and WSQ by the participants. The general drivers in these figures mean the average scores of general drivers in Japan obtained by the research of Ishibashi et al. [8] and Akamatsu et al. [9]. Compared with the general average, the average scores of the participants have the following features:

- The scores for "Anxiety about traffic accident" and "Hesitation for driving" are high.
- The scores for "Confidence in driving skill" is low.

It is considered that these results are caused by lacking driving experience of the participants. In addition, there is few difference between the participants and the general drivers excepting the points above. Therefore, in terms of the types of driving, it is said that there is few difference between the participants and the general driver.

Table 1. The details of the participants.

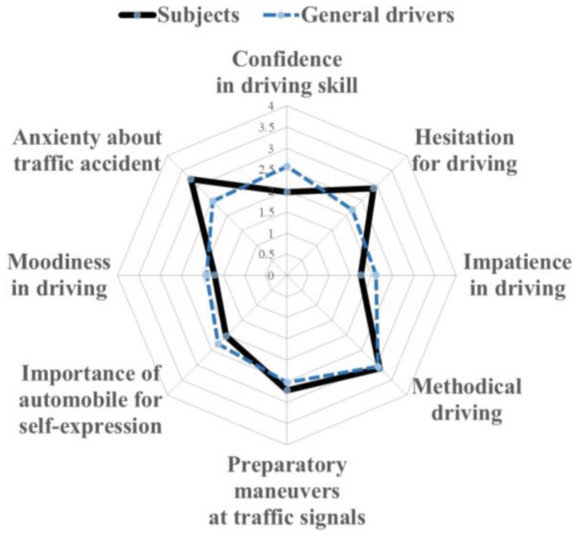| Item	Values
Number of participants	14
Gender	Male:10, Female:4
Age	19–25 years (average:21.7, SD:1.27)
Period of getting driving license	1–4 years
Driving frequency	Less than once a month

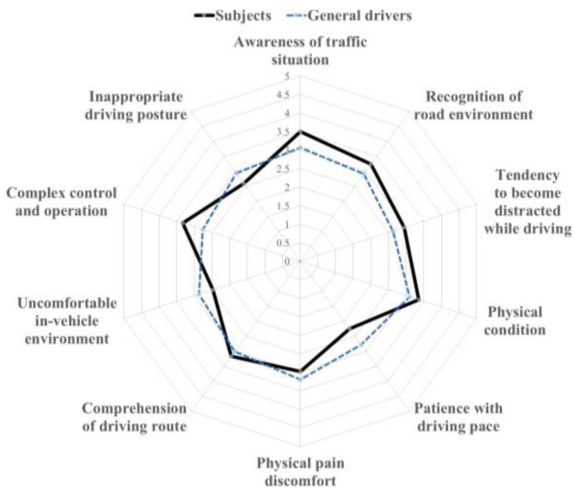**Fig. 1.** The average scores about the results for driving style.

Fig. 2. The average scores about the results for workload sensitivity.

4.3   Procedure

Creating Soliloquy Voice Navigation

In the experiment, we create suitable SVN for each participant using the soliloquy of the In the experiment, we create suitable SVN for each participant using the soliloquy of the participant. This process is called the SVN create process hereinafter. This process uses three videos, called SVN create videos. Figure 3 shows the outline of the
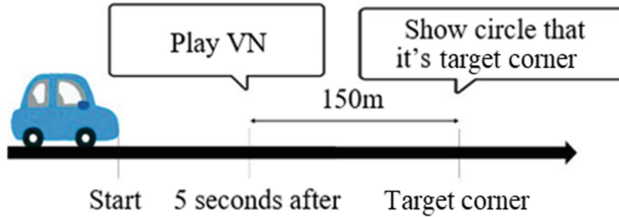
Fig. 3. The outline of SVN create video.

Fig. 4. The SVN create video capture. The red circle points out the target corner instructed by VN. (Color figure online)

SVN create video, 150 m version. At first, the VN starts at 5 s after the video starts. Next, the corner instructed by VN is pointed out by a red circle. Figure 4 shows one of the SVN create video capture, in which the corner is pointed out by the circle. In each video, the different contents of VN for the different driving roads are recorded. The SVN create videos were recorded the actual driving scenes from the driver's viewpoints in Nishi-oji-dori and Kita-oji-dori in Kyoto, Japan. We used GoProHERO4 camera to record. This camera was fixed to the installation position where does not cover the driver's operation. Regarding the parts of the recorded videos that were considered to have visibility problems due to the influence of sunlight, we adjusted the brightness so that it was in a general range during driving. The SVN create video was edited by adding VN at a point such that there was a corner at a distance 150 m, 250 m, or 325 m away from that point. The contents of three types of VN using in SVN create video are showed in Table 2. Each VN contents were created based on the voice guidance expression in the general car navigation systems in Japan.

Next, we explain the procedure of the SVN create process. The participants watch the three SVN create videos one by one. After watching one of the videos, we ask the soliloquy expression of the participants. More concretely, the procedure is as follows:

1. The participant watches the three SVN create videos.
2. The participant watches the SVN create video (150 m) three times.

Table 2. The contents of 3 types of VN using in SVN create video. The upper parts are the contents in Japanese, and the lower parts are the translated one in English.

Distance to a corner	Contents
150 m	"Mamonaku hidari houkou desu."
	(Please turn to the left soon.)
250 m	"Konosaki hidari houkou desu."
	(Please turn left in a little while.)
325 m	"Oyoso 300 m saki hidari houkou desu."
	(Please turn left after 300 meters away.)

3. We ask a question "How do you talk to yourself when the voice guidance was played?" and record the answer.
4. The participant watches the SVN create video (250 m) three times.
5. We ask the same question and record the answer.
6. The participant watches the SVN create video (325 m) three times.
7. We ask the same question and record the answer.
8. We generate the voices expressing the three answers above by the synthesis.

In this process, three types of SVN that is personalized to each participant are generated by using the SVN create videos.

At first, the participant watches all of the three SVN create videos. It is because that the participant is made understand the sense of distance to the corner instructed by the VN. Second, the participant watches SVN create video 150 m version. After that, we ask a question to get a soliloquy expression of the participant for the case of the preceding video.

The "talk to yourself" used for the question is synonymous with "soliloquy of the participant during driving", which is explained to the participants before the experiment. We record the answer to the question. The same process in the cases of the SVN create video (250 m) and the video (325 m) is performed. An example of the generated three types of soliloquy is showed in Table 3.

Table 3. The examples of the contents of SVN. The upper parts are the contents in Japanese, and the lower parts are the translated one in English.

Distance to corner	Contents
150 m	"Mousugu hidari houkou yana"
	(OK, turn left.)
250 m	"Kekkou saki susunndekara hidari
	(I'll turn left in a little bit.)
325 m	"Daibu sakimade susunndekara hidari"
	(After a time, I'll turn left.)

For video play, we used a system created using PsychoPy [10], which is an application for a psychological experiment environment. In addition, we used Rospeex API [11] to generate the synthetic voice. Moreover, It is considered that if the

participants watch the video on the desktop display, the sense of distance will be changed from the feel of them. To solve this problem, we used Oculus Rift CV 1, which is a Head Mounted Display (HMD), for display the SVN create video.

Evaluation Experiment

We compared the efficiency of the SVN and the VN in terms of the understanding time and the accuracy of decision. We call this process the evaluation experiment. In the evaluation experiment, we used a system constructed using PhychoPy. The system outputs the videos, called the evaluation experiment videos. This video is created based on the actual driving scenes from the driver's view-points in Nishi-oji-dori and Kita-oji-dori in Kyoto, Japan. In each video, the different contents of VN for the different driving roads are recorded. The VN or SVN, indicating a corner 150 m ahead starts at 5 s after the video starts. The evaluation experiment video, 150 m version was edited by adding VN or SVN at a point such that there was a corner at a distance 150 m away from that point. In this video, the corner instructed by VN or SVN is not pointed out like the SVN create video. SVN was used synthetic voice, created by the SVN create process. Therefore, SVN contents was changed for each participant in this video. There are 20 types of the videos; we have 8 types of video for the SVN (150 m) or VN (150 m), 6 types for both of them (250 m), and 6 types for both of them (325 m).

In the videos, the color and the position of the circle for performing the tasks changes according to the driving behavior of the driver in the video:

- When the driver accelerates or drives at a constant speed, the color of the circle changes to blue.
- When the driver decelerates or is braking, the color of the circle changes to red.
- When the driver turns the wheel to the right, the circle moves to the right.
- When the driver turns the wheel to the left, the circle moves to the left.

Examples of changes of the circle are shown in Figs. 5 and 6.

Fig. 5. An example of changes the color of the circle when the driver is braking. (Color figure online)

Fig. 6. An example of changes of the color of the circle when the driver is accelerating. (Color figure online)

In the evaluation experiment, the participants watched the videos, and performed the following tasks:

- Driving task: It is the task for imposing the participants the work load like driving.
- Understanding task: It is the task for measuring the understanding time of the participants.
- Judgement task: It is the task for measuring the number of the correct decision of the target corner instructed by the guidance.

In the driving task, the participants performed the driving tasks like actual driving operations using the racing wheel, the brake and the accelerator pedal for games.

The movement of the wheel corresponds to the movement of the mouse cursor on the display. So, the participants can look the changes due to the wheel operation. The driving task is that the participant operates the steering wheel or the pedals according to the change of circle in the evaluation experiment videos. The operations performed by the participants are as follows:

- When the color of the circle is red, step on the brake pedal.
- When the color of the circle is blue, step on the accelerator pedal.
- When the position of the circle changes, operate the wheel so that move the mouse cursor to the center of the circle.

We conducted the preliminary experiments on 16 participants to evaluate the driving task. As a result, 11 participants evaluated that the task is similar to actual driving. Therefore, it is suggested that the task simulated the actual driving operation.

In the understanding task, the participants click the button on the wheel when they under-stand the instruction contents from the SVN or the VN. We record the length of time from the start of the evaluation experiment video to time the button pressed.

In the judgement task, the participants click another button on the wheel when they think the car in the video reached the target corner instructed by the SVN or the VN. We record the length of time from the start of the evaluation experiment video to the time the button pressed. To evaluate the accuracy of decision, we set a time range of
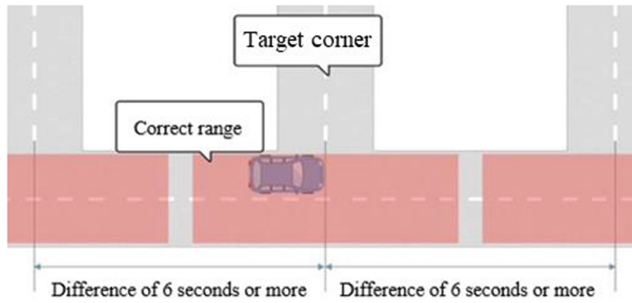
Fig. 7. The correct range in the judgement task.

correct answer for the understanding task. The correct answer range is set to 3 s before and after the time when the car reached the target corner. The correct range is shown in Fig. 7. The time length between all corners is more than 6 s in the videos. We set the correct answer range so that there is no overlap between the correct answer ranges and that there is no misjudgment. For the judgment task, we also conducted the preliminary experiment on 16 participants. As a result, we judged all of the indicated corners by the participants based on the recorded time in the judgement task. Therefore, there is no problem of the setting the correct answer range.

The outline of the evaluation experiment is shown in Fig. 8. At first, we separated the participants into group A and group B. After that, the groups A and B evaluated the SVN and the VN in a different order. In the evaluation experiment, we used the system for playing the evaluation experiment video and recording data about the tasks. At first, the system started playing the video and recording the data about the tasks. Next, the system started playing the SVN or the VN when 5 s after the video started playing. The participants performed the tasks while watching the videos and listening to the SVN or the VN. After repeating 5 times this flow, the type of voice guidance was changed. For example, at first, the group A evaluated the VN 5 times. after that, changed the type of voice guidance from the VN to the SVN. After that, the group A evaluated the SVN 5 times in the same way. It is because that avoid the influence of the order effect. By using Oculus Rift CV 1, the participant watched the videos and listened the SVN or the VN. The experimental environment is shown in Fig. 9. Finally, we applied the questionnaire about DSQ and WSQ to examine the participant's driving characteristics.
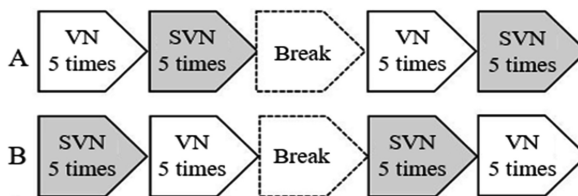
Fig. 8. The outline of the evaluation experiment.

Fig. 9. The experimental environment.

4.4   Results

The results of the understanding tasks are shown in Table 4. The average of the understanding time when using the SVN was 4.71 s (SD: 1.38), and when using the VN was 5.16 s (SD: 1.40). The 13 participants of 14 participants had the shorter time for using the SVN than using the VN. In addition, in each participant, we analyzed the understanding time of the SVN and the time for the VN, using an unpaired t test. As a result, a significant difference was appeared for 4 participants ($p < 0.05$). The time for the SVN of all the participants who were recognized significant differences were shorter than the time for the VN.

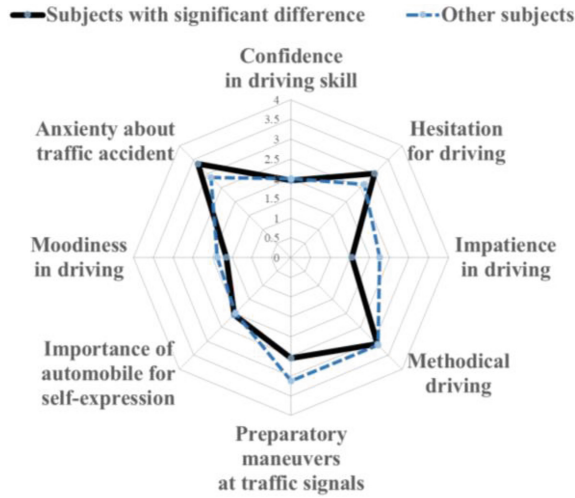**Table 4.** The results of the understanding tasks.

Participant	A	B	C	D	E	F	G	H	I	J	K	L	M	N
VN	3.97	5.05	5.44	4.47	5.54	5.52	4.84	5.06	4.98	4.55	9.25	5.73	5.8	5.07
SVN	3.37	3.93	5.56	3.59	5.21	4.64	3.75	3.69	3.84	4.44	9.17	4.95	5.24	4.25
t test	n.s.	**	n.s.	n.s.	n.s.	n.s.	**	**	**	n.s.	n.s.	n.s.	n.s.	n.s.

n.s.: Not significant **: $p < 0.05$

Furthermore, using the results of DSQ and WSQ, we compared the participants who were recognized significant differences to the other participants. The results of DSQ and WSQ are shown in Figs. 10 and 11. As the result, the participants with the significant differences have the following features:

- The scores for "Impatience in driving" is low.
- The scores for "Preparatory maneuvers at traffic signals" is low.
- The scores for "Patience with driving pace" is low.

The results of the numbers of the correct answers in the judgement task are shown in Table 5. The average number of correct answers in the judgement task when using

Fig. 10. The comparison of the results for DSQ between the participants with and without significant difference in the understanding time.
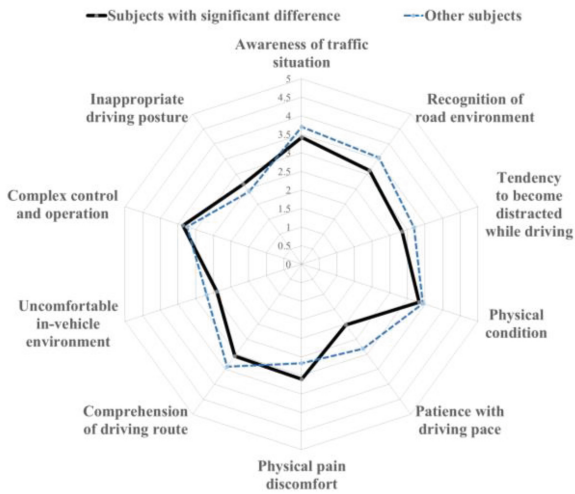
Fig. 11. The comparison of the results for WSQ between the participants with and without significant difference in the understanding time.

the SVN was 5.28 (SD: 1.94), and when using the VN was 4.21 (SD: 1.93). The 7 participants of 14 participants had the larger number of the correct answers for the SVN, 5 participants had the smaller number of the correct answers for the SVN, and 1 participants had the same number of the correct answers. In addition, in each participant, we analyzed the numbers of the correct answers in the judgement task for the SVN and the VN, using the $\chi 2$ test (Fisher two-sided test). As a result, a significant

difference was appeared for 2 participants (p < 0.10). For all the participants who were recognized significant differences, the numbers of the correct answers for the SVN was larger than the numbers for the VN. In the results of the judgement task, there were only 2 participants who was recognized significant difference. For this reason, we didn't compare the results of DSQ and WSQ of the participants who were recognized significant difference to the other participants.

The contents used in the SVN as the soliloquy of the participants are shown in Table 6. From this result, we classified the types of the SVN as follows:

- Time (Abstract): It is the SVN using abstract time expression (e.g. "a little later").
- Time (Concrete): It is SVN using concrete time expression (e.g. "after 10 s").
- Distance (Abstract): It is the SVN using abstract distance expression (e.g. "a bit far").
- Distance (Concrete): It is SVN using concrete distance expression (e.g. "300 meters away").
- Intersection: It is SVN using information of intersection (e.g. "the second intersection").
- Other: It is SVN not applicable to the above types.

Table 5. The numbers of the correct answers in the judgement tasks.

Participant	A	B	C	D	E	F	G	H	I	J	K	L	M	N
VN	7	5	4	4	4	2	6	4	5	3	7	0	6	2
SVN	3	6	7	7	4	7	5	9	3	7	4	2	4	6
x^2 test	n.s.	n.s.	n.s.	n.s.	n.s.	*	n.s.	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

n.s.: Not significant *: p < 0.10

The result of classifying the SVN is shown in Table 7. A feature of the SVN for 150 m distance is that many participants expressed the SVN of Time (Abstract) or Intersection type relatively. In addition, a feature of the SVN for 250 m is that most participants used Distance (Abstract) type. As a feature of the SVN for 325 m, there are various types of the SVN. In English, the participant F's SVN for 150 m means "It is about time I'll move to the left a bit.". The action of going to the left side is done before turning the corner to the left. Since this does not match to the above types, we classified it as Other. In addition, in English, the participant N's SVN for 325 m means "I will keep driving on this lane.". This does not directly represent the distance or time to the corner, and then we classified it as Other. The combination of types of the SVN which was used most frequently was "150 m: Time (Abstract), 250 m: Distance (Abstract), 325 m: Distance (Abstract)", and 4 participants used it.

4.5   Discussion

In the result of understanding task, we analyzed the understanding time of the SVN and the time for the VN, using an unpaired t test. As a result, a significant difference was appeared for 4 participants (p < 0.05). The time for the SVN of all the participants who were recognized significant differences were shorter than the time for the VN. That was, it was suggested that SVN is effective in reducing understanding time. Furthermore, from the

Table 6. The contents of all the SVN.

Participant 150 m	250 m	325 m
A	mousugu hidari houkou	mousukoshi saki de hidari houkou	mousukoshi tattara hidari houkou
B	soko hidari dayone	ato sukoshi de hidari houkou	2tume no shinngou wo hidari
C	tugi no magarikado wo hidari houkou	supu-do sonomama de 15byou teido	300m saki hidari kana
D	sukoshi saki hidari houkou	kono saki hidari houkou	15byougo hidari houkou
E	mousugu hidarihoukou yana	kekkou susunnde kara hidari	daibu saki made ittara hidari yana
F	sorosoro hidari yottkou kana-	kono saki hidari ikanakyana	moutyotto shitara hidari magannnakyana
G	mousukoshi de hidari	tyotto ittara hidari	kono saki moutyoi ittara hidari
H	tugikurai hidari	moutyotto saki hidari	30byou gurai ato hidari
I	sorosoro hidari yana	mousukoshi saki de hidari yana	ato mousukoshi hashittara hidari yana
J	2ko saki wo hidari houkou	100m saki wo hidari houkou	300m saki hidari yana-
K	mousugu hidari kana	mousukoshi saki hidari yana-	mousukoshi saki de hidari yana-
L	mousugu hidari kana	mousukoshi saki hidari yana-	tyotto ookime no kousatenn no tokoro wo hidari kana
M	2tume no tokoro hidari	moutyoi saki hidari	tyotto oodoori no tokoro wo hidari kana
N	shinngou no tugi no kado wo hidari yana	moutyotto saki yana	mada massu de ikka

results of the scores for DSQ and WSQ, the average score of the participants who had significant differences in their understanding time was smaller than one of the other participants in terms of "Impatience in driving", "Preparatory maneuvers at traffic signals", "Patience strength", "Patience with driving pace". Thus, it is considered that the SVN is especially effective for the drivers who prefer to drive at their own pace in the understanding time. However, in this experiment, since we focused only on the inexperienced drivers as the participants, it is not clear whether the same fact is confirmed for all common drivers.

In the result of the judgement task, we analyzed the average number of correct answers of the SVN and the number of the VN, using an unpaired t test. As a result, a significant difference was appeared for 2 participants ($p < 0.10$). For all the participants who were recognized significant differences, the numbers of the correct answers for the SVN was larger than the numbers for the VN. That was, it was suggested that SVN is effective in improving the accuracy of decisions. However, it is hard to conclude the effectiveness of the SVN because there are only a few participants who have significant differences. We expect that the reason of this result is the small count of trials. For example, in this

Table 7. The result of classifying SVN.

Participant	150 m	250 m	325 m
A	Time (Abstract)	Distance (Abstract)	Time (Abstract)
B	Distance (Abstract)	Time (Abstract)	Intersections
C	Intersections	Time (Concrete)	Distance (Concrete)
D	Distance (Abstract)	Distance (Abstract)	Time (Concrete)
E	Time (Abstract)	Distance (Abstract)	Distance (Abstract)
F	Other	Distance (Abstract)	Time (Abstract)
G	Time (Abstract)	Distance (Abstract)	Distance (Abstract)
H	Intersections	Distance (Abstract)	Time (Concrete)
I	Time (Abstract)	Distance (Abstract)	Distance (Abstract)
J	Interse tions	Distance (Concrete)	Distance (Concrete)
K	Time (Abstract)	Distance (Abstract)	Distance (Abstract)
L	Time (Abstract)	Distance (Abstract)	Intersections
M	Intersections	Distance (Abstract)	Intersections
N	Intersections	Distance (Abstract)	Other

experiment, the participant does not have a significant difference unless the difference between the correct answers of the SVN and the VN is 5 or more. That is, if the participant answers the target corners correctly more than half of trials in both cases of the VN and the SVN, there is no significant difference among them. Therefore, it is necessary to increase the number of trials and to evaluate them again. However, if the number of trials is increased, the participant will become familiar with the SVN and the VN. Thus, the appropriate number of trials should be set in the examination.

In addition, by analyzing the contents of the SVN, it was revealed that the tendency of the types of content varies depending on the distance to the instructed target corner. However, it is possible that the tendency was affected by the SVN create video. The contents of the VN was placed in the types of "150 m: Time (Abstract)", "250 m: Distance (Abstract)", "325 m: Distance (Concrete)". Many participants answered their tweets categorized in "Distances (Abreast)" for the video of 250 m, and we generated the SVN (250 m) based on the tweets. Therefore, we need further consideration to yield any facts about the types of the SVN. On the other hand, there were the "Intersections" contents of the SVN, while the information of intersection is not used in the VN. The "Intersections" contents express the number of the intersections to the target corner in the SVN creation video. Such the number does not correspond to the number of the intersections in the evaluation experiment video. We explained this fact in the experiment, however, many participants used the expression using intersections. In addition, 6 participants of 7 participants using Intersection type had more number of the correct answers for the SVN than the VN. Furthermore, the contents for the SVN (250 m) of the participant J "100 m ahead to the left". Even though the correct distances to the target corner and the distance in the SVN were different, the participant J had the more number of the correct answers for the SVN than the VN. Therefore, it is suggested that an exact expression of the voice guidance does not lead drivers to easy understanding. In the analyze of the types of the SVN, 2 contents were classified as Other. The content

of the participant F for the SVN (150) means "It is about time I move to the left a bit.". The action of going to the left side is done before turning the corner to the left. The expression about the action does not indicate the distance to the target corner clearly.

However, the participant F has the more number of the correct answers for the SVN than the VN, and the participant F was recognized significant difference. It seems that the distance to the target corner was recognized by imagining the action performed 150 m before. Since this result is seen only for the participant F, it is necessary to study in the future whether other drivers may have similar tendencies. The combination of types of the SVN which was used most frequently was "150 m: Time (Abstract), 250 m: Distance (Abstract), 325 m: Distance (Abstract)", and 4 participants used it. Each of the other combination were used only by one participant. In the analysis, we classified the contents of SVN based on the expression about distance, time, or intersections excepting the endings of the tweets and dialect, the patterns of the SVN were dispersed. Therefore, it is said that SVN has so many patterns, and it is needed to further study and consideration to yield any findings about the patterns.

5   Conclusion

In this research, we proposed a new expression of voice guidance, called SVN, which uses soliloquy of a driver. We considered the situation where the navigation adopts the voice guidance only. Our goal was to verify the hypothesis that the use of SVN is more effective than that of a general voice guidance (VN). To evaluate the efficiency, we focused on the time required for understanding the instruction content, called understanding time, and the accuracy of identifying the instruction, called the accuracy of decision. That was, we evaluated the understanding time and the accuracy of decision for SVN and VN in the experiment. In addition, we showed the suitable type of driver to the SVN and the types of soliloquy during driving.

In each participant, we analyzed the understanding time of the SVN and the time for the VN, using an unpaired t test. As a result, a significant difference was appeared for 4 participants ($p < 0.05$). The time for the SVN of all the participants who were recognized significant differences were shorter than the time for the VN. From the result, it was suggested that SVN is effective in reducing understanding time. Furthermore, using the results of DSQ and WSQ, we compared the participants who were recognized significant differences to the other participants. From the results, it was speculated that SVN is especially effective for reducing understanding time of the drivers who prefer to drive at their own pace.

In each participant, we analyzed the numbers of the correct answers in the judgement task for the SVN and the VN, using the $\chi 2$ test (Fisher two-sided test). As a result, a significant difference was appeared for 2 participants ($p < 0.10$). For all the participants who were recognized significant differences, the numbers of the correct answers for the SVN was larger than the numbers for the VN. From the result, it was suggested that SVN is effective in improving the accuracy of decisions. In the results of the judgement task, there were only 2 participants who was recognized significant difference.

In addition, by analyzing the contents of the SVN, it was revealed that the tendency of the types of content varies depending on the distance to the instructed target corner. In addition, it is suggested that an exact expression of the voice guidance does not lead drivers to easy understanding.

References

1. Summala, H., Dave, L., Matti, L.: Driving experience and perception of the lead car's braking when looking at in-car targets. Accid. Anal. Prev. **30**, 401–407 (1998)
2. Freundschuh, S.M., Mercer, D.J.: Spatial cognitive representations of story worlds acquired from maps and narrative. Geogr. Syst. **2**, 217–233 (1995)
3. Otani, R., Kansaki, H.: The presentation of route guidance information and better wayfinding. Chukyo Univ. Bull. Psychol. **1**, 45–59 (2001)
4. Taylor, H., Tversky, B.: Assessing spatial representation using text. Geogr. Syst. **2**, 235–254 (1995)
5. Kawai, M., Kato, S., Minobe, N., Tsugawa, S.: Driver adaptive audio route guidance information for car navigation systems. Soc. Automot. Eng. Jpn. Acad. Lect. Meet. Prepr. **37**, 173–178 (2006)
6. Ishibashi, M., Okuwa, M., Doi, S., Akamatsu, M.: Indices for characterizing driving style and their relevance to car following behavior. In: SICE, 2007 Annual Conference, pp. 1132–1137 (2007)
7. Akamatsu, M., Kurahashi, T., Ishibashi, M.: Driver's status assessment using physical measures. In: XVth Triennial Congress of the International Ergonomics Association on Proceedings (2003)
8. Galea, L.A., Kimura, D.: Sex differences in route-learning. Pers. Individ. Differ. **14**, 53–65 (1993)
9. Tomita, Y., Kuriyagawa, Y.: Effect of sexuality, driving experience and habit on driving attitude and workload consciousness. Soc. Autom. Eng. Jpn. Acad. Lect. Meet. Prepr. 7–10 (2005)
10. Peirce, J.W.: PsychoPy - psychophysics software in Python. J. Neurosci. Methods **162**, 8–13 (2007)
11. Sugiura, K., Zettsu, K.: Rospeex a cloud robotics platform for human-robot spoken dialogues. In: Proceedings of IEEE/RSJ IROS. pp. 6155–6160 (2015)