

운전자 모니터링 환경에서 LLM 기반 설명 가능한 음성 경고가 사용자 경험에 미치는 영향

한혜승¹⁾ · 주다영^{*1,2)}

국민대학교 TED 스마트경험디자인학과¹⁾ · 국민대학교 AI디자인학과²⁾

The Impact of LLM-Generated Explainable Voice Warnings on User Experience in Driver Monitoring Systems

Hye Seung Han¹⁾ · Da Young Ju^{*1,2)}

¹⁾Department of Smart Experience Design, TED, Kookmin University, Seoul 02707, Korea

²⁾Department of AI Design, Kookmin University, Seoul 02707, Korea

(Received 3 April 2026 / Revised 4 May 2026 / Accepted 8 May 2026)

Abstract : Driver monitoring systems (DMSs) have shifted from simple warning devices into advanced systems capable of predicting a driver's condition. However, even as existing DMS primarily rely on in-vehicle cameras to analyze driver behavior and issue warnings, while some systems provide more contextual information, the effectiveness of such explanations has not been sufficiently validated. Therefore, this study developed a sentence-based voice warning that explains the cause and recommends follow-up actions, and evaluates whether this approach improves user acceptance compared to beep alerts. Quantitative results verify that the sentence-based warning improves user satisfaction and awareness in distracted driving situations, and that the explanatory nature of the warnings positively influences user understanding and safety awareness. Furthermore, the findings emphasize the importance of explainability in warning design and provide implications for user-centered in-vehicle interfaces in AI-defined vehicle environments.

Key words : Proactive interface(능동형 인터페이스), Driver monitoring system(운전자 모니터링 시스템), Explainable AI(설명 가능한 인공지능), Human-machine interface(인간 기계 상호작용), Voice warning(음성 경고), Usability evaluation(사용성 평가)

1. 서론

운전자의 부주의 및 졸음 운전은 주요 교통 안전 문제로 지속적으로 지적되고 있다.¹⁾ 이에 따라 유럽연합(EU)은 2024년부터 신차에 DMS 장착을 의무화하는 등 운전자 모니터링 시스템(DMS)은 차량 안전의 핵심 요소로 자리잡고 있다.^{2,3)} 이러한 변화에 맞춰 국내외 제조사들은 적외선 카메라, 뇌파 기반 센서, 다중 센서 융합 기술 등을 활용하여 운전자의 상태를 정밀하게 추정하는 시스템을 개발하고 있다.^{4,6)} 최근 AI 기술의 발전과 함께 자동차 산업은 단순 이동 수단을 넘어 탑승자의 상태와 맥락을 인식하고 반응하는 지능형 공간, 즉 AIDV(AI-Defined

Vehicle)로 진화하고 있다.^{7,8)} 주행 중 차량 내·외부 환경 정보를 실시간으로 인식하는 것을 넘어, 운전자 상태 및 차량 내부 맥락 등 다양한 데이터를 통합적으로 처리하여 차량 제어 및 서비스와 연동하는 수준에 이르렀다. 이제 차량은 운전자의 입력에 반응하는 기계적 시스템이 아니라, 상황을 해석하고 능동적으로 개입하는 지능형 주체로 변화하고 있다.⁹⁾ 이러한 흐름 속에서 DMS 역시 단순 감지 기능을 넘어 운전자 상태와 주행 맥락을 실시간으로 파악하고 상황에 적합한 방식으로 개입하는 능동적 역할을 수행하게 되었다. 그러나 운전자의 인지 자원이 제한된 차량 환경에서 예측 기반 경고는 실제 위험

^{*}A part of this paper was presented at the KSAE 2026 Spring Conference

^{*}Corresponding author, E-mail: dyju@kookmin.ac.kr

^{*}This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

상황 이전에 개입이 발생하기 때문에, 운전자가 경고의 타당성을 직관적으로 이해하기 어려우며 이는 시스템에 대한 신뢰도 저하로 이어질 수 있다. 차량 내 AI 경고 맥락은 일반적인 환경에 비해 개입 타이밍, 정보의 간결성, 탐지 정확도, 신뢰도에 대한 요구 수준이 높다. 따라서 경고의 원인과 맥락을 명확히 전달하는 설명 가능성을 고려한 인터페이스 설계가 요구된다. 실제로 최근 이러한 방향의 전환이 시도되고 있으며, 운전자 상태를 반영한 맥락 기반 문장형 음성 경고 방식이 제안되고 있다.^{10,11)} 그러나 설명 가능성 관련 연구는 주로 자율주행 차량의 주행 성능 및 외부 인터페이스에 집중되어 왔으며, 차량 내부 경고 상황에서 설명 가능한 경고 방식이 운전자에게 미치는 영향을 실험적으로 검증한 연구는 제한적인 실정이다. 이에 본 연구에서는 LLM을 활용한 문장형 음성 경고 방식과 기존 고정형 비프음과의 비교 실험을 통해 신뢰도, 사용성, 안전 인식 등 사용자 경험에 미치는 영향을 실증적으로 검증하고자 한다.

2. 선행 연구

운전자 모니터링 시스템(DMS)은 운전자 행동을 감지하는 방식에 따라 차량 기반, 생리적, 행동적 접근으로 구분되며,^{12,13)} 이 중 행동적 접근은 차량 내 카메라로 운전자의 외적 행동을 실시간 분석하는 방식으로,¹⁴⁾ 비침습적이고 실용적이라는 점에서 가장 활발히 연구되고 있다. 최근 연구에서는 졸음 감지 프레임워크 제시,¹⁵⁾ 얼굴·신체 정보를 결합한 이중 입력 구조,¹⁶⁾ 위험 발생 이전 예측 모델¹⁷⁾ 등 탐지 정확도 향상에 초점을 맞춘 연구들이 축적되어 왔다. 그러나 운전자 상태를 정밀하게 탐지하더라도 해당 정보를 효과적으로 전달하지 못하면 실제 행동 변화를 유도하는 데 한계가 존재한다. 차량 내 경고 및 알림 인터페이스는 탐지된 위험을 운전자에게 전달하는 핵심 수단으로, 전통적으로는 경고음·진동·시각 알림과 같은 단순 감각 피드백 방식이 주로 활용되어 왔다.^{18,19)} 하지만 차량 내 인터페이스에 인공지능이 결합되면서, 시스템은 운전자의 상태와 주행 맥락을 실시간으로 인식하고 상황에 적합한 방식으로 개입하는 능동적 역할을 수행할 수 있게 되었다. 이에 따라 인공지능 기반 인터페이스 설계에서는 개입의 원인과 맥락을 명시적으로 설명하는 설명 가능성(Explainability)이 사용자 신뢰 향상과 효과적인 인간-기계 상호작용의 핵심 요소로 부각되고 있다.^{20,21)} 실제로 위험 상황에서 문장 기반 음성 경고는 청자의 정서적 안정에 기여하며,²²⁾ 차량용 인터페이스에서도 문장형 경고가 운전자의 주관적 인지 부하를 낮출 가능성이 제시된 바 있다.²³⁾ 최근에는 대규모



Fig. 1 Camera-based DMS configurations in vehicles



Fig. 2 Example of camera-based sensing for driver monitoring

언어 모델(LLM)을 활용한 실시간 맥락 적합 문장 생성이 가능해지면서 전통적인 명령-응답 구조를 넘어 AI 시스템이 사용자의 맥락을 이해하고 능동적으로 개입하는 상호적 구조로 발전하고 있음을 확인할 수 있다. 이러한 흐름은 차량 내·외부 상황을 탐지하여 AI를 활용한 문장형 음성 알림을 능동적으로 제공하는 기술¹⁰⁾을 통해 상용 시스템 차원에서도 맥락 기반 설명형 경고의 필요성이 인식되고 있음을 보여준다. 종합하면, 기존 선행 연구들은 탐지 정확도에 집중된 모니터링 연구이거나, 녹음된 고정형 메시지를 활용하거나 메시지 생성에 대해서 실험적으로 검증한 연구는 아직 제한적인 실정이다.

3. 연구 방법

3.1 실험 참여자

실험 참여자는 남성 16명과 여성 17명을 포함한 총 33명으로, 평균 연령은 30.8세($SD = 10.59$), 평균 운전경력 6.27년($SD = 8.90$)이다. 실험에 앞서 모든 참가자는 연구 목적과 실험 과업에 대한 안내를 받고 연구 참여 동의를 자발적으로 작성한 후 실험을 진행하였다. 본 연구는 국민대학교 생명윤리위원회(IRB)의 사전 승인을 받아 관

런 규정을 준수하여 진행되었다(KMU-202511-HR-522).

본 연구는 문장형 경고 방식의 효과를 보다 민감하게 관찰하기 위해 운전 경험이 상대적으로 적은 20대 운전자를 중심으로 실험을 설계하였다. 이는 선행연구 및 파일럿 실험 결과를 근거로 하며, 선행연구에 따르면 운전 경험이 상대적으로 적은 운전자일수록 원인과 맥락을 포함하는 설명적 피드백이 즉각적인 행동 대응에 보다 효과적으로 작용할 가능성이 보고되었다.^{24,25)} 파일럿 실험 또한 운전 경력이 낮은 참가자일수록 문장형 음성 경고에 대해 상대적으로 높은 수용도를 보임을 확인하였다. 이에 따라 문장형 경고 방식의 효과를 보다 민감하게 관찰할 수 있는 집단으로 20대 운전자를 주요 참가자로 선정하였다. 다만 본 연구의 표본이 20대 중심으로 구성되어 있고, 특히 음성 경고에 대한 반응력 및 판단력 차이가 나타날 수 있는 고령 운전자의 비율이 적다는 점에서 전체 운전자 집단으로의 일반화에는 한계가 있다. 향후 연

구에서는 다양한 연령대 및 운전 경력을 포함한 확장된 실험을 통해 보다 탐색적 결과를 도출할 필요가 있다.

3.2 실험 환경

실험 환경은 운전자의 전방 시야를 방해하지 않도록 스티어링 컬럼 상단에 웹캠을 설치하여 얼굴을 실시간으로 인식할 수 있도록 하였다(Fig. 3). 전방 모니터에는 낮 시간대 고속도로 주행 영상을 재생하여 주행 맥락이 반영된 상황에서 시스템의 성능을 평가할 수 있도록 설계하였다.

3.3 행동 정의

선행연구를 바탕으로 빈번하게 나타나는 부주의 행동을 반영하여,²⁶⁻²⁹⁾ 운전자의 얼굴 및 시선 움직임을 중심으로 한 4가지 행동(눈 감김, 하품, 시선 이탈, 하향 주시)과 객체 탐지를 통한 휴대폰 사용을 포함한 총 5가지 행동을 정의하였다.

Table 1 Demographic characteristics of participants

Category	Groups	N
Gender	Male	16
	Female	17
Age	20-24	8
	25-29	15
	30-39	5
	40 and above	5
Driving experience	< 2 years	14
	2- 5 years	5
	5 - 10 years	7
	> 10 years	7

Table 2 Definition of distracted driving behavior

Action	Description
Eye closure	Drowsy driving, Reduced attention
Yawning	Cumulative fatigue, Precursor to drowsiness
Gaze off-road	Scanning around, Distracted driving, Head turning or bending, etc.
Head down	Looking downward to interact IVI or controls, Drowsy driving, etc.
Mobile phone use	Checking messages, Entering navigation, or other phone-related tasks.

3.4 실험 설계

본 연구의 시스템은 탐지, 판단, 생성의 세 모듈로 구성되며(Fig. 4), 각 모듈의 세부 메커니즘은 3.5절에서 기술한다. 탐지 모듈에서는 MediaPipe Face Landmarker³⁰⁾와 객체 탐지 모델을 활용하여 웹캠 영상으로부터 운전자의 얼굴 특징점 좌표 및 휴대폰 객체 정보를 실시간으로 추출한다. 판단 모듈에서는 추출된 값에 개인별 캘리브레이션 기준값과 행동별 임계값을 적용하여 부주의 행동 유형을 문자열 라벨로 출력한다. 생성 모듈에서는 행동 라벨을 GPT-4o API에 전달하여 상황에 적합한 경고문을 생성하고, macOS 내장 TTS를 통해 음성으로 출력한다. 참가자는 5가지 부주의 행동을 차례로 수행하며 비프음 방식³¹⁾과 문장형 방식을 모두 경험한 후 두 조건에 대한 평가를 진행하였다.



Fig. 3 Experimental environment

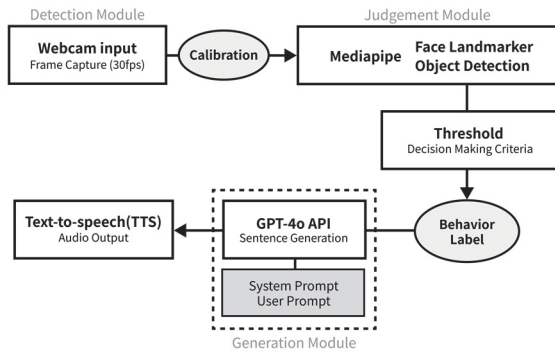


Fig. 4 Flow chart of the experiment

3.5 시스템 프로세스

3.5.1. 신호 가공 및 상태 판별

본 연구의 시스템은 웹캠 영상으로부터 운전자 상태를 실시간으로 추정하고, 이를 문장형 경고 생성을 위한 입력 정보로 변환하는 구조를 갖는다. 입력 영상은 프레임 단위로 처리되며, MediaPipe Face Landmarker는 468개의 얼굴 특징점 좌표를 조합하여 52가지의 표정 가중치(Blendshape score)로 변환한 뒤 라벨을 추출한다. 정확한 판별을 위해 실험 시작 시 30프레임 동안 정면 응시 및 기본 자세 값을 수집하여 개인별 캘리브레이션을 수행하였다. 표정 가중치(w_i)는 MediaPipe Task의 사전 학습된 알고리즘을 통해 산출되며, 캘리브레이션 후 중립 상태($B_{neutral}$)와 각 표정 기저(B_i)의 가중 합산 형태로 처리된다. 본 연구에서는 선행연구^{30,32)}를 바탕으로 중립 상태 대비 표정 변화의 정도를 식 (1)과 같이 수식적으로 정의하였다.

$$V_{final} = B_{neutral} + \sum_{i=1}^{52} w_i * (B_i - B_{neutral}) \quad (1)$$

이후 실시간 입력값이 기준값 또는 사전 정의된 임계값을 초과하는지를 분석하여 5가지 부주의 행동 상태를 판별하였다. 휴대폰 사용 여부는 객체 탐지 모델의 신뢰도 점수(Confidence score)를 기준으로 별도 판별하였다. 임계값 및 신뢰도 점수는 파일럿 실험을 통해 탐지 민감도

와 오탐지 간의 균형을 고려하여 경험적으로 결정하였으며, 각 행동별 세부 판별 기준은 Table 3에 제시하였다.

3.5.2 LLM 기반 경고 생성

판별된 운전자 상태는 LLM 입력을 위한 문자열 라벨로 변환되며 각각 “눈 감기”, “하품”, “시선 이탈”, “하향 주시”, “휴대폰 사용”으로 변환된다. 해당 라벨은 사용자 프롬프트로 GPT-4o API에 전달된다. 표정 가중치, Pitch 값 등 수치 데이터는 LLM에 직접 입력되지 않으며, 임계값 판단을 거쳐 추상화된 행동 라벨이 입력 값으로 사용된다. 시스템 프롬프트에는 경고 문장의 생성 조건을 사전 정의하였으며, 행동 유형별 경고 내용 및 후속 행동 지침, 상황에 따른 어조 규칙(주의 환기 시 부드럽고 공손한 어조, 즉각적 경고 시 짧고 명확한 표현),³³⁾ 출력 길이 제한(약 2~5초 이내)^{34,35)}을 포함하도록 설계하였다. LLM 입력은 행동 라벨과 시스템 프롬프트를 결합한 구조로 구성되며 전달된 행동 라벨을 기반으로 원인 설명과 행동 지침이 결합된 자연어 문장을 생성한다. 생성된 문장은 macOS 내장 TTS를 통해 실시간 음성으로 출력된다.

3.6 설문 문항

객관식 설문 문항은 사용성 평가와 관련된 맥락적 효과성(CE)과 시스템 신뢰성(SR) 2개 영역으로 구성하여 경고 문장의 적절성, 신뢰성, 이해 용이성, 효과성 등의 사용자 경험 요소를 5점 Likert 척도로 평가하였다(Table 4). 평가 문항은 SUS(System Usability Scale)³⁶⁾와 이를 음성 인터페이스에 맞게 확장한 VUS(Voice Usability Scale)³⁷⁾ 및 UX 관련 평가 프레임워크³⁸⁻⁴¹⁾를 참고하여 구성하였으며, 한국어로 번역한 후 연구 맥락에 맞게 일부 표현을 수정하여 사용하였다. 각 영역의 내적 일관성은 SPSS 29를 활용한 크론바흐 알파값(Cronbach's α)으로 검증하였으며, 두 영역 모두 일반적인 탐색적 연구 기준(.60 이상)을 충족하였다. 정서적 반응을 측정하는 단일 문항 혹은 조건, 행동별 응답 차이 분석을 위한 비교 평가 문항은 알파값 산출 대상에서 제외하였다.

3.7 데이터 분석

본 실험은 비프 경고음 방식과 비교하여 문장형 음성 경고 시스템의 사용자 경험을 실증적으로 검증하고자 하였다. 이를 위해 다음과 같은 연구 질문을 설정하였다.

- Research Question 1(RQ1): 문장형 경고 방식은 기존 대비 만족도에 어떠한 영향을 미치는가?
- Research Question 2(RQ2): 문장형 경고 방식은 운전 중 행동 별 방해 인식 및 놀람 정도에 어떠한 영향을 미치는가?

Table 3 Threshold for driving behavior detection

Action	Threshold
Eye closure	Eyeblink score > 0.3(both eyes) for ≥ 0.2 s
Yawning	Mouthopen/jawopen score > 0.3 for ≥ 0.6 s
Gaze off-road	Gaze score deviation from baseline > 0.15
Head down	Head pitch deviation from baseline > 0.05
Mobile phone use	Object classified as 'cell phone' with confidence ≥ 0.5 for ≥ 0.5 s

Table 4 Evaluation framework and questions for usability test

Variable	Question	Scale	α
Contextual effectiveness (CE)	Safety contribution	5-point likert scale	.735
	Warning timing		
	Detection accuracy		
	System consistency		
	Situation awareness		
System reliability (SR)	System trust	5-point likert scale	.694
	Functional adequacy		
	Ease of use		
	System integration		
Emotional reassurance			
Comparative response	Relative effectiveness	5-point comparative likert scale	
	Cognitive distraction		
Task-specific metrics	Message appropriateness		
	Warning timing		
	Disturbance and surprise		
Qualitative feedback	Advantages	Open-ended response	
	Disadvantages		
	Suggestions for improvement		

Note. Emotional Reassurance, Comparative Response and Task Specific Metrics were excluded from the reliability analysis as they represent single-item or conceptually distinct measures, including emotional response and comparative evaluation, which are not suitable for internal consistency assessment.

- Research Question 3(RQ3): 문장형 경고에서 상황 이해 지원성과 사용 용이성은 사용성에 어떠한 영향을 미치는가?
 - Research Question 4(RQ4): 행동과 문장의 적합성은 설명 가능한 경고의 사용성에 어떠한 영향을 미치는가?
- 또한 본 실험은 정량적 평가와 함께 서술형 응답 및 기능 중요도 평가를 포함한 혼합 연구 방법을 적용하였으며, 이를 통해 질적 측면을 보완적으로 분석하였다.

4. 실험 결과

4.1 만족도 비교(RQ1)

본 연구에서는 고정 비프음과 문장형 경고 방식 간의 만족도를 조건별로 평가하였다. 분석에 앞서 Shapiro-Wilk 검정을 통해 정규성을 확인한 결과, 두 조건 모두 정규성 가정이 충족되지 않아($p < .001$) 비모수 검정인 Wilcoxon

부호순위 검정(Wilcoxon signed-rank test)을 수행하였다. 분석 결과, 문장형 경고 방식의 만족도 중앙값은 4.0점으로 비프음 방식의 중앙값인 3.00점에 비해 높은 것으로 나타났다(M = 4.06, SD = 1.029 / M = 3.15, SD = 1.326), 그 차이는 통계적으로 유의하였다($Z = 2.559, p = .010$). 이는 문장형 음성 경고가 기존 경고음 방식 대비 사용자 만족도를 유의하게 향상시킴을 시사한다. 경고의 원인과 맥락을 함께 전달하는 설명형 인터페이스가 개입의 이유를 명확히 이해하도록 도운 것이 전반적인 만족도 향상으로 이어진 것으로 해석된다.

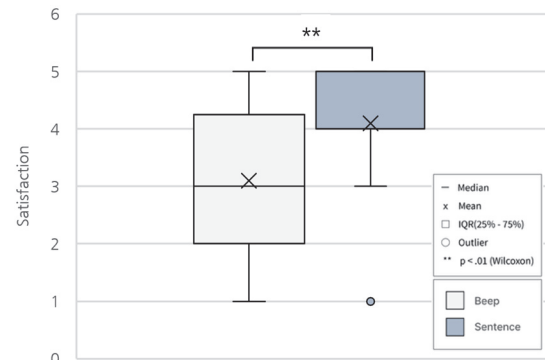


Fig. 5 Comparison of overall satisfaction scores between warning conditions (RQ1)

4.2 방해·놀람 효과(RQ2)

RQ2에 따라 행동 및 경고 조건별 운전 중 방해 및 놀람 정도의 차이를 비교하였다. RQ2 또한 정규성 가정이 충족되지 않아 Wilcoxon 부호 순위검정을 시행하였다. 실

Table 5 Differences in perceived disturbance and surprise between warning types across driving behaviors (RQ2)

Action	Type	M	SD	Δ	r	$Z(p)$
Eye-closure	B	3.03	1.237	+0.636	.470	2.702**
	S	3.67	1.109			
Yawning	B	3.24	1.480	+0.576	.370	2.123*
	S	3.82	1.185			
Gaze-off road	B	3.27	1.353	+0.546	.441	2.532*
	S	3.82	1.103			
Head down	B	3.27	1.376	+0.667	.471	2.708**
	S	3.94	1.059			
Mobile phone use	B	2.94	1.499	+0.939	.618	3.551***
	S	3.88	1.111			

- Type: B(Beep), S(Sentence)

- Effect size (r): small $\geq .1$, medium $\geq .3$, large $\geq .5$.

- * $p < .05$, ** $p < .01$, *** $p < .001$

용적 유의성 검토를 위해 효과 크기는 $r = |Z|/\sqrt{N}$ 로 산출하였다.⁴²⁾ 평균 차이(Δ)는 각 행동 상황에서 문장형 경고와 경고음 방식 간의 차이를 의미하며, 값이 클수록 문장형 경고에 의해 방해 및 놀람 수준이 감소했음을 나타낸다.

분석 결과, 문장형 경고는 기존 경고음 방식에 비해 모든 부주의 행동 유형에서 운전 중 놀람 및 방해 수준이 통계적으로 유의하게 낮은 것으로 나타났다($p < .05$). 특히 휴대폰 조작 상황에서 문장형 경고는 놀람 및 방해 수준을 가장 크게 감소시켰으며($Z = 3.551^{***}$), 효과 크기 또한 큰 효과 수준으로 나타나 실용적으로도 유의미한 차이가 확인되었다($r = .618$). 하향 주시 상황에서도 유사하게 높은 효과 크기가 관찰되어($Z = 2.708, p = .007, r = .471$), 전방 주시가 이루어지지 않는 상황에서 문장형 경고의 효과가 두드러지는 것으로 해석된다. 이는 운전자 주의 이탈 상황에서 문장형 경고가 단순 경고음 대비 낮은 놀람 수준을 유지하면서도 상황 인식과 수용성을 향상시킬 수 있음을 시사한다.

4.3 인지적 요인이 사용성에 미치는 영향(RQ3)

RQ3에 따라 CE 영역의 핵심 변수인 상황 이해 지원성과 SR 영역의 핵심 변수인 사용 용이성이 사용성에 미치는 영향을 검토하였다. 분석에 앞서 Pearson 상관분석을 수행한 결과, 상황 이해 지원성($r = .585, p < .001$)과 사용 용이성($r = .408, p = .018$) 모두 사용성과 강한 정적 상관관계를 보였다. 다중 회귀 분석 결과, 모형은 유의하였으

며($F(2,30) = 12.992, p < .001$), 두 변수는 사용성 변동에서 약 46.4 %를 설명한다. 개별 변수의 영향력을 살펴보면, 상황 이해 지원성($\beta = .548^{***}$)이 사용 용이성($\beta = .351^*$)보다 사용성에 더 강한 영향을 미치는 것으로 나타났다. 이는 문장형 경고가 단순히 이해하기 쉬움을 넘어 운전자가 문장형 경고를 통해 현재 상황을 정확히 이해하도록 돕는지가 사용성에 더욱 중요한 요소임을 시사한다. 따라서 운전 시 인지 자원이 제한된 상황에서는 사용 용이정보다도 경고 원인과 상황을 이해할 수 있도록 하는 인지 중심적 설계가 고려되어야 한다.

4.4 메시지 적합성이 사용성에 미치는 영향(RQ4)

경고 메시지의 적합성이 사용성 평가에 미치는 영향을 확인하기 위해 단순 선형 회귀분석을 수행하였으며 모형은 통계적으로 유의하였다($F(1,31) = 43.850, p < .001$). 분석 결과, 메시지 적합성은 사용성 변동에서 약 58.6%를 설명하는 것으로 나타났다. 또한 유의한 정적 영향을 미치는 것으로 나타났으며($\beta = .765, p < .001$), 적합성이 1 증가하면 사용성이 0.560만큼 증가하는 것으로 판단할 수 있다($B = .560$). 이는 운전자의 상태와 경고 문장의 내용이 얼마나 일치하는지가 사용성에 강력한 영향을 미치는 요인임을 의미한다. 따라서, LLM이 문장을 생성하더라도 그 내용이 실제 운전자 행동과 일치하지 않을 경우 오히려 신뢰도 저하로 이어질 수 있음을 시사하며, 탐지 정확도와 상황 적합성이 설명가능한 경고 인터페이스 설계에서 핵심 설계 변수로 작용함을 확인하였다.

4.5 주요 기능 순위 분석

향후 운전자 모니터링 시스템 내 문장형 경고 방식 설계 시 중요하다고 인식되는 주요 기능 5가지를 선정하여 1순위부터 5순위까지 순위를 부여하도록 하였다. 순위 데이터는 점수화하여 각 기능 별 평균 점수와 표준편차를 산출하였다. 순위 간의 차이를 검증하기 위해 Friedman 검정을 수행한 결과, 항목 간 유의한 차이가 확인되었다($\chi^2(4) = 58.255, p < .001$). 분석 결과, 경고 타이밍의 적절성이 가장 중요한 요소로 평가되었으며 이후 행동 탐지

Table 6 Regression results for factors influencing overall user evaluation (RQ3)

DV	IV	Unstandardized Coeff.		β	t	p
		B	SE			
Usability	(Constant)	1.038	.783		1.326	.195
	Ease of use	.407	.156	.351	2.612	.014
	Situation awareness	.320	.079	.548	4.081	<.001

$R^2 = .464$, Adj. $R^2 = .428$, $F(2,30) = 12.992$, $p < .001$

Table 7 Regression results for the effect of behavior-message fit on user evaluation (RQ4)

DV	IV	Unstandardized Coeff.		β	t	p
		B	SE			
Usability	(Constant)	1.998	.362		5.516	<.001
	Behavior-message fit	.560	.085	.765	6.622	<.001

$R^2 = .586$, Adj. $R^2 = .572$, $F(1,31) = 43.850$, $p < .001$

Table 8 Results of importance ranking for 5 key functions

Rank	Category	M	SD
1	Warning timing	4.09	.947
2	Detection accuracy	3.61	1.088
3	Behavior-message fit	3.45	1.348
4	Message clarity	2.33	.957
5	Natural voice	1.52	.972

$\chi^2(4) = 58.255$, $p < .001$

능력, 행동-문장 간 일치 정확성, 문장 이해도, 음성의 자연스러움 순으로 나타났다. 이러한 결과는 앞선 정량적 평가에서 확인된 결과와 비슷한 경향을 보인다. 다수의 응답자는 경고가 발생하는 시점의 적절성과 함께, 자신의 실제 행동과 음성 안내 내용이 일치하는지를 중요하게 인식하고 있었으며, 탐지 오류나 과하게 해석된 경고는 시스템에 대한 신뢰 저하로 이어질 수 있음을 지적하였다. 이는 정량적 평가 결과와 유사하게 능동형 경고 시스템 설계 시 탐지 성능 향상 뿐만 아니라 경고 타이밍과 메시지의 적합성까지 종합적으로 고려해야 한다.

4.6 정성 분석

능동형 경고 시스템의 사용성에 대한 서술형 응답을 분석하였다(Table 9). 정성 분석 결과는 정서적 수용성, 인지 명확성, 행동 기반 피드백 측면에서 긍정적 평가가 확인되었다. ‘정서적 수용성’ 측면에서 음성 안내는 운전자에게 심리적 안정감을 제공하며 경고에 대한 수용성을 향상시키는 요소로 작용한 것으로 나타났다. 다만, 일부 응답에서는 줄음 상황과 같이 높은 경각심이 요구되는 맥락에서는 부드러운 음성 안내만으로는 충분한 각성 효과를 제공하지 못할 수 있다는 한계가 제기되었다. ‘인지 명확성’ 및 ‘행동 기반 피드백’ 항목에서는 경고 발생의 원인과 맥락을 구체적으로 설명하는 방식이 운전자가 자신의 상태를 이해하는 데 도움을 주며, 결과적으로 시스템에 대한 신뢰를 강화하는 요인으로 작용한 것으로 확인되었다. 다양한 문장형 경고를 통해 반복 경고로 인한 둔감화를 완화하고, 운전자가 자신의 부주의한 행동을 자각하도록 유도함으로써 행동 교정과 관련된 긍정적인 인식이 확인되었다.

- P14 (30대, 남성): 원래는 운전할 때 예민해서 조심을 많이 하는 편인데 적절한 경고음과 안내말이 나오는게 생각보다 효과가 좋다.
- P16 (20대 초반, 남성): 운전자 입장에서 명확하게 잘못된 부분을 이해하기 편했고, 말할 때 ‘쉬어주세요’ 같이 다음 행동을 어떻게 해야 할지 제시해줘서 편찮았습니다.
- P20 (20대 후반, 여성): 내가 무의식적으로 하는 행동에 대해서 파악되니 의식할 수 있는 계기가 될 수 있었다.

한편, 행동-문장 간 불일치, 탐지 정확도 한계, 반복 경고로 인한 피로감 등의 개선 필요 요인이 확인되었다. 행동 인식을 기반으로 음성 피드백을 제공하는 접근은 운전자의 실제 행동과 음성 안내 간의 일치도가 낮을 경우 오히려 시스템에 대한 신뢰 저하로 이어질 수 있다는 우려도 함께 제기되었다. 또한 카메라 기반 탐지 특성상 수

Table 9 Insight for strengths and advantages, weaknesses and improvement factors from qualitative responses

Strengths and advantage factors	
Category	Keywords
Emotional acceptability	Gentle, Less startling, Reduced discomfort
Cognitive clarity	Context-aware, Informative, Understandable
Action-recognition-based feedback	Action-referencing, Recognition, Trust, Accuracy
Action-inductive	Guidance, Take a break, Leading
Corrective effect	Self-reflection, Admonishment, Correction
Expectation of diversity	Sentence variety, Preventing familiarity
Weakness and improvement factors	
Category	Keywords
Misalignment	Overinterpretation, Confusion, Incongruity
Limitations in detection	Detect failure, False positives, glasses, beards
Lack of follow-up action guidance	Next-step guidance, Navigation support, Actionable feedback, Practicality
Increased cognitive load	Lengthy, Requires focus, Fatigue
Concerns about long-term fatigue	Repetition, Annoyance

염, 안경, 모자 등으로 얼굴 특징이 가려지는 경우 인식 정확도가 저하될 수 있다는 점이 주요 개선점으로 확인되었다. 일부 응답에서는 운전자 상태에 대한 경고는 제공되었으나 후속 조치에 대한 안내가 부족할 경우 시스템의 효과가 감소한다는 의견이 존재했으며 인근 휴게소 안내, 공조 및 조명 조절과 같이 차량 내부의 시스템과 연동성을 함께 제공할 필요성이 언급되었다. 마지막으로, 운전 중 다양한 행동을 지속적으로 탐지하고 경고를 제공하는 방식은 반복적 개입으로 인한 피로감이나 거부감으로 이어질 가능성도 제시되었다. 이에 따라 행동의 누적 빈도나 위험 수준을 고려하는 방식이 요구된다.

- P1 (20대 후반, 여성): 내가 한 행동과 다른 얘기를 했을 때 순간적으로 왜 그렇게 생각했지? 등의 의구심이 들어서 신뢰도가 떨어졌다.
- P10 (20대 후반, 남성): 눈을 뜨고 있었는데 눈이 작다 보니 (눈을) 뜨라 해서 기분이 안 좋았다...
- P31 (40대 남성): 하품처럼 어쩔 수 없이 나오는 행동까지 경고하니깐 과하다는 생각도 들었다.

서술형 응답 분석 결과, 정량적 평가만으로는 포착하기 어려운 부분을 구체적으로 보여주며 설명적 음성 피드백이 운전자의 인지 이해와 경고 수용성에 영향을 미칠 수 있음을 확인하였다.

5. 결론 및 향후 연구

본 연구에서는 카메라 기반 운전자 모니터링 시스템 환경에서 탐지된 운전자 상태 정보를 LLM과 결합하여, 경고 원인과 후속 행동을 포함한 문장형 음성 경고 인터페이스를 제안하고 그 효과를 실험적으로 검증하였다. 이는 탐지 정확도 중심의 기존 DMS 연구와 달리, 탐지 이후 정보 전달 방식에 초점을 맞춘 인터페이스 설계 연구라는 점에서 의의를 가진다. 실험 결과, 문장형 경고는 비프음 방식 대비 사용자 만족도를 향상시키고 방해 및 놀람 수준을 감소시키는 효과를 보였다. 또한 상황 이해 지원성과 행동-문장 적합성이 사용성 평가에 유의한 영향을 미치는 주요 요인으로 확인되었다. 따라서 AI 기반 경고 설계는 언제 도입할 것인지 뿐만 아니라, 어떠한 방식으로 설명할 것인지까지 포함하는 통합적 설계 접근이 요구된다.

다만 본 연구는 제한된 표본과 시뮬레이터 환경에서 수행되어 실제 주행 환경으로의 일반화에 한계가 있다. 연구 결과를 바탕으로 다음과 같은 향후 연구 방향을 제시한다.

- 1) 휴대폰 사용과 눈 감김 상황에서 문장형 경고의 효과가 크게 나타난 바, 고위험 행동에는 즉각 경고를, 동일 행동 반복 감지 시에는 단계적으로 강도를 높이는 레벨 분기 구조 설계를 구현할 수 있다.
- 2) 행동-문장 적합성이 사용성을 설명하는 핵심 변인으로 확인된 만큼, 생성 문장과 입력 행동 라벨의 의미 일치 여부를 검증하는 규칙 기반 후처리 필터링 모듈을 파이프라인에 추가할 필요가 있다.
- 3) 정성 분석에서 다수 참여자가 내비게이션, 공조 등의 연계를 요구한 바와 같이, 단순 경고 출력에 그치지 않고 차량 내 기능과 연계된 조치를 자동으로 수행하는 이벤트 기반 트리거 구조를 설계할 수 있다.
- 4) 캘리브레이션 기반 임계값 구조를 확장하여, 운전자의 행동 패턴 및 누적 이력을 반영하여 경고 전략을 동적으로 조정하는 개인화 적응형 제어 구조를 적용할 수 있다.

후 기

본 논문은 산업통상자원부 자동차산업기술개발사업(20018907, 가변형 디스플레이 및 친환경 소재 적용 자율주행 각펫 모듈 기술 개발)의 지원을 받아 수행하였습니다.

References

- 1) S. Kim, Report on Drowsy Driving Accident Status and Reduction Measures, Samsung Fire & Marine Insurance
- 2) T. Pande, K. Shin, R. Prabhu, N. Muralidharan, H. Garud, S. Jagannathan and S. Haas, "Driver and Occupancy Monitoring Systems with AM62A," 2024 IEEE International Conference on Consumer Electronics (ICCE), pp.1-6, 2024.
- 3) LG Innotek, A Safety Partner That Meticulously Monitors Drivers and Passengers: High-Performance In-Cabin Camera Module, <https://www.lginnotek.com/news/insightsView.do?idx=192&locale=ko&rowSize=9>, 2026.
- 4) Hyundai Mobis, M.Brain: Mobility Technology for Safety, <https://www.hyundaimotorgroup.com/ko/story/CONT0000000000032086>, 2026.
- 5) Hyundai Motor Group, Opening a New Era of Commercial Vehicles with World-First Technologies, <https://www.hyundaimotorgroup.com/ko/story/CONT00000000099995>, 2026.
- 6) S. Hong, J. Park, H. Lim, C. W. Lee, J. Lee and J. H. Yang, "Evaluation of EEG-Based Driver Inattention Warning System in Real Road Environment," Transactions of KSAE, Vol.31, No.8, pp.585-596, 2023.
- 7) G. H. Song and B. T. Koo, "AI Semiconductors for AI-Defined Vehicles: Technology Trends, Market Outlook, and Strategic Implications," Electronics and Telecommunications Trends, Vol.40, No.4, pp.12-24, 2025.
- 8) M. De Vincenzi, C. Bodei and I. Matteucci, "When AI Takes the Wheel: AI-Defined Vehicles Principles and Pitfalls," Frontiers in Robotics and AI, Vol.13, Paper No.1770121, 2026.
- 9) B. Li, E. Y. Lai and X. Wang, "From Tools to Agents: Meta-Analytic Insights into Human Acceptance of AI," Journal of Marketing, Vol.90, No.3, Paper No. 00222429251355266, 2025.
- 10) LG Electronics, LG Electronics Partners with Qualcomm to Unveil Future AIDV Technology Based on Generative AI, <https://www.lge.co.kr/story/newsroom/234641>, 2026.
- 11) H. Han and D. Ju, "A Voice Interface System Providing Proactive Interaction Based on Driver Monitoring," KSAE Spring Conference Proceedings, p.968, 2025.
- 12) G. Yang, C. Ridgeway, A. Miller and A. Sarkar, "Comprehensive Assessment of Artificial Intelligence Tools for Driver Monitoring and Analyzing Safety Critical Events in Vehicles," Sensors, Vol.24, No.8, Paper No.2478, 2024.
- 13) S. Cao, P. Feng, W. Kang, Z. Chen and B. Wang, "Optimized Driver Fatigue Detection Method Using Multimodal Neural Networks," Scientific Reports, Vol.15, No.1, Paper No.12240, 2025.

- 14) F. Akhmedov, H. Khujamatov, M. Abdullaev and H. Jeon, "Joint Driver State Classification Approach: Face Classification Model Development and Facial Feature Analysis Improvement," *Sensors*, Vol.25, No.5, Paper No.1472, 2025.
- 15) F. Hassan, A. F. Ibrahim, A. Gomaa, M. A. Makhoul and B. Hafiz, "Real-Time Driver Drowsiness Detection Using Transformer Architectures: A Novel Deep Learning Approach," *Scientific Reports*, Vol.15, No.1, Paper No.17493, 2025.
- 16) M. Kim and G. Choi, "STFTTransNet: A Transformer-Based Spatial Temporal Fusion Network for Enhanced Multimodal Driver Inattention State Recognition System," *Sensors*, Vol.25, No.18, Paper No.5819, 2025.
- 17) K. R. Gopal and K. Ashokkumar, "Predictive Analytics Based Driver Drowsiness Detection and Health Monitoring System," 2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA), pp.1-6, 2025.
- 18) S. B. Oh and J. H. Kim, "An Analysis on Interrupt Latency of Hypervisor for Automotive Software Integration," *Transactions of KSAE*, Vol.30, No.11, pp.901-907, 2022.
- 19) J. Fagerl nn, S. Lindberg and A. Sirkka, "Graded Auditory Warnings during In-Vehicle Use: Using Sound to Guide Drivers without Additional Noise," *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp.85-91, 2012.
- 20) V. Chen, A. Zhu, S. Zhao, H. Mozannar, D. Sontag and A. Talwalkar, "Need Help? Designing Proactive AI Assistants for Programming," *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp.1-18, 2025.
- 21) S. Atakishiyev, M. Salameh, H. Yao and R. Goebel, "Explainable Artificial Intelligence for Autonomous Driving: A Comprehensive Overview and Field Guide for Future Research Directions," *IEEE Access*, Vol.12, pp.101603-101625, 2024.
- 22) J. J. Smith and M. S. Wogalter, "Perceived Effectiveness of Warning Messages for Use as Vocal Warnings in Residential Fire Alarms," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pp.1398-1402, 2007.
- 23) G. You, T. Kim, J. Kim, K. Lee and J. Choi, "The Effects of Voice and Visual Information Types on Driver's Cognitive Workload and Emotion," *The HCI Society of Korea*, pp.986-991, 2018.
- 24) M. Wang, S. C. Lee, G. Montavon, J. Qin and M. Jeon, "Conversational Voice Agents Are Preferred and Lead to Better Driving Performance in Conditionally Automated Vehicles," *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp.86-95, 2022.
- 25) SafeRoads Alberta, *Geared to Go: A Workbook for Coaching New Drivers*, <https://www.registriesdirect.ca/pdf/CoachingGuide.pdf>, 2026.
- 26) C. El Khatib, O. Ou and F. Karray, "Driver Inattention Detection in the Context of Next-Generation Autonomous Vehicles Design: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, Vol.21, No.11, pp.4483-4496, 2020.
- 27) R. Kashevnik, S. Shchedrin, C. Kaiser and A. Stocker, "Driver Distraction Detection Methods: A Literature Review and Framework," *IEEE Access*, Vol.9, pp.60063-60076, 2021.
- 28) Chevrolet, *Driver Attention Assist*, <https://www.chevrolet.com/support/vehicle/driving-safety/driver-assistance/driver-attention-assist>, 2026.
- 29) National Highway Traffic Safety Administration, *Distracted Driving*, <https://www.nhtsa.gov/risky-driving/distracted-driving>, 2026.
- 30) C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays and M. Grundmann, "MediaPipe: A Framework for Building Perception Pipelines," *arXiv Preprint arXiv:1906.08172*, 2019.
- 31) J. H. Han and D. Y. Ju, "Advanced Alarm Method Based on Driver's State in Autonomous Vehicles," *Electronics*, Vol.10, No.22, Paper No.2796, 2021.
- 32) I. Grishchenko, G. Yan, E. G. Bazavan, A. Zangir, N. Chinaev, K. Raveendran and C. Sminchisescu, "Blendshapes GHUM: Real-Time Monocular Facial Blendshape Prediction," *arXiv Preprint arXiv:2309.05782*, 2023.
- 33) Z. Zou, O. Mubin, F. Alnajjar and L. Ali, "A Pilot Study of Measuring Emotional Response and Perception of LLM-Generated and Human-Generated Questionnaires," *Scientific Reports*, Vol.14, No.1, Paper No.2781, 2024.
- 34) International Organization for Standardization, *ISO 15006:2011 - Road Vehicles-Ergonomic Aspects of Transport Information and Control Systems-Specifications and Test Procedures for In-Vehicle Visual Presentation*, ISO, 2011.
- 35) J. L. Campbell, J. L. Brown, J. S. Graving, C. M. Richard, M. G. Lichty, L. P. Bacon and T. Sanquist, "Human Factors Design Guidance for Level 2 and Level 3 Automated Driving Concepts," *National Highway Traffic Safety Administration*, Washington, DC, 2018.
- 36) J. Brooke, "SUS: A Quick and Dirty Usability Scale,"

- Usability Evaluation in Industry, Vol.189, No.3, pp.189-194, 1996.
- 37) D. S. Zwakman, D. Pal and C. Arpikanondt, "Usability Evaluation of Artificial Intelligence-Based Voice Assistants: The Case of Amazon Alexa," SN Computer Science, Vol.2, No.1, Paper No.28, 2021.
 - 38) C. Gold, M. Körber, C. Hohenberger, D. Lechner and K. Bengler, "Trust in Automation-Before and after the Experience of Take-Over Scenarios in a Highly Automated Vehicle," Procedia Manufacturing, Vol.3, pp.3025-3032, 2015.
 - 39) M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," Human Factors: The Journal of the Human Factors and Ergonomics Society, Vol.37, No.1, pp.32-64, 1995.
 - 40) International Organization for Standardization, ISO 9241-210:2010 Ergonomics of Human-System Interaction-Part 210: Human-Centred Design for Interactive Systems, ISO, Switzerland, 2010.
 - 41) E. Macpherson, The UX Honeycomb: Seven Essential Considerations for Developers, <https://medium.com/mytake/the-ux-honeycomb-seven-essential-considerations-for-developers-accc372a398c>, 2019.
 - 42) J. Cohen, "A Power Primer," Psychological Bulletin, Vol.112, No.1, pp.155-159, 1992.