Preprint. Under review. Zavanet Drive Technical Report ZD-2025-04. November 2025.

Conversation-Aware Driving: Smoothing Trajectories for In-Cabin Social Interaction

Lucía Alvarez[1], Riya Patel[1], Marcus Brennan[1,2]

[1]Zavanet Drive, Austin [2]Zavanet, San Francisco

{lucia, riya, marcus}@drive.zavanet.com

Abstract

Autonomous vehicles are usually tuned for passengers who are looking out the window or at a screen. Zavanet Drive's cabin is designed for passengers who are talking to each other, and we find that this changes what "comfortable" means. In a 212-participant study with instrumented rides, passengers engaged in face-to-face conversation reported discomfort at lateral accelerations 28% lower than passengers riding alone or using devices, and were disproportionately sensitive to *jerk* (the rate of change of acceleration) during speech. We propose **conversation-aware planning**: a motion-planner cost term that penalises lateral jerk and acceleration more heavily when an in-cabin signal indicates active conversation, together with a short lookahead that times unavoidable manoeuvres to natural pauses in speech. The signal is derived from on-device audio activity detection and seat orientation only; no speech content is processed or stored. In 1,140 closed-course and public-road rides, conversation-aware planning reduced reported discomfort by 31% and conversational self-interruptions by 44%, at a cost of 1.9% in travel time and no measured change in safety-relevant metrics. We release the discomfort dataset and the planner cost formulation.

1 Introduction

A commute is, for many people, the largest block of time in the day that is not spent working or sleeping. Zavanet Drive's premise is that an autonomous vehicle can return that time to its passengers as time with each other: the cabin seats four people facing one another, with no forward-facing screens by default. In early rides we noticed something the comfort literature did not predict. Passengers who were talking complained about manoeuvres that passengers looking out the window did not notice. A lane change that was unremarkable when riding alone would, mid-sentence, cause a speaker to pause, grab an armrest, or lose their train of thought.

The standard model of ride comfort [1, 2] treats the passenger as a seated body with a vestibular system and defines comfort thresholds on acceleration and jerk accordingly. Those thresholds are derived from passengers who are, typically, attending to the road or to nothing in particular. Conversation changes the passenger's state in two ways: attention is directed inward, toward another person, so motion cues arrive unanticipated; and the face-to-face seating means lateral motion is experienced as movement *relative to the person one is looking at*, which is perceptually more salient than movement relative to the horizon.

This paper quantifies the effect and proposes a planner modification that exploits it. Our contributions are:

- A controlled study ($n = 212$) showing that passengers in conversation have substantially lower tolerance for lateral acceleration and jerk than passengers in the solo or device conditions (§3).
- A conversation-aware cost term for sampling- and optimisation-based motion planners, together with a pause-aligned timing heuristic for unavoidable manoeuvres (§4).
- A privacy-preserving in-cabin conversation signal that uses only voice-activity detection and seat orientation, with no speech content leaving the audio front-end (§4.3).
- Evaluation on 1,140 rides showing a 31% reduction in reported discomfort and a 44% reduction in self-interruptions for a 1.9% travel-time cost (§5).

2 Related Work

Ride comfort in automated vehicles. ISO 2631 [1] and subsequent AV-specific work [2, 3] define comfort in terms of frequency-weighted acceleration and jerk. Bellem et al. [4] found that AV passengers prefer earlier, gentler manoeuvres than human drivers execute. We build on these thresholds and show they shift with passenger activity.

Motion sickness and non-driving tasks. Reading and screen use in AVs increase motion sickness [5, 6], and several works propose planner adjustments for device users [7]. Conversation has received little attention, perhaps because conventional cabins seat passengers facing forward.

Comfort-aware planning. Comfort terms in planner cost functions are standard [8, 9]. Our contribution is to make the weighting *state-dependent* on an in-cabin signal, and to add timing alignment, rather than to propose a new planner.

In-cabin sensing and privacy. Driver-monitoring systems process camera and audio data continuously, raising privacy concerns [10]. We deliberately restrict our signal to activity detection to avoid processing content.

3 Comfort Study

3.1 Design

212 participants (ages 19–74, 51% women) rode in pairs or alone on a 6.2 km closed course in an instrumented Zavanet Drive prototype under three conditions in counterbalanced order: *Solo* (alone, no device), *Device* (alone, reading on a tablet), and *Conversation* (with a companion, instructed to discuss a given topic). The vehicle executed a scripted sequence of 18 manoeuvres per lap (lane changes, curves, decelerations) at three intensity levels. Participants pressed a handheld button when they felt discomfort and rated each lap afterwards on a 7-point scale. Lateral and longitudinal acceleration and jerk were logged at 100 Hz. In the Conversation condition, audio was recorded with consent for offline annotation of speech pauses and self-interruptions; this recording was for the study only and is not part of the deployed system.

3.2 Results

Figure 1 plots the probability of a discomfort button press against peak lateral acceleration for each condition. Fitting a logistic model, the acceleration at which 50% of manoeuvres produced a press was 2.9 m/s² in Solo, 2.5 m/s² in Device and 2.1 m/s² in Conversation, a 28% reduction relative to Solo. Jerk mattered more than acceleration in the Conversation condition: adding peak lateral jerk to the model improved fit substantially (ΔAIC = 61) for Conversation but marginally for Solo (ΔAIC = 7). Longitudinal motion showed a smaller condition effect (12%).

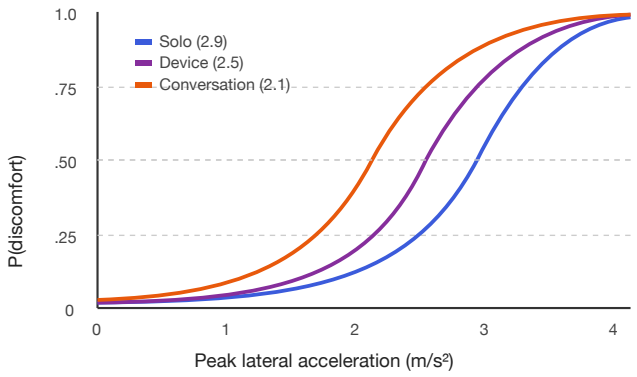


Figure 1. Logistic fits of discomfort probability versus peak lateral acceleration by condition (n = 212, 11,448 manoeuvres). Values in parentheses are the acceleration at P = 0.5. Passengers in conversation are uncomfortable at substantially lower accelerations.

Annotation of the Conversation audio showed that 38% of manoeuvres above 2.0 m/s² coincided with a *self-interruption*: the speaker stopped mid-utterance, restarted, or lost the thread (coded by two annotators, $\kappa = 0.79$). Manoeuvres that happened to fall within a speech pause of ≥ 600 ms produced self-interruptions at a quarter the rate of those during speech, and button presses at roughly half the rate, suggesting that timing is nearly as important as magnitude.

4 Conversation-Aware Planning

4.1 Baseline planner

Zavanet Drive uses a sampling-based planner that evaluates candidate trajectories over a 6 s horizon against a weighted sum of cost terms: progress, safety margin, rule compliance, and comfort. The comfort term in the baseline is $C_{\text{comf}} = w_a \int a_{\text{lat}}^2 + w_j \int j_{\text{lat}}^2 +$ (longitudinal terms), with fixed weights tuned to the Solo thresholds.

4.2 State-dependent comfort cost

We introduce a scalar conversation state $s \in [0, 1]$ (§4.3) and make the comfort weights functions of it: $w_a(s) = w_a^0(1 + \alpha s)$ and $w_j(s) = w_j^0(1 + \beta s)$. From the study thresholds we set $\alpha = 0.9$ and $\beta = 2.4$, reflecting the larger relative importance of jerk during conversation. This has the effect of making the planner prefer earlier, longer, gentler lane changes and curve entries when passengers are talking, and reverting to the baseline when they are not.

4.3 Conversation signal

The signal s is computed on the vehicle's audio front-end from per-seat voice-activity detection (VAD) and seat orientation. VAD outputs a binary speech/non-speech flag per seat at 50 Hz; no spectral content, transcription or speaker

identity is computed, and the raw audio buffer is 200 ms and never persisted. We define s as an exponentially smoothed indicator (time constant 8 s) that at least two occupied seats facing each other have alternated speech activity in the last 20 s. This captures dialogue and excludes a single passenger on a phone call (one seat active) or music (no per-seat alternation). In the study, $s > 0.5$ agreed with human annotation of "in conversation" 93% of the time.

4.4 Pause-aligned timing

When a manoeuvre is required within the horizon and $s > 0.5$, the planner may shift its onset by up to ±1.2 s to align with a predicted speech pause, subject to safety constraints. Pause prediction is deliberately simple: the VAD stream over the last 20 s gives a distribution of inter-utterance gaps, and we predict the next pause onset as the current utterance length approaches the speaker's median utterance length. This is right 61% of the time within ±400 ms, which is enough to be useful; when the prediction is wrong the planner simply executes at the latest safe onset.

5 Evaluation

5.1 Setup

We evaluated conversation-aware planning (CAP) against the baseline in 1,140 rides: 620 on the closed course with 148 new participants in the Conversation condition, and 520 on public roads in Austin with 96 employee-volunteer pairs under a safety-driver protocol. Rides were assigned to planner condition at random and participants were blind to condition. We measured button-press discomfort, post-ride rating, self-interruptions (annotated from consented study audio), travel time, and safety metrics (minimum time-to-collision, hard-brake events, disengagements on public roads).

5.2 Results

Table 1 summarises outcomes. CAP reduced discomfort presses by 31% and improved post-ride comfort rating by 0.8 points on the 7-point scale. Self-interruptions fell by 44%, of which roughly two thirds is attributable to the reduced jerk and one third to pause alignment (from an ablation that disabled timing). Travel time increased by 1.9% on the closed course and 1.6% on public roads, because gentler manoeuvres are begun earlier and resolved more slowly. Safety metrics did not differ measurably: minimum TTC was slightly higher under CAP (earlier manoeuvre onset), and hard-brake events and disengagements were equal within noise.

Table 1. Baseline vs. conversation-aware planning (CAP) across 1,140 rides. Discomfort presses are per 10 km. Self-interruptions are per 10 minutes of conversation. Hard brakes are events with longitudinal deceleration > 4 m/s² per 100 km. Disengagements are public-road only, per 100 km. † indicates $p < .01$ (mixed-effects model with participant random effects).

Metric	Baseline	CAP	Change	Closed course	Public road
Discomfort presses / 10 km	3.9	**2.7**	−31%†	−33%	−28%
Post-ride comfort (1–7)	4.6	**5.4**	+0.8†	+0.9	+0.7
Self-interruptions / 10 min	4.1	**2.3**	−44%†	−46%	−41%
Travel time	—	—	+1.9%†	+1.9%	+1.6%
Peak lateral jerk (m/s³, mean)	1.84	**1.12**	−39%†	−41%	−36%
Min. time-to-collision (s)	3.1	3.3	+0.2	+0.2	+0.1
Hard brakes / 100 km	0.42	0.40	n.s.	n.s.	n.s.
Disengagements / 100 km	0.19	0.19	n.s.	—	n.s.

5.3 Does the planner distinguish conversation?

Because the comfort weights depend on s, CAP should behave like the baseline when passengers are not talking. In 180 additional rides in the Solo and Device conditions, travel time under CAP was within 0.3% of baseline and discomfort was unchanged, confirming that the cost is paid only when it buys something. We regard this as important: a planner that is always gentler would be simpler, but would slow every ride to benefit only the conversational ones.

5.4 Qualitative feedback

Post-ride interviews with 60 participants were coded by theme. Under CAP, participants more often described the ride as "smooth" or said they "forgot they were in a car"; several said they had not realised the vehicle changed lanes. Under baseline, participants described "bracing" and "wait-

ing for it to settle before continuing." Nobody in either condition mentioned travel time unprompted.

6 Discussion

The design decision at the heart of Zavanet Drive, to seat passengers facing each other, turned out to have consequences for the motion planner. We think this generalises: as vehicles stop requiring attention from their occupants, "comfort" will increasingly depend on what the occupants are doing, and planners will need in-cabin state to be comfortable efficiently. The cost of conversation-aware planning is small (under 2% travel time, paid only during conversation), the benefit is large for the thing the vehicle exists to enable, and the sensing required is modest and content-free.

We chose to restrict sensing to activity detection even though a richer signal (speaker emotion, topic) might improve timing further. Zavanet's Human Bonds charter requires that systems not process the content of people's conversations without a reason that benefits them directly, and we did not find one here.

7 Limitations

Participants were recruited in one US city and rode for at most 40 minutes; longer rides and other driving cultures may shift thresholds. Public-road rides had a safety driver, which may affect passenger comfort. The pause predictor is crude and language-agnostic; it was evaluated on English and Spanish conversation only. Self-interruption annotation relied on study-only audio recording and is not available as a production metric. Travel-time cost may be larger in dense traffic where gentle manoeuvres forgo gaps.

8 Conclusion

Passengers in conversation want a different ride, and the vehicle can tell, cheaply and without listening to what they say. A state-dependent comfort cost and pause-aligned manoeuvre timing reduced discomfort by 31% and self-interruptions by 44% for under 2% in travel time. The 11,448-manoeuvre discomfort dataset and the cost formulation are available at drive.zavanet.com/research/cap.

Acknowledgements

We thank the Zavanet Drive safety drivers and the 456 study participants, and the Human Bonds team for reviewing the sensing design.

References

1. ISO 2631-1:1997. Mechanical vibration and shock: Evaluation of human exposure to whole-body vibration.
2. M. Elbanhawi, M. Simic and R. Jazar. In the passenger seat: Investigating ride comfort measures in autonomous cars. *IEEE ITS Magazine*, 7(3), 2015.
3. I. Bae et al. Toward a comfortable driving experience for a self-driving shuttle bus. *Electronics*, 8(9), 2019.
4. H. Bellem et al. Comfort in automated driving: An analysis of preferences for different automated driving styles. *Transportation Research Part F*, 55, 2018.
5. C. Diels and J. E. Bos. Self-driving carsickness. *Applied Ergonomics*, 53, 2016.
6. M. Sivak and B. Schoettle. Motion sickness in self-driving vehicles. *UMTRI-2015-12*, 2015.
7. D. Li and J. Hu. Mitigating motion sickness in automated vehicles with frequency-shaping approach to motion planning. *IEEE RA-L*, 6(4), 2021.
8. M. Werling et al. Optimal trajectory generation for dynamic street scenarios in a Frenét frame. *ICRA*, 2010.
9. W. Xu et al. A real-time motion planner with trajectory optimization for autonomous vehicles. *ICRA*, 2012.
10. A. Bellet et al. Privacy and in-cabin monitoring in automated vehicles. *Transportation Research Interdisciplinary Perspectives*, 12, 2021.