DAWZY: Human-in-the-Loop Natural-Language Control of REAPER

1 Description

We demo DAWZY, a voice- and text-driven assistant that translates natural-language requests into reversible, state-aware actions in REAPER, a scriptable Digital Audio Workstation (DAW). DAWZY shortens the gap between musical intent (e.g., ‘warm the vocals’) and precise DAW edits, keeping the DAW as the creative hub. The system queries live session states, explains each change, and supports rapid beat prototyping with an AI beat generation model and a Hum-to-MIDI model. *DAWZY* uses LLM-based code generation as a novel way to significantly reduce the time users spend familiarizing themselves with large interfaces. While related efforts (e.g., Mozart AI [Mozart AI, 2025]) explore closed-source adjacent ideas, DAWZY emphasizes open-source availability and ReaScript-specific reliability. The video highlights reliability on common production tasks. [1]

Modern music production centers on Digital Audio Workstations (DAWs) Leider [2004], which democratize pro-quality creation but burden users with option overload that disrupts flow [Kjus, 2024]. A gap persists between high-level intent (e.g., “make the vocals warmer”) and the low-level steps to realize it.

Primary Contributions

- **System design & open-source prototype.** REAPER-targeted pipeline mapping natural language to safe, reversible ReaScript grounded in live state (Sec. 2).
- **Minimal-GUI, voice-first interaction.** Natural-language control with buttons for common tasks (“start,” “stop,” “record,” “undo”) to reduce GUI micromanagement.
- **Explain-as-you-go pedagogy.** Plain-language rationales accompany each edit to support learning and auditability.

2 DAWZY Architecture

- A minimal **Electron.js** GUI [OpenJS Foundation, 2024] forwards text, speech, and humming to the processing layer.
- **OpenAI GPT-5** [OpenAI, 2025a,b] interprets the user’s intent, calls MCP tools, and emits Lua ReaScript.
- **Model Context Protocol (MCP)** exposes the following functions to the LLM:
  - **State query.** Provides live REAPER context including tracks, items, FX, and routing.
  - **FX parameterization (`fxparam`).** Converts human units (dB, ms) to ReaScript slider ranges (e.g., 0–1, 0–4) to prevent scaling errors.
  - **Beat generation (1:54).** Meta’s **MusicGen-small (300M)** model is run locally to create an audio waveform based on a text description [Meta AI, Copet et al., 2023].
- **Hum-to-MIDI (2:41):** The open-source **Spotify BasicPitch** model is run locally to convert hums into MIDI data [Spotify, Bittner et al., 2022].

[1] Code Demo

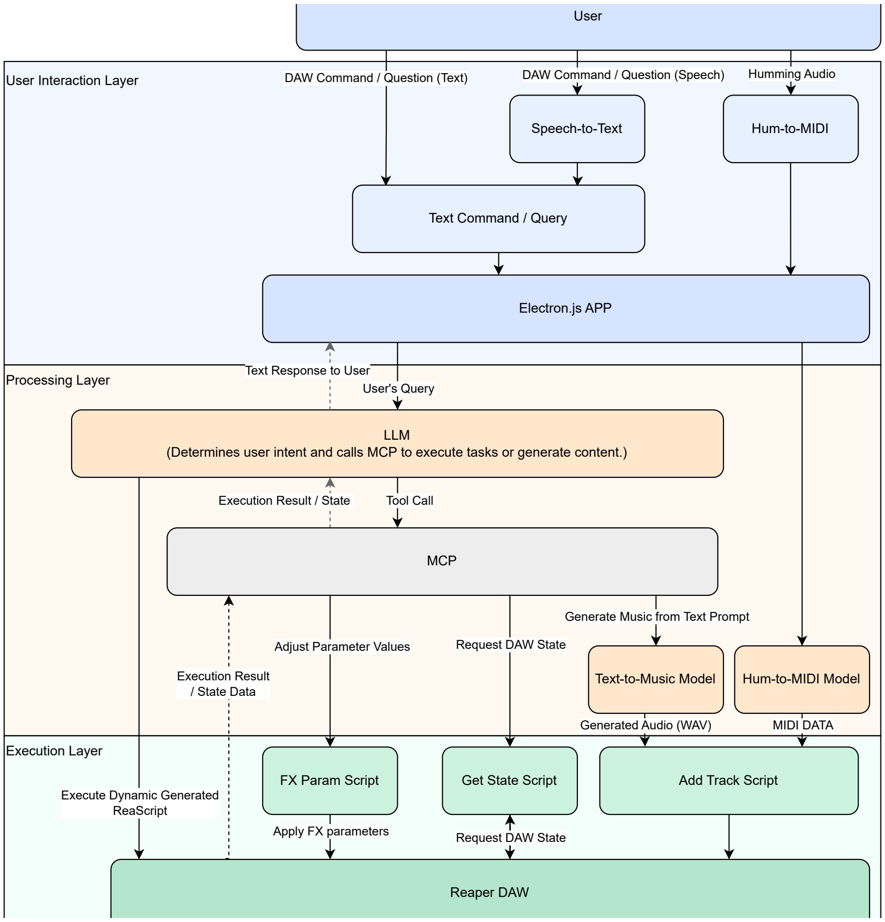

Figure 1: **DAWZY Architecture.** User intent (text/speech/hum) flows through the Electron gateway to the LLM and MCP tools, then executes as reversible ReaScripts in REAPER. Rounded rectangles denote AI/MCP components; sharp rectangles denote DAW/runtime components; dashed arrows indicate data queries; solid arrows indicate state-changing actions.

3 Evaluation

To assess our model's reliability, we tested four reproducible tasks during the demo video: (1) **Multi-instruction FX processing** — "Double the first track's volume, increase the decay, and set the attack to 10 ms," (2) **GUI navigation** — "Open the FX browser for the first track," (3) **Workflow automation** — "Duplicate the first track, pitch it up one octave, and blend it in at 20%," and (4) **Educational interaction** — "What does attack time do in the second track's compressor?"

We conducted a **Mean Opinion Score (MOS)** test with *21* participants, who rated DAWZY's *Enjoyment* as 4.48, *Learning* as 4.38, *Collaboration* as 4.29, *Usability* as 4.14, and *Control* as 3.81 out of 5.

4 Conclusion

DAWZY packages state-grounded scripting, reversible edits, and explain-as-you-go guidance into a cohesive, human-in-the-loop workflow that fits naturally into existing creative practice. We believe this work opens promising directions for AI-assisted creativity across domains where scriptable interfaces can enable sophisticated automation while maintaining precision and reliability.

References

Mozart AI. Mozart ai — ai-powered music production daw, 2025. URL `https://getmozart.ai/`. Product site; Accessed 2025-08-18.

Colby N. Leider. *Digital Audio Workstation*. McGraw-Hill, Inc., USA, 1 edition, 2004. ISBN 0071422862.

Yngvar Kjus. The platformization of music production: How digital audio workstations are turned into platforms of labor market relations. *New Media & Society*, 2024. doi: 10.1177/14614448241304660. First published online December 11, 2024.

OpenJS Foundation. Electron. `https://www.electronjs.org/`, 2024. Accessed 18 Aug 2025.

OpenAI. Introducing gpt-5. `https://openai.com/index/introducing-gpt-5/`, 2025a. Accessed 18 Aug 2025.

OpenAI. Gpt-5 system card. `https://openai.com/index/gpt-5-system-card/`, 2025b. Accessed 18 Aug 2025.

Meta AI. facebook/musicgen-small. `https://huggingface.co/facebook/musicgen-small`. Hugging Face model card; accessed 18 Aug 2025.

Jade Copet, Felix Kreuk, Itai Gat, Tal Remez, David Kant, Gabriel Synnaeve, Yossi Adi, and Alexandre Défossez. Simple and controllable music generation, 2023.

Spotify. Basic pitch — about. `https://basicpitch.spotify.com/about`. Product page; accessed 18 Aug 2025.

Rachel Bittner, Min Kim, Juan José Bosch, Aren Jansen, Gordon Wichern, and Longshaop Hantrakul. A lightweight polyphonic pitch transcription model. *arXiv preprint arXiv:2203.09893*, 2022. URL `https://arxiv.org/abs/2203.09893`.