

SYNESTHESIA MACHINE.

A practical field guide to translating words, beats and sound into visible compositions.

CREATED BY SAHIL TAYAL • A SAHILVERSE EXPERIMENT

This booklet accompanies synesthesiamachine.sahiltayal.online. It is educational and exploratory; it is not a clinical assessment, scientific instrument or account of what synesthesia is like for any particular person. Its colours and shapes are design decisions, not innate properties of words or frequencies.

01 / The question underneath the spectacle

Human perception combines signals across modalities. A drumbeat can seem “heavy”; a voice can seem “bright”; a sentence can have “sharp” edges. Some of these descriptions are conventions, some may be influenced by shared perceptual tendencies, and some are personal associations. A translation machine makes a specific creative choice explicit: it chooses a function that maps measurable input features to visible properties.

Synesthesia, by contrast, refers to experiences in which a stimulus can elicit an additional concurrent experience or association. Forms and individual experiences differ. An arbitrary web mapping is not a simulation or detection of it.

02 / What the machine actually measures

Text: the machine computes a deterministic 32-bit fingerprint by iterating over Unicode code points. Word count and punctuation influence energy; character variety influences visual complexity. A fingerprint determines a hue starting point. These are reproducible, illustrative design features, not sentiment analysis.

Rhythm: four input pads collect timestamps and pad identities for a recent window. The number of recent strikes changes energy. Interval variation and variety of pad identities change complexity. A deterministic mapping from pad selections creates hue. Timing is captured live, so identical human performances may differ.

Microphone: only after permission, the local Web Audio API analyses a small frequency spectrum and waveform. Root-mean-square signal level is used as a rough loudness proxy; a magnitude-weighted spectrum position approximates brightness. It never interprets words or

classifies a person's emotions. Signal quality depends on microphone hardware, ambient noise, gain and browser.

Output: a Canvas visualisation uses layered paths and bounded dots. An optional synthesised sine tone is played only following audio opt-in and user interaction. No remote API, tracking or storage service is necessary.

03 / Three different mapping grammars

PRISM: angular, faceted contours and bright accents. Energy changes the composition's scale; complexity changes the number of layers, with colour offset by layer. It often reads as precise, brittle or excited. Those readings are interpretations, not inherent meaning.

TIDE: a cooler hue shift, continuous curved bands and sinusoidal variation around a loop. The same input can seem gentle or fluid, although no semantic analysis has taken place.

ORBIT: a warmer hue shift, nested rings with small radial oscillations. Visitors may perceive order, planetary motion or repetition. Here, the style is a visual convention.

04 / Three guided experiments

EXPERIMENT A — THE PUNCTUATION BUTTERFLY

1. Set Text + Prism. Compose "We are here." and pin the result. 2. Replace the period with "!"; compose again. 3. Observe any shift in hue, energy and detail. 4. Ask: what changed in the code, and what changed in your reading of the sentence?

Interpretation: the output changes because punctuation participates in the fingerprint and energy calculation. The feeling of excitement may come from conventional language reading, not from a naturally red or jagged exclamation mark.

EXPERIMENT B — ONE INPUT, THREE WORLDS

1. Enter "A small signal travels far." 2. Pin Prism. 3. Choose Tide, then Orbit without altering the text. 4. Compare the visual captions as well as the images. 5. Write two contradictory emotional interpretations of the same source.

Interpretation: different renderers make the same features appear emotionally distinct. The experiment exposes the agency of the designer rather than proving that the signal contains a unique visual truth.

EXPERIMENT C — THE SPACE BETWEEN BEATS

1. Select Rhythm; press 1 four times at a roughly constant tempo. 2. Observe the output while the beat window is active. 3. Clear; tap a syncopated pattern with two pad identities. 4. Compare energy and complexity. 5. Discuss whether the irregularity or the pad mixture changed your perception more.

Interpretation: the machine computes a local proxy for interval variation. It does not measure musical quality, groove, cultural background, or human emotion.

05 / Field journal

Record the phrase or rhythm: _____

Observed input features (not interpretation): _____

My first association: _____

An alternative explanation or opposite association: _____

06 / What can—and cannot—be claimed

Can claim: specific encoded text or captured signal features drive specific browser-rendered parameters; changing the preset changes the mapping; typed input is reproducible for a given version and settings; file exports represent the current recipe or rendered state.

Cannot claim: that a sound has an objectively correct colour; that the device diagnoses synesthesia; that any image recreates another person's perception; that frequency-centroid approximations reveal mood, identity or mental health; or that all observers attach the same associations to the shapes.

Many cross-modal correspondences are experimentally studied—for example, associations between pitch and spatial height or between sounds and shapes. Context, culture, individual variation and task design matter. This interface is not an empirical validation of those findings. The sound-shape “bouba/kiki” literature is a useful starting point but should not be treated as universal law.

07 / Privacy, safety and accessible use

The website is a static application: there is no login, backend or analytics endpoint. Microphone permission is explicitly requested by Start microphone and can be revoked at any time by Stop or browser controls. Raw audio is not stored. Avoid speaking private information around other services or extensions you do not trust. Microphone permission usually requires HTTPS or localhost.

Sound is off until explicitly enabled. Set volume to zero or mute for a silent experience. Pausing motion freezes the animation; browsers with reduced-motion preference begin paused. Every visual has a written description and numeric features; rhythm pad buttons also work with number keys 1–4 when the rhythm panel is active. Export SVG or PNG for a still image; JSON preserves numerical choices, not raw microphone data.

08 / Further reading: questions worth following

Cytowic, R. E. & Eagleman, D. M. (2009). *Wednesday Is Indigo Blue: Discovering the Brain of Synesthesia*. MIT Press. An accessible overview of varieties of synesthetic experience.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995. A research-level review of how features across sensory domains become associated.

Ramachandran, V. S. & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3–34. A historical theoretical discussion; revisit claims critically.

Parise, C. V. & Spence, C. (2013). Audiovisual crossmodal correspondences in the general population. Oxford Handbook of Synesthesia. Further context on tendencies without assuming they apply uniformly.

A closing question: When you describe a melody as purple or a sentence as jagged, how much comes from the stimulus, how much from personal history, and how much from the vocabulary available to you?

SAHIL TAYAL · SAHILVERSE · <https://sahiltayal.com/> · sahil@sahiltayal.com