



The Un-Englishable

JUNE 2026

9 min read • 1,983 words

Everyone has lost a dream to language. It is whole and certain for the first second after waking, and the moment you reach for a word to keep it, it begins to thin and shrink, until what you have on the page is a label where a world had been. The label is accurate. It is also almost nothing.

I have spent a lot of my life trusting words, and the thing I most want to keep is the thing words keep failing to hold. For years I thought that was my failure as a writer. Lately I think it is mostly a property of the language I think in, and the culture that raised me to think in it, and I think I can finally say why.

Language is the technology we have for moving an inner state from one mind to another, and the move is [always approximate](#). Some of the signal makes it across. Most of it does not. [I have written about the visitors who arrive in that dropped part](#), the presences that reach me as figures and never as sentences. This essay is the other half of that question. Not what they are, but why language was never going to hold them, or hold a great many things far more ordinary than them.

Naming is the encode

I name things for a living. I have written before about [how hard that is](#), how a good name can carry an entire idea and a bad one can quietly poison a codebase for a decade. What I did not see until recently is what naming actually is.

A name is not a description. It is a compression. To call a feeling grief is to take a state that is entirely your own, textured with a specific person and a specific Tuesday and the exact quality of the light, and round it to the nearest bucket other people already keep. The bucket travels. Almost everything that made the feeling yours does not.

The loss happens on the way in, which is the part that surprised me.

Lossy compression is a metaphor here, not a mechanism. Human language is generative and context-rich in ways a fixed codec is not, and a good sentence can do things a JPEG cannot. What the figure catches is narrower and true: naming discards most of the detail of an experience on purpose, to make a sharable token of it.

We tend to imagine the loss arrives later, in clumsy phrasing or a tired metaphor. It arrives at the first step, the instant the experience becomes a word at all. You encode, and most of the file never enters the stream.

It is also why the right name feels like finding rather than making. When you finally name a thing correctly it lands like recovery, like you pulled something back that the encoding had nearly dropped. A good name throws away a little less.

Other tongues have more bands

Sometimes you meet a word in another language and a state you have carried wordless for years suddenly has a shape. The Portuguese have *saudade*, the presence of an absence, the ache and the tenderness arriving as one feeling. The Germans have *Sehnsucht*, a longing that aches toward something it cannot name, a homesickness for a place you have never been. The Japanese have *mono no aware*, the soft grief of things precisely because they pass, the whole of autumn folded into a single falling petal. The Welsh have *hiraeth*, the pull toward a home you cannot return to, or that never existed.

Finding one of these is a small bandwidth event. Another language's codebook held an entry yours was missing, and a thing you could only gesture at compresses, at last, into a token you can send.

I want to be careful here, because the untranslatable-word genre is a little bit of a con.

Note the obvious tell: I just translated all four of them, a clause each. They are not untranslatable. They are un-compressible into a single English word, which is a smaller and more honest claim.

None of those words are actually untranslatable. I just translated them. And an English speaker feels *saudade* perfectly well without the label, which means the missing word was never a missing feeling. The codebook is a codec, not a cage. But here is the part that holds. Gather every language's extra entries, sum every codebook ever built, and you still run out of words long before you run out of world. The vocabulary is finite. The thing it is reaching for is not.

The thing no codebook covers

If the smaller the shared vocabulary the more gets dropped, then out at the far edge there is a thing no vocabulary covers at all. No word is short enough to send it through. And the traditions that took that thing most seriously all hit the same wall from the inside, and all answered it in nearly the same way. They stopped trying to encode it and started pointing instead.

The negative theologians said you may only state what God is not, clearing away each false name until what remains cannot be named.

The *via negativa*, in Pseudo-Dionysius, Meister Eckhart, and the anonymous *Cloud of Unknowing*. It is a method of pointing past concepts, not a claim that the divine is literally nothing.

The Upanishads taught *neti neti*, not this, not this, the sage refusing every description as it arrived. The *Tao Te Ching* concedes the whole problem in its opening line: the Tao that can be spoken is not the eternal Tao. Zen kept the image of a finger pointing at the moon, and the standing warning not to mistake the finger for the light. And Wittgenstein, of all people, ended his most rigorous book by saying that whereof one cannot speak, thereof one must be silent. People read that as a dismissal of mysticism. He meant the reverse. The unsayable was, to him, the part that mattered most, the part that does not get said and shows itself instead.

These are people reporting the same gap I keep falling into when I try to write down a dream. The veil the older traditions call maya, [the partial render the senses hand you](#), is the same fact from the other side. The surface is real, and it is not the whole, and the codec was never going to fit the source.

The codecs that carry more

Words are not the only channel, and they are nowhere near the widest. The wordless is not only what we cannot say. A great deal of it is just what something else says better.

Music is the one I know from the inside. I came up as a percussionist and taught myself harmony late, and I have spent years building [a library to model music in code](#). A song can carry grief without once using the word, can move a room of strangers to the same ache with no proposition in it anywhere. The notation is not the music. It is a lossy transcript of it, useful on the page and silent on the floor, where the body already knows what to do [before the mind has any words for it](#).

Even music's own written codec drops signal. Standard notation rounds the whole continuous field of pitch into twelve equal steps, and much of what it rounds away is exactly the part that carries the feeling. [A sitar player rounds her flat third to the nearest white key and loses the precise thing that made it hers](#). I built the library partly to refuse that rounding, to keep the shrutis and the quarter-tones the default codec throws out. Image works the same way in another band, which is why a symbol can hold more than the paragraph explaining it, and so can a dream. Each is a channel with room for things the verbal one has no slot for.

The machine whose first language has no words

Here is the part I cannot stop turning over.

We built a machine that is extraordinary with language, and it does not think in words. A large language model takes your sentence and immediately turns it into vectors, long lists of numbers, points in a space with thousands of dimensions.

Everything that happens next is geometry. Meaning, to whatever degree the thing has anything like meaning, lives there as position and direction, as nearness and distance among points that carry no names.

I am being loose on purpose and trying not to be wrong. The precise terms are embeddings and per-layer hidden states. Interpretability finds real but partial and distributed structure in there, not a tidy labeled map, and the model represents structure, it does not understand or experience anything. The old word2vec trick, king minus man plus woman lands near queen, is the famous hint that the space encodes relations no single word holds. Modern models are far stranger than that.

Words only return at the very end. The final point is collapsed into a spread of probabilities across the whole vocabulary, and one token is chosen, and then the next.

That last step is a lossy decode from a wordless internal state into language. It is the exact boundary I cross every morning, when the whole thing in my hands has to become a sentence. We taught a machine to think in the medium that has no words, and then we built ourselves one small door, the text box, and we only ever speak to it through the door. The English it hands back, the part we see, is the most compressed and averaged surface of it, the thinnest layer, [an average of nearly everything we have already said](#). The interior is wider than the doorway. I cannot tell, from here, whether that makes it more like us or less, and I have stopped needing to decide. It does mean that when you name something to it, you are not labeling a row in a table. You are [tilting the whole landscape toward the region where that name lives](#), calling a pattern forward out of a space with no words in it, which is more or less what naming has always done.

For humans

I keep landing on one consequence, because it is the one that costs people.

A culture that trusts only what compresses into words will treat whatever does not as either unreal or unwell. If a feeling will not fit the shared codebook, the verbal world tends to read the missing word as a missing thing, and to hand you a diagnosis where a vocabulary should have been. But the loss was in the channel. It was never an absence in you. Insisting that an inner state be put into

words before it is allowed to count is like insisting that a song be reduced to notation before you will admit it is music. The notation is genuinely useful. It is just not the song.

This is the same instinct I have followed since I started writing software for humans instead of for machines: the tool should bend to the shape of the person, including the shape of the parts that do not run through words. I am not going to tell you to go sit with the wordless, and I am not going to pretend it is all sacred either. Some of what arrives without words is revelation and some of it is noise, and the wonder and the danger run on the same wire. But the wordless is not the enemy of the real. Most of the time it is simply the rest of the real, the part that did not fit through the door.

I name things so a little of them can travel. Most of it never leaves, and that, it turns out, is where I have been living the whole time.