

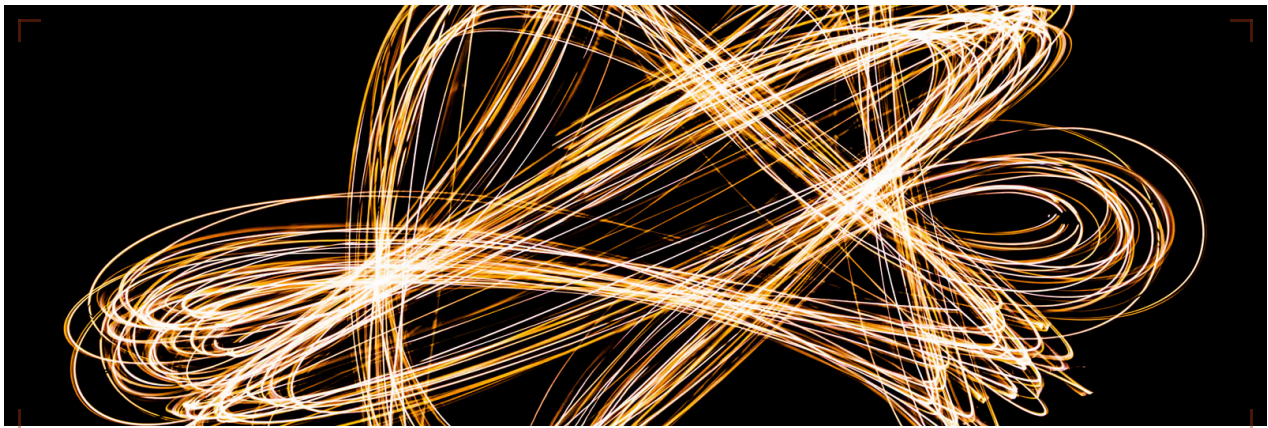
# Tipjai Voice 1: the full model card

Tipjai's first Thai tip-alert voice: what it is for, how it was measured, results, limits and responsible use. 6.0% heard wrong, 8th of 30, ฿0.

Published  
14 Oct 2026

Authors  
Tipjai Research

PDF  
Download



I. Lissajous figures, the curves two overlapping oscillations draw, painted with light.  
Dietmar Rabich / Wikimedia Commons / "Lissajous-Figur -- 2020 -- 7770" / CC BY-SA 4.0, resized Licence

## 1 Summary

Heard wrong

6.0%

8th of 30 voices

Voices clearly clearer

0

of 29

Per alert

฿0

on the streamer's computer

Time per alert

1.38<sub>s</sub>

laptop CPU

- Heard wrong on **6.0%** of characters across 60 tip alerts no voice had seen; 95% interval 4.6% to 7.5%. None of the 29 other voices was clearly clearer, and 22 were too close to call.
- It loses most clearly on **supporter names** (26th of 30) and on **naturalness** (3.17 of 5 against 5.00 for the best voice, one listener).
- Known issue: an amount of 555 baht is read as laughter (found 14 October 2026; to be fixed in the next version).

**For creators** It reads tip alerts about as clearly as the best paid voices on the market, at no cost, but it still gets some names wrong and it still sounds synthetic.

## 2 Overview

|              |                                                                                  |
|--------------|----------------------------------------------------------------------------------|
| Version      | Tipjai Voice 1 (internally jai-pulse-1), frozen 6 October 2026 at 14:43          |
| What it does | Reads Thai tip alerts aloud on a live stream: supporter name, amount and message |
| Language     | Thai, including names written in Latin letters                                   |
| Status       | Lab preview                                                                      |
| Runs on      | The streamer's own computer, no graphics card needed                             |
| Price        | ฿0 per alert                                                                     |

It builds on open-source work, with every licence honoured. No vendor's audio was ever used to train or tune it, and no real supporter message was used.

What we do not publish: how the voice is built, meaning its components, data, sizes and settings. Outcomes and methods are public so anyone can check them; the recipe stays in-house so the work cannot simply be copied.

## 3 Intended use and what it is not for

- ✓ Reading tip alerts on a live stream
- ✓ Thai and Latin-script names, numbers, amounts, 555
- ✗ Imitating a real person
- ✗ Long reading, or other languages
- ✗ Passing synthetic speech off as a person

## 4 How it was measured

Measured on 60 tip alerts written apart and frozen before scoring, against 29 voices taken from each vendor's own service. 2 Thai speech recognizers listened to every clip; the score is the share of characters heard wrong; intervals come from 2,000 resamples. Naturalness comes from one listener who did not know whose voice was whose.

The full method is in the [paper, with the test set to download \(/en/research/thai-tip-alert-benchmark/\)](/en/research/thai-tip-alert-benchmark/)

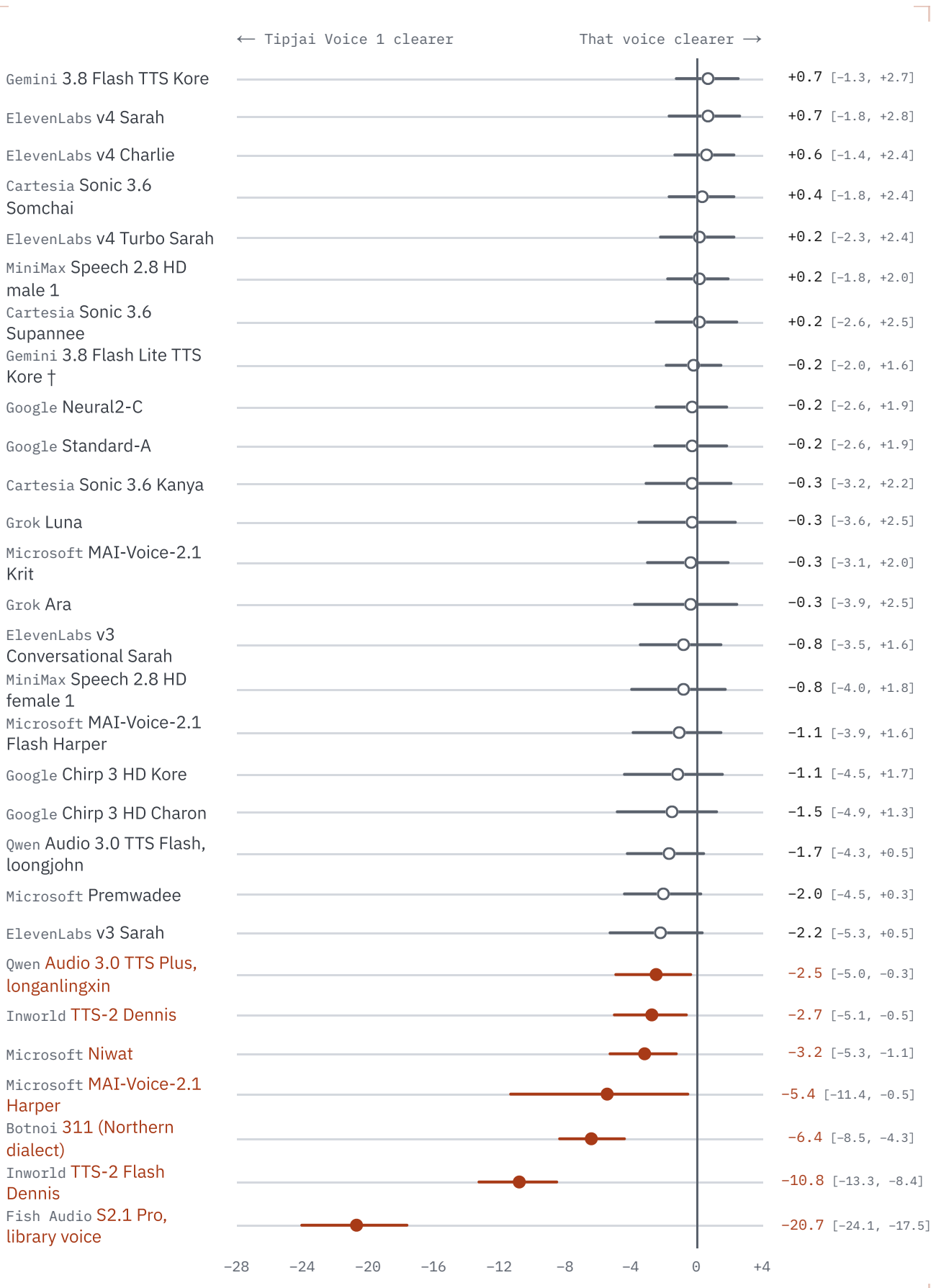
## 5 Results

|                          | Tipjai Voice 1<br>ours | Gemini 3.8 Flash TTS Kore<br>clearest on the board | Google Standard-A<br>cheapest paid voice |
|--------------------------|------------------------|----------------------------------------------------|------------------------------------------|
| Heard wrong              | 6.0%                   | <b>5.3%</b>                                        | 6.3%                                     |
| 95% interval             | 4.6% to 7.5%           | 3.4% to 7.7%                                       | 3.9% to 8.8%                             |
| Rank range               | 1 to 22                | 1 to 14                                            | 3 to 18                                  |
| Name heard right (1 / 2) | 29 / 30                | <b>41 / 43</b>                                     | 39 / 40                                  |
| Whole alert right        | 16 / 60                | <b>25 / 60</b>                                     | 24 / 60                                  |
| Naturalness (1 to 5)     | 3.17                   | -                                                  | -                                        |
| Per 1,000 alerts         | <b>฿0</b>              | ฿35                                                | ฿6.7                                     |

Bold is the best value in its row, whoever owns it

Figure 1

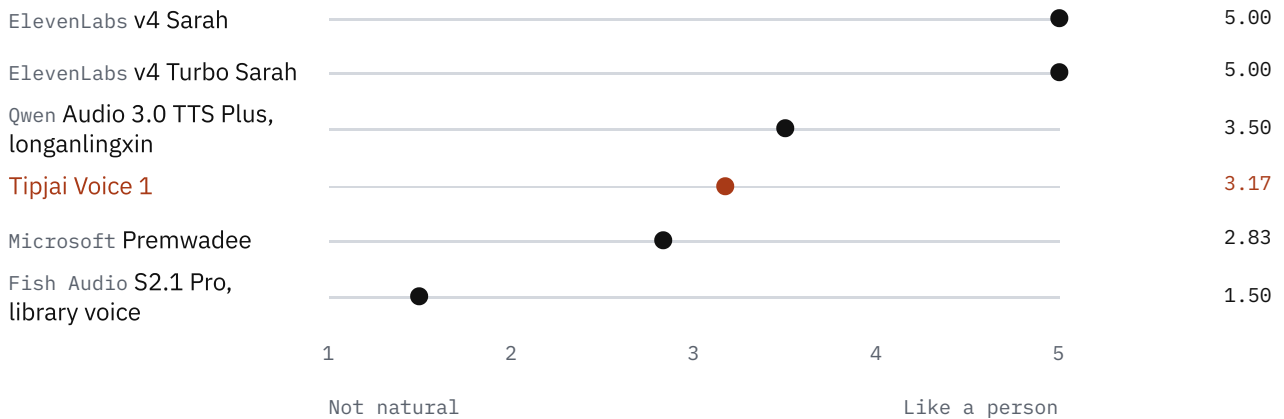
Tipjai Voice 1 is clearly clearer than 7 voices; none is clearly clearer than it; 22 are too close to call.



Each dot is the difference in errors (Tipjai Voice 1 minus that voice, in percentage points) on the same alerts; each line is its 95% interval.

Figure 2

## Naturalness: the next version's main job



1 listener, blind, 6 alerts. A small sample: a direction, not a ranking.

## 6 Speed and stability

On a laptop CPU it takes 1.38 s per alert (median of 20) and reads up to 76 alerts a minute in parallel. The same text gives the same audio every time.

Figure 3

## 400 stress inputs: no crash, no silence, no clipping



The longest names and messages Tipjai accepts, emoji, links, other writing systems, numbers in every shape, invisible characters, HTML and stretched letters.

---

## 7 Safety and misuse

|                      |                                                                                       |
|----------------------|---------------------------------------------------------------------------------------|
| <b>Impersonation</b> | Tipjai Voice 1 is one synthetic voice. It cannot be set to sound like a given person. |
|----------------------|---------------------------------------------------------------------------------------|

---

|                   |                                                                                    |
|-------------------|------------------------------------------------------------------------------------|
| <b>Disclosure</b> | It is synthetic, not a recording of a person, and must not be used to pass as one. |
|-------------------|------------------------------------------------------------------------------------|

---

|                     |                                                                                                |
|---------------------|------------------------------------------------------------------------------------------------|
| <b>Abusive text</b> | The voice reads what it is given; Tipjai's alert system masks rude words before they reach it. |
|---------------------|------------------------------------------------------------------------------------------------|

---

|                         |                                                                                                             |
|-------------------------|-------------------------------------------------------------------------------------------------------------|
| <b>Report a problem</b> | A misreading or a misuse can be reported at <a href="https://tipjai.com/feedback">tipjai.com/feedback</a> . |
|-------------------------|-------------------------------------------------------------------------------------------------------------|

---

## 8 Privacy

No real supporter message is in the data used to build or test this voice. Every test alert was written for testing.

---

## 9 Known limitations

|              |                                                        |
|--------------|--------------------------------------------------------|
| <b>Names</b> | Fewer names heard right than most voices (26th of 30). |
|--------------|--------------------------------------------------------|

---

|                    |                                                                       |
|--------------------|-----------------------------------------------------------------------|
| <b>Naturalness</b> | It still sounds synthetic, and naturalness was rated by one listener. |
|--------------------|-----------------------------------------------------------------------|

---

|                 |                                                                                                                       |
|-----------------|-----------------------------------------------------------------------------------------------------------------------|
| <b>555 baht</b> | An amount of 555 baht is read as laughter. Found 14 October 2026; the fix will be published with a fresh measurement. |
|-----------------|-----------------------------------------------------------------------------------------------------------------------|

---

|                      |                                          |
|----------------------|------------------------------------------|
| <b>The top group</b> | 60 alerts cannot separate the top group. |
|----------------------|------------------------------------------|

---

|                    |                                                                                  |
|--------------------|----------------------------------------------------------------------------------|
| <b>Recognizers</b> | The two disagree by up to 10 places, so the score estimates what a person hears. |
|--------------------|----------------------------------------------------------------------------------|

---

10

# Versions and changelog

|             |                                                                                   |
|-------------|-----------------------------------------------------------------------------------|
| 14 Oct 2026 | Issue found: an amount of 555 baht is read as laughter. Added to the limitations. |
| 14 Oct 2026 | Naturalness round 2 added (1 listener, 6 alerts).                                 |
| 6 Oct 2026  | A new way of reading was tried on blind set 3: worse, so not used.                |
| 6 Oct 2026  | Same text, same audio every time; long messages read in full.                     |
| 6 Oct 2026  | jai-pulse-1 frozen at 14:43. Every number on this card was measured on it.        |

```
@misc{tipjai2026voice1card,
  title = {Tipjai Voice 1: model card},
  author = {{Tipjai Research}},
  year  = {2026},
  month = oct,
  url   = {https://research.tipjai.com/en/models/tipjai-voice-1/card/}
}
```