COST ANALYSIS

On average, **1 token ≈ 4 characters of English text**

Example Reference:

- **The word `"ChatGPT"` = 1 token**
- **`"OpenAI is great!"` = 4 tokens**
- **`"The quick brown fox."` = 5 tokens**

GEMINI MODELS

- **COST OF Gemini 2.0 Flash Exp:**
    - Price for token
        - **Input: $0.07 / 1M tokens**
        - **Output: $0.30 / 1M tokens**
    - Average Response (200 tokens)
        - **Input: $0.000014**
        - **Output: $0.00006**

The pricing of text and audio token is the same for the experimental version

---

- **COST OF Gemini 2.0 Flash Live (001):**
    - Text Tokens
        - Price for token
            - **Input: $0.50 / M Tokens**
            - **Output: $2 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.0001**
            - **Output: $0.0004**
    - Audio Tokens
        - Price for token
            - **Input: $3 / M Tokens**
            - **Output: $12 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.0006**
            - **Output: $0.0024**

Both models are very similar, the only difference is that the first one is experimental (its a combination of some gemini 2.0 models).

OPENAI MODELS

- **COST OF GPT-4o Realtime:**
    - Text Tokens
        - Price for token
            - **Input: $5.00 / M Tokens**
            - **Output: $20 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.001**
            - **Output: $0.004**
    - Audio Tokens
        - Price for token
            - **Input: $40 / M Tokens**
            - **Output: $80 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.008**
            - **Output: $0.016**

- **COST OF GPT-4o mini Realtime:**
    - Text Tokens
        - Price for token
            - **Input: $0.6 / M Tokens**
            - **Output: $2.4 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.00012**
            - **Output: $0.00048**
    - Audio Tokens
        - Price for token
            - **Input: $10 / M Tokens**
            - **Output: $20 / M Tokens**
        - Average Response (200 tokens)
            - **Input: $0.002**
            - **Output: $0.004**

The Gemini models are much more affordable and the performance is not much different between Gemini and Openai, so the best option is either of the Gemini models.