

■ // GPT-LIVE · CHATGPT VOICE

Hören und sprechen gleichzeitig

GPT-Live hat ChatGPTs Sprachmodus im Juli 2026 ersetzt. Statt abzuwarten, bis du fertig bist, entscheidet das Modell laufend neu, was als Nächstes passiert. Dieser Überblick zeigt, wie du es einschaltest, welche Stufe auf welchem Plan läuft, was Live zwischen zwei Gesprächen behält und wo es heute noch hakt.

FULL DUPLEX

GPT-5.5 IM HINTERGRUND

PLAN-GRENZEN

GEDÄCHTNIS

QUELLEN INKLUSIVE

Worum es geht: OpenAI hat GPT-Live am 08.07.2026 vorgestellt und damit den Advanced Voice Mode abgelöst. Der Unterschied steckt in einem Satz aus der Ankündigung: Das Modell behandelt ein Gespräch nicht mehr als Folge getrennter Nachrichten, sondern **entscheidet fortlaufend neu, ob es spricht, weiter zuhört, pausiert, unterbricht oder ein Werkzeug aufruft**. Über Voice und Dictation reden nach OpenAIs Angabe jede Woche mehr als 150 Millionen Menschen mit ChatGPT.

1 Einschalten und umschalten

Der Einstieg ist ein Symbol, keine Einstellung: Tippe in der Eingabezeile auf das **Sprachsymbol** und erlaube den Mikrofonzugriff, wenn du gefragt wirst. Das war es für den ersten Start.

Zwischen den Modi wechseln

```
// in ChatGPT  
Einstellungen → Sprache → Live | Advanced | Standard
```

Die drei Optionen stehen nebeneinander, du bist also nicht auf Live festgelegt. Wenn dich das Unterbrechen stört oder eine Verbindung wackelig ist, bringt der Rückweg auf **Advanced** die alte, abwartende Gesprächsführung zurück.

Was während des Gesprächs geht

- Dich selbst **stummschalten**
- ChatGPT **unterbrechen**, mitten im Satz
- Eine Nachricht **tippen**, ohne das Gespräch zu verlassen
- Ein **Bild anhängen**, wo das für dein Konto verfügbar ist
- Danach den entstandenen **Textverlauf nachlesen**

2 Welche Stufe auf welchem Plan

Es gibt zwei Modelle und darüber drei Intelligenzstufen. Das ist der Teil, der in den meisten Zusammenfassungen fehlt, und er entscheidet, wie gut Live bei schweren Fragen wirklich ist.

```
// Modelle
GPT-Live-1      volle Version, ab dem Go-Plan
GPT-Live-1 mini  auf Free, begrenzte Zeit pro 24 Stunden

// Intelligenzstufen und was dahinter rechnet
Instant         GPT-5.5 Instant
Medium          GPT-5.5 Thinking, mittlerer Denkaufwand
High            GPT-5.5 Thinking, hoher Denkaufwand
```

Die Zeitkontingente laufen als rollende Fenster, nicht als Monatspaket. Go und Plus nennen aktuell je eine rollende Stunde pro Intelligenzstufe und zwei rollende Stunden für **mini**; Business nennt bis zu einer Stunde Instant plus einer Stunde Medium oder High. Ein einzelnes Live-Gespräch ist auf **zwei Stunden** begrenzt.

Praktische Folge: Wenn du Live für eine echte Arbeitssitzung nutzen willst, plane die schwere Recherche bewusst ein. Die Instant-Stufe reicht für Gespräche, aber die Denkstufen sind das knappe Gut.

3 Die Übergabe an GPT-5.5

Das ist die Funktion, die den Unterschied im Alltag macht, und sie läuft unsichtbar. Kommt eine Frage, die Suche oder echtes Nachdenken braucht, gibt Live sie im Hintergrund an GPT-5.5 ab und hält das Gespräch vorne weiter am Laufen. Du hörst keine Wartepause und keine Stille.

Dazu kann Live während der Unterhaltung **Websuche und Gedächtnis** nutzen und Ergebnisse als sichtbare Karten einblenden — etwa zu Wetter, Aktien oder Sport. Text und Bilder liegen im selben Chat.

4 Was Live sich merkt — und was nicht

Hier liegt das größte Missverständnis. Innerhalb eines Gesprächs führt Live ein Transkript, und du kannst dort später wieder anknüpfen. Über Gespräche hinweg greift dagegen **nur ChatGPTs allgemeines Gedächtnis**.

Und dieses Gedächtnis speichert **zusammengefasste Fakten über dich**, nicht deine Sätze. Was du gesagt hast, ist im nächsten Gespräch weg; was daraus als Fakt abgeleitet wurde, bleibt.

Für den Alltag heißt das: Wichtiges, das weiterleben soll, gehört ausdrücklich benannt, damit es als Fakt landet. Wer erwartet, dass Live sich an den Wortlaut von vorgestern erinnert, wird enttäuscht.

5 Wasserzeichen und Nachweis

Von OpenAI erzeugte Audios tragen inzwischen ein **SynthID**-Wasserzeichen, entwickelt von Google DeepMind. Das öffentliche Prüfwerkzeug von OpenAI deckt damit neben Bildern auch unterstützte Audiodateien ab.

Was es beantwortet, ist eine einzige Frage: ob eine Datei von OpenAI erzeugt wurde. Es sagt nichts darüber, wer sie erzeugt hat oder wozu. Der Vorteil gegenüber Metadaten liegt darin, dass das Signal auch dann bleibt, wenn die Datei-Informationen beim Hochladen verloren gehen.

6 Was noch nicht rund läuft

- In Nutzerberichten ist der häufigste Kritikpunkt eine **Überenthusiasmus**-Note im Ton — Live klingt vielen zu aufgeregt.
- Die **Zeitkontingente ändern sich**. Die Zahlen hier sind ein Stand, keine Zusage; prüfe sie in deinem Konto nach.
- Ein Modell, das unterbrechen darf, unterbricht manchmal falsch. Genau dafür ist der Rückweg auf **Advanced** in den Einstellungen da.
- Es gibt **kein separates Abo** für Live und keine Guthaben-Pakete. Der Zugang hängt ausschließlich an deinem ChatGPT-Plan.

7 Quellen

- **Ankündigung:** openai.com/index/introducing-gpt-live/
- **ChatGPT Voice (Help Center):** help.openai.com/en/articles/20001274
- **Wasserzeichen:** openai.com/index/advancing-content-provenance/

■ // AUF EINEN BLICK

Die Kurzfassung

Vier Punkte für den Start.

- ✓ **Sprachsymbol** in der Eingabezeile antippen, Mikrofon erlauben
- ✓ Umschalten unter **Einstellungen > Sprache**: Live, Advanced, Standard
- ✓ Denkstufen sind das knappe Gut — **je eine rollende Stunde** auf Go und Plus
- ✓ Über Gespräche hinweg bleiben nur **Fakten**, nicht deine Sätze

■ // BLEIB DRAN

KI-Tools testen, auch das Kleingedruckte.

Bei einem Voice-Release schaue ich zuerst auf Plan-Grenzen und Gedächtnis, weil sich dort entscheidet, ob es im Alltag trägt. Folg mir für die nächste Einordnung.

@ki.borish

Fakten verifiziert am 03.09.2026 aus OpenAIs Ankündigung „Introducing GPT-Live“ (08.07.2026), dem Help-Center-Artikel „ChatGPT Voice“ und OpenAIs Beitrag zur Content-Provenienz. Zeitkontingente, Planzuordnungen und Menüpfade sind ein Stand und können sich mit neuen Releases ändern — prüfe sie in deinem Konto nach. Die Angabe von mehr als 150 Millionen Menschen pro Woche bezieht sich auf Voice und Dictation zusammen.