

Sprachprobe und Sprechtexte

Wie Sie eine saubere Sprachprobe aufnehmen und Texte so vorbereiten, dass Ihre KI-Stimme sie fehlerfrei vorliest – auch mit Formeln aus Mathe und Chemie.

Stand: Oktober 2026 · Ergänzt die „Technische Anleitung“ zur Installation



1 Sprachprobe aufnehmen

XTTS v2 braucht eine saubere Referenzaufnahme Ihrer Stimme. Mindestens 6 Sekunden funktionieren, **15 bis 30 Sekunden** klingen deutlich stabiler.

- **Ruhiger Raum, festes Mikrofon:** gleicher Abstand während der ganzen Aufnahme, kein Hall. Ein Raum mit Vorhängen oder Teppich ist ideal.
- **2 bis 3 Sekunden Stille zuerst:** Nehmen Sie vor dem ersten Wort kurz nur den Raum auf. Diese Stelle brauchen Sie später für das Rauschprofil.
- **Natürliches Tempo:** Sprechen Sie so, als würden Sie vorne stehen und Ihre Klasse begrüßen. Nicht überdeutlich, nicht extrem langsam.
- **Leichtes Lächeln:** Man hört es in der Stimme. Die KI-Stimme klingt dann später freundlicher.
- **Umlaute deutlich formen:** Die KI übernimmt Ihre Art, ü, ö und ä zu bilden.
- **Stimmung passend zum Einsatz:** Die KI übernimmt auch Tonfall und Energie. Für Erklärtex te also ruhig und zugewandt sprechen, nicht vorlesend-monoton.

REFERENZTEXT DEUTSCH · CA. 20–25 SEKUNDEN

Über den weiten Mauern leuchtet das herbe Licht der frühen Morgensonne. In der Küche duftet es bereits süß nach frischem Gebäck und Kräutertee. Völlig ungestört fließen die Gedanken durch den Raum, während draußen der kühle Wind in den Bäumen spielt. Jedes Wort und jeder Laut zählt, um ein klares Bild dieser Atmosphäre zu zeichnen. Wir prüfen nun sorgfältig, ob alle Töne und Nuancen richtig erfasst wurden.

The quick brown fox jumps over the lazy dog. Beyond the quiet mountains, a bright light shines through the morning mist. We are exploring new ways to create engaging lessons for our students today. It is truly amazing how technology helps us to share knowledge and express our thoughts clearly. Let's see how well this digital twin captures the natural rhythm of my voice.

2 Sprachprobe in Audacity bereinigen

Reihenfolge beachten

Erst das Rauschprofil aus der Stille holen, dann die Stille abschneiden. Umgekehrt fehlt Ihnen sonst die Stelle für das Profil.

- 1 **Rauschprofil ermitteln:** Die 2 bis 3 Sekunden reine Raumstille am Anfang markieren (kein Atmen, kein Klicken). *Effekt → Entfernung und Reparatur → Rauschminderung...* → „Rauschprofil ermitteln“.
- 2 **Rauschen entfernen:** Gesamte Spur markieren (**Strg** + **A**), wieder *Rauschminderung...* aufrufen und mit OK bestätigen.
- 3 **Stille abschneiden:** Leere Bereiche vor dem ersten und nach dem letzten Wort markieren und mit **Entf** löschen. Die Spur beginnt direkt mit dem ersten Laut.
- 4 **Mundgeräusche entfernen:** Einzelne Schmatzer oder Klicks gezielt markieren und löschen oder mit *Effekt → Stille erzeugen* ersetzen. Den Klick-Entferner von Audacity nur vorsichtig nutzen: Er ist für Schallplattenknacken gedacht und kann Konsonanten wie t und k beschädigen.
- 5 **Normalisieren:** Gesamte Spur markieren, *Effekt → Lautstärke und Kompression → Normalisieren...*, „Spitzenamplitude normalisieren auf“ -1,0 dB.
- 6 **Exportieren:** *Datei → Exportieren → Als WAV exportieren*. Mono, 22.050 Hz oder 44.100 Hz, Dateiname ohne Leerzeichen, z. B. `sprechprobe_sauber.wav`.

Nach der Rauschminderung anhören

Klingt die Stimme blechern oder wie unter Wasser, war die Rauschminderung zu stark. Mit **Strg** + **Z** zurück und mit weniger Dezibel erneut versuchen.

3 Regeln für die sprechtext.txt

Grundsatz: Alles so schreiben, wie es gesprochen werden soll. XTTS liest sehr wörtlich und kennt keine Formeln.

Allgemein

Regel	Beispiel oder Erklärung
Korrekte Rechtschreibung	Groß- und Kleinschreibung hilft bei der Betonung.
Kurze Sätze	Bandwurmsätze teilen. Bei langen Einheiten verschluckt die KI eher das Satzende.
Leerzeile bei neuem Gedanken	z. B. zwischen Einleitung und Aufgabe
Zahlen ausschreiben	105 → einhundertfünf · 3,5 → drei Komma fünf
Einheiten ausschreiben	10 ml → zehn Milliliter · 25 °C → fünfundzwanzig Grad Celsius
Abkürzungen auflösen	z. B. → zum Beispiel · bzw. → beziehungsweise · ca. → ungefähr
Sonderzeichen ersetzen	Keine Symbole wie $\sqrt{\quad}$, π , $\int$, $\rightarrow$, 2 , $^+$ im Text lassen. Sie führen zu Fehlern oder seltsamen Geräuschen.
Listen	Jeder Punkt in eine eigene Zeile, mit Punkt am Ende: „Erstens: Reagenzglas vorbereiten.“
Lange Fachwörter	Nur trennen, wenn die Stimme tatsächlich stolpert („Titrations Kurve“). Nicht vorsorglich, sonst entstehen unnötige Pausen.

Einzelne Buchstaben (Variablen und Elementsymbole)

Einheitliche Schreibweise mit Doppelvokal oder Doppelkonsonant, ohne „h“ am Ende. Gilt für Mathe und Chemie gleichermaßen. Klingt ein Buchstabe mit Ihrer Stimme seltsam, passen Sie ihn hier zentral an.

a aa	b bee	c zee	d dee	e ee	f eff
g gee	h haa	i ii	k kaa	l ell	m emm
n enn	o oo	p pee	r err	s ess	t tee
u uu	v fau	x iks	y üpsilon	z zett	α Alpha

Weitere griechische Buchstaben ebenfalls als Wort: Δ → Delta, λ → Lambda, π → Pi.

Chemie

Original	Sprechtext
H_2O	Haa zwei Oo
CO_2	Zee Oo zwei
NaCl	Enn aa, Zee ell
$2 \text{H}_2\text{O}$	zwei, Haa zwei Oo
$\text{Ca}(\text{OH})_2$	Zee aa, Klammer auf, Oo Haa, Klammer zu, zwei
Na^+	Enn aa plus
SO_4^{2-}	Ess Oo vier, zwei minus
$\rightarrow$	reagiert zu
$\rightleftharpoons$	steht im Gleichgewicht mit
(aq) (s) (l) (g)	gelöst, fest, flüssig, gasförmig

Mathe

Original	Sprechtext
x^2	iks hoch zwei (bei Flächen: iks Quadrat)
-3^2	minus, drei hoch zwei
$(-3)^2$	minus drei, in Klammern, hoch zwei
$\frac{3}{4}$	drei Viertel
$(a+b)/2$	aa plus bee, geteilt durch zwei
$\sqrt{x+1}$	Wurzel aus, iks plus eins
$f(x)$	eff von iks
$f'(x)$	eff Strich von iks
$a \cdot b$	aa mal bee
$= \neq < \leq$	ist gleich, ist ungleich, ist kleiner als, ist kleiner gleich

Fachlich korrekt bleiben

Elementsymbole aus zwei Buchstaben immer einzeln buchstabieren, nie als Wort lesen lassen. Den Stoffnamen („Wasser“) nur verwenden, wenn er im Originaltext steht. Ein Minuszeichen nie als Bindestrich stehen lassen, sondern immer als „minus“ ausschreiben, und Klammern so sprechen, dass der Unterschied hörbar ist.

4 Satzzeichen und Pausen

Mit Satzzeichen steuern Sie Rhythmus und Satzmelodie. Zusätzliche Punkte („...“ oder „. .“) erzeugen bei XTTS oft Störgeräusche, deshalb lieber diese Mittel nutzen:

Zeichen	Wirkung	Gut für
Komma	kurzer Atemzug, Stimme bleibt oben	Aufzählungen, nach einzelnen Variablen oder kurzen Fachbegriffen
Semikolon	mittlere Pause	zwei zusammenhängende Formeln oder Gedanken
Doppelpunkt	Pause mit Erwartung	vor Erklärung oder Liste
Fragezeichen	Stimme hebt sich am Ende	echte Fragen an die Lernenden
Gedankenstrich	kurze, natürliche Zäsur	Einschübe; nie als Minuszeichen verwenden
Leerzeile	neues Einatmen, frischer Ansatz	Themenwechsel
Anführungszeichen	etwas präzisere Betonung	Fachbegriffe, die sonst untergehen

Kleine Tricks

Nach einzeln stehenden Variablen ein Komma statt eines Punktes setzen, wenn nur eine kurze Pause folgen soll. Das verhindert, dass die Stimme absackt. Klingt ein Wort zu abgehackt, hilft manchmal eine Dehnung mit Doppelvokal („Mool“ statt „Mol“) – ein Notbehelf für Einzelfälle, keine Grundregel.

5 TTS-Vorbereiter-Prompt (Deutsch)

Diesen Prompt kopieren Sie in eine beliebige KI und hängen Ihren Rohtext an. Eine KI kann den Text auch direkt aus einem Screenshot auslesen.

Überarbeite den folgenden Text so, dass ihn eine Text-to-Speech-KI (XTTS) optimal vorlesen kann. Gib nur den fertigen Sprechtext aus, ohne Erklärungen. Beachte streng diese Regeln:

1. Inhalt: Ändere nichts am fachlichen Inhalt. Ersetze keine Formel durch einen Stoffnamen oder eine Umschreibung, die nicht im Text steht.

2. Einzelne Buchstaben (Variablen und Elementsymbole) schreibst du so aus:

a = aa, b = bee, c = zee, d = dee, e = ee, f = eff, g = gee, h = haa, i = ii, k = kaa, l = ell, m = emm, n = enn, o = oo, p = pee, r = err, s = ess, t = tee, u = uu, v = fau, x = iks, y = üpsilon, z = zett. Griechische Buchstaben als Wort (Alpha, Delta, Lambda).

3. Chemie:

- Elementsymbole Buchstabe für Buchstabe nach Regel 2, Indizes als Zahlwort: H₂O = Haa zwei Oo, CO₂ = Zee Oo zwei, NaCl = Enn aa, Zee ell.
- Koeffizienten mit Komma vom Teilchen trennen: 2 H₂O = zwei, Haa zwei Oo.
- Klammern aussprechen: Ca(OH)₂ = Zee aa, Klammer auf, Oo Haa, Klammer zu, zwei.
- Ladungen: Na⁺ = Enn aa plus, SO₄²⁻ = Ess Oo vier, zwei minus.

- Pfeile: $\rightarrow$ = reagiert zu, $\rightleftharpoons$ = steht im Gleichgewicht mit.
- Aggregatzustände: (aq) = gelöst, (s) = fest, (l) = flüssig, (g) = gasförmig.

4. Mathe:

- Rechenzeichen als Wort: plus, minus, mal, geteilt durch, ist gleich, ist ungleich, ist kleiner als, ist größer als, ist kleiner gleich, ist größer gleich.
- Potenzen: x^2 = iks hoch zwei.
- Vorzeichen und Klammern hörbar machen: -3^2 = minus, drei hoch zwei; $(-3)^2$ = minus drei, in Klammern, hoch zwei.
- Brüche: $3/4$ = drei Viertel; $(a+b)/2$ = aa plus bee, geteilt durch zwei.
- Wurzeln: $\sqrt{x+1}$ = Wurzel aus, iks plus eins.
- Funktionen: $f(x)$ = eff von iks; $f'(x)$ = eff Strich von iks.
- Ein Minuszeichen wird nie als Bindestrich geschrieben, sondern immer als „minus“.

5. Zahlen und Einheiten: Alle Zahlen als Wörter (105 = einhundertfünf, 3,5 = drei Komma fünf). Einheiten ausschreiben (ml = Milliliter, °C = Grad Celsius, mol/l = Mol pro Liter).

6. Abkürzungen vollständig auflösen (z. B. = zum Beispiel, bzw. = beziehungsweise, ca. = ungefähr).

7. Keine Sonderzeichen im Ergebnis: keine Formelzeichen, hochgestellten Zahlen, Pfeile oder Emojis.

8. Struktur: Kurze Sätze. Leerzeile bei Themenwechsel. Listenelemente jeweils in eine eigene Zeile mit Punkt am Ende. Nach einzeln stehenden Variablen oder kurzen Fachbegriffen ein Komma statt eines Punktes, wenn nur eine kurze Pause folgen soll. Keine Auslassungspunkte.

9. Lange Fachwörter nicht trennen, außer sie stehen in der Liste: [hier ggf. Wörter eintragen, an denen die Stimme stolpert].

Hier ist der Text:
[IHR TEXT]

6 Fremdsprachen (Beispiel Englisch)

XTTS v2 ist mehrsprachig. Für englische Texte setzen Sie im Skript nur den Sprachparameter von "de" auf "en". Die deutsche Sprachprobe reicht meist, weil das Modell den Klang Ihrer Stimme auf die englische Aussprache überträgt. Eine eigene englische Probe (Referenztext in Abschnitt 1) klingt oft noch natürlicher.

Rewrite the following text so that a text-to-speech AI (XTTS, English) can read it aloud optimally. Output only the finished speech text, no explanations. Follow these rules strictly:

1. Content: Do not change the subject matter. Never replace a formula with a substance name that is not in the text.

2. Single letters (variables and element symbols): a = ay, b = bee, c = cee, d = dee, e = ee, f = ef, g = gee, h = aitch, i = eye, k = kay, l = el, m = em, n = en, o = oh, p = pee, r = ar, s = ess, t = tee, u = you, v = vee, x = ex, y = why, z = zed (British) or zee (American).

3. Chemistry: element symbols letter by letter, subscripts as number words: H2O = aitch two oh, CO2 = cee oh two, NaCl = en ay, cee el. Separate coefficients with a comma: 2 H2O = two, aitch two oh. Charges: Na+ = en ay plus. Arrows: $\rightarrow$ = reacts to form, $\rightleftharpoons$ = is in equilibrium with.

4. Maths: + = plus, - = minus, $\times$ or $\cdot$ = times, / = divided by, = = equals. x^2 = ex squared, x^3 = ex cubed, x^4 = ex to the power of four. $\sqrt{}$ = the square root of. $f(x)$ = ef of ex. Never write a minus sign as a hyphen.

5. Numbers and units: all numbers as words (105 = one hundred and five). Units in full (ml = millilitres).

6. Abbreviations in full (e.g. = for example, approx. = approximately).

7. No special characters, superscripts, arrows or emojis in the result.

8. Structure: short sentences, a blank line for topic changes, one list item per line. Use a comma for short pauses after single variables. No ellipses.

Here is the text:

[YOUR TEXT]

7 Wichtige Hinweise

Immer gegenlesen und anhören

Beim Ausschreiben von Zahlen, Formeln und Ladungen macht auch die KI Fehler. Den Sprechtext vor der Vertonung prüfen und die fertige Audiodatei einmal komplett anhören.

- **Nur die eigene Stimme:** Stimmen anderer Personen nur mit deren ausdrücklicher Einwilligung klonen.
- **Kennzeichnen:** Lernende sollten wissen, dass sie eine KI-Stimme hören, z. B. mit dem Satz „Diese Anleitung wird von einer KI-Version meiner Stimme gesprochen.“
- **Lizenz des Modells:** XTTS v2 steht unter der Coqui Public Model License, die nach aktuellem Kenntnisstand nur nicht-kommerzielle Nutzung erlaubt. Für Unterricht und frei geteilte Materialien passt das in der Regel. Bei Material, das im bezahlten Auftrag entsteht, vorher die Lizenz prüfen.
- **Software-Stand:** Die Firma Coqui existiert nicht mehr, das ursprüngliche Paket `TTS` wird nicht weiterentwickelt. Der gepflegte Nachfolger `coqui-tts` läuft auch mit neueren Python-Versionen – Details in der Technischen Anleitung.