



---

≡ Media and devices > **Talk Mode**

---

Media and devices

## Talk Mode

Talk mode is a continuous voice conversation loop:

1. Listen for speech
2. Send transcript to the model (main session, chat.send)
3. Wait for the response
4. Speak it via ElevenLabs (streaming playback)

## Behavior (macOS)

**Always-on overlay** while Talk mode is enabled.

**Listening** → **Thinking** → **Speaking** phase transitions.

On a **short pause** (silence window), the current transcript is sent.

Replies are **written to WebChat** (same as typing).

**Interrupt on speech** (default on): if the user starts talking while the assistant is speaking, we stop playback and note the interruption timestamp for the next prompt.

## Voice directives in replies

The assistant may prefix its reply with a **single JSON line** to control voice:



```
{ "voice": "<voice-id>", "once": true }
```

## Rules:

First non-empty line only.

Unknown keys are ignored.

`once: true` applies to the current reply only.

Without `once`, the voice becomes the new default for Talk mode.

The JSON line is stripped before TTS playback.

## Supported keys:

`voice` / `voice_id` / `voiceId`

`model` / `model_id` / `modelId`

`speed`, `rate` (WPM), `stability`, `similarity`, `style`, `speakerBoost`

`seed`, `normalize`, `lang`, `output_format`, `latency_tier`

`once`

## Config ( `~/.openclaw/openclaw.json` )

```
{
  talk: {
    voiceId: "elevenlabs_voice_id",
    modelId: "eleven_v3",
    outputFormat: "mp3_44100_128",
    apiKey: "elevenlabs_api_key",
    interruptOnSpeech: true,
  },
}
```

## Defaults:



```
interruptOnSpeech : true
```

```
voiceId : falls back to ELEVENLABS_VOICE_ID / SAG_VOICE_ID (or first
ElevenLabs voice when API key is available)
```

```
modelId : defaults to eleven_v3 when unset
```

```
apiKey : falls back to ELEVENLABS_API_KEY (or gateway shell profile
if available)
```

```
outputFormat : defaults to pcm_44100 on macOS/iOS and pcm_24000 on
Android (set mp3_* to force MP3 streaming)
```

## macOS UI

Menu bar toggle: **Talk**

Config tab: **Talk Mode** group (voice id + interrupt toggle)

Overlay:

**Listening:** cloud pulses with mic level

**Thinking:** sinking animation

**Speaking:** radiating rings

Click cloud: stop speaking

Click X: exit Talk mode

## Notes

Requires Speech + Microphone permissions.

Uses `chat.send` against session key `main`.

TTS uses ElevenLabs streaming API with `ELEVENLABS_API_KEY` and incremental playback on macOS/iOS/Android for lower latency.

`stability` for `eleven_v3` is validated to `0.0`, `0.5`, or `1.0`; other models accept `0..1`.

`latency_tier` is validated to `0..4` when set.



Android supports `pcm_16000` , `pcm_22050` , `pcm_24000` , and `pcm_44100` output formats for low-latency `AudioTrack` streaming.

---

&gt;

&lt; Camera Capture

Voice Wake &gt;

Powered by [mintlify](#)