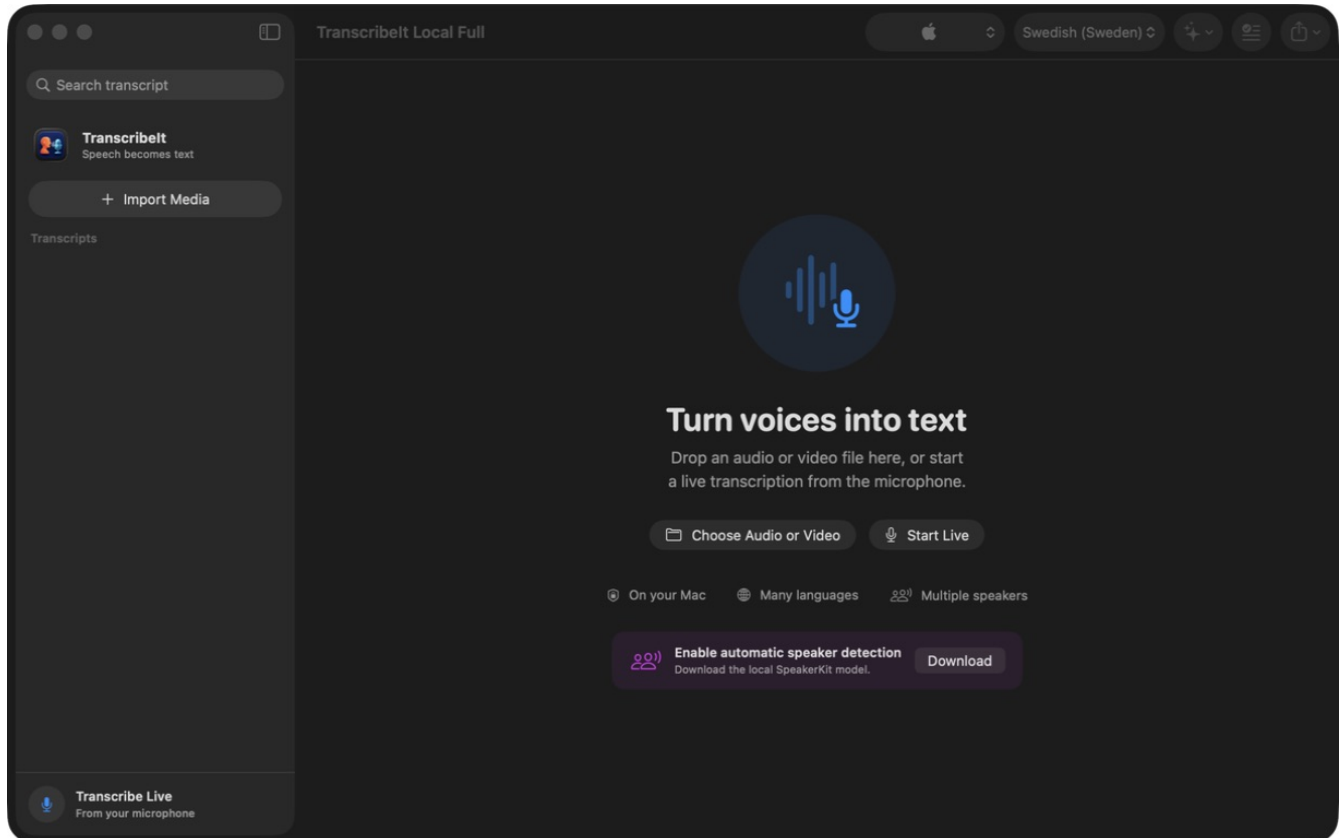


Transcribelt Local

Kort svensk manual - fullversion 1.0

All bearbetning kan ske lokalt. När valda modeller är nedladdade kan transkribering, talardetektering, sammanfattning, korrektur och neural talsyntes användas utan API-krediter. Internet behövs första gången en modell hämtas.



Kom igång

1 Installera

Öppna DMG-filen och dra **Transcribelt Local** till mappen Program. Appen kräver Apple Silicon och macOS 14 eller senare.

2 Välj motor och språk

Använd kontrollerna i verktygsfältet. **Apple Speech** är inbyggt. **Whisper** ger lokal transkribering med nedladdad modell. Systemspråket används första gången appen startar.

3 Transkribera

Välj **Import Media** för ljud eller film, eller **Transcribe Live** för mikrofonen. Tillåt mikrofon och taligenkänning när macOS frågar.

4 Granska och exportera

Spela upp materialet, redigera text och talarnamn, skapa AI-resultat och välj sedan önskat format i delningsmenyn.

Knapp- och menyname anges på engelska eftersom fullversionens gränssnitt är engelskt.

1. Transkribera ljud, film och möten

Importera en fil

- Klicka **Import Media** eller **Choose Audio or Video**. Du kan även dra filer till fönstret.
- Välj en eller flera ljud- eller videofiler. Flera filer behandlas i ordning.
- Appen sparar en lokal arbetskopia av hela mediefilen innan analysen startar. För film extraheras ljudet automatiskt.
- Status och förlopp visas i appen. Långa filer tar längre tid, särskilt med stora Whisper-modeller och talardetektering.

Live-transkribering

- Klicka **Transcribe Live** eller **Start Live** och tala mot datorns valda mikrofon.
- Klicka **Stop Live** för att avsluta. Vänta tills slutbehandlingen är klar - inspelningen sparas tillsammans med transkriberingen.
- Om en Whisper-modell är vald görs den slutliga transkriberingen med Whisper. Annars används Apple Speech.

Motor, modell och språk

Val	När passar det?	Att tänka på
Apple Speech	Snabbt och utan modellhämtning.	Körs på enheten när macOS och valt språk stöder det.
Whisper	Hög och jämn kvalitet, många språk.	Modellen måste vara nedladdad och vald.
Automatic language	När källspråket är okänt eller varierar.	Finns för Whisper under Settings - General .
Bestämt språk	När du vet vilket språk som talas.	Brukar ge stabilare stavning och mindre feltolkning.

Tips för bästa resultat: använd en bra mikrofon, undvik överlappande tal och välj ett bestämt språk när du känner till det. För högsta kvalitet väljer du Large v3 Turbo; för snabbare provkörning väljer du Base eller Small.

Talardetektering

- Aktivera **Identify different speakers automatically** i **Settings - General**.
- Välj **Automatic** eller ett känt antal talare, 1-10. Analysen körs efter transkriberingen.
- Talarnamn kan alltid rättas i efterhand. Resultatet blir bäst när rösterna inte talar samtidigt och inspelningen är tydlig.

2. Granska, spela upp och redigera

Synkroniserad uppspelning

- Tryck på en tidskod eller textrad för att hoppa till motsvarande plats i ljudet eller filmen.
- Använd spela/pausa, hoppa 10 sekunder bakåt eller framåt, tidsreglaget och hastighet 0.75x, 1x, 1.25x, 1.5x eller 2x.
- För film kan videopanelen visas eller döljas. **Show Current Transcript Row** hoppar från filmens aktuella position till rätt transkriptionsrad.
- Om mediefilen inte längre hittas väljer du **Choose Media File...** för att koppla den igen. Appen sparar då en ny lokal kopia.

Rätta text och talare

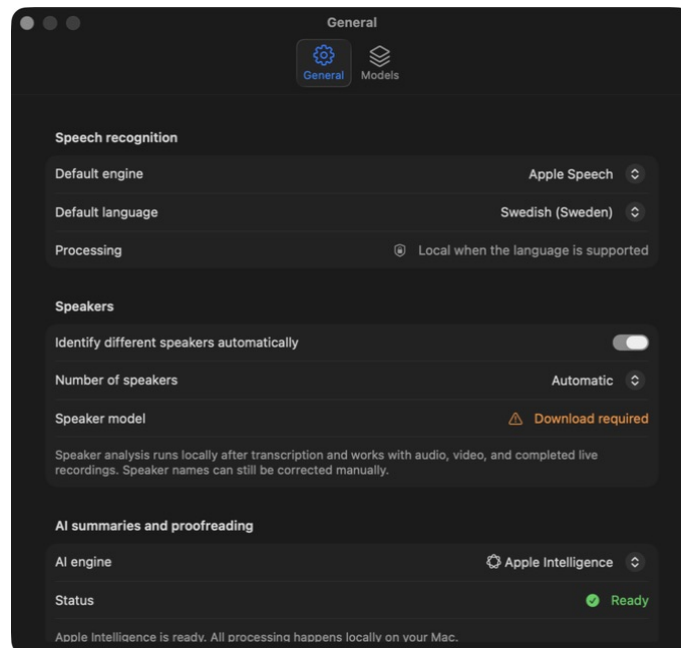
- Klicka i texten på en rad och skriv korrigeringen direkt.
- Klicka på ett talarnamn och byt namn, till exempel från Speaker 1 till Anna. Alla segment med samma talare uppdateras.
- På varje rad kan talarmenyn användas för att välja ett befintligt namn eller Speaker 1-8.
- Kopiera eller ta bort en enskild rad med radens meny. Felaktiga eller hallucinerade slutrader kan därmed tas bort.

Ordna transkriberingar

- Redigera rubriken högst upp för att byta namn på transkriberingen.
- Sök i samlingen med sökfältet i sidofältet.
- Högerklicka på en transkribering för **Copy Text** eller **Delete**. Radering går inte att ångra.

Rekommenderat arbetsflöde: lyssna igenom i 0.75x eller 1x, rätta talarnamn först, korrigera därefter texten och kör sist AI-korrektur. Exporten bygger på den sparade, korrigerade texten.

Inställningar i korthet



General innehåller standardmotor, standardspråk, talardetektering, AI-motor och standardformat för export.

3. AI: sammanfattning och korrektur

Välj AI-motor

Motor	Fördel	Krav
Apple Intelligence	Integrerat med macOS.	Kompatibel Mac, aktiverat Apple Intelligence och språkstöd.
Downloaded local model	Gemma kör lokalt utan API-krediter.	Gemma 3 eller Gemma 4 måste vara nedladdad och vald.

Skapa en sammanfattning

- Öppna **Summarize** i verktygsfältet.
- Välj **Summary** och ange ungefär 100-2000 ord i steg om 50. Texten kan bli något längre eller kortare för att sammanfatta hela materialet på ett naturligt sätt.
- Välj **Meeting minutes** för mötesprotokoll med ämnen, diskussion, beslut, åtgärder och öppna frågor.
- Välj utdata-språk: **Same as Transcript** eller bland annat svenska, engelska, tyska, franska, spanska, italienska, portugisiska, nederländska, nordiska språk, polska, tjeckiska, ukrainska, turkiska, arabiska, hebreiska, hindi, japanska, koreanska eller kinesiska.
- Långa transkriberingar behandlas i delar och sammanfogas till ett komplett resultat.
- Redigera resultatet direkt och använd **Copy**, **Save Changes** eller **Create Again**.

AI-korrektur

- Öppna **AI Proofreading**. AI:n letar efter stavnings- och skrivfel men försöker behålla innehåll och innebörd.
- Original och förslag visas sida vid sida. Markera de ändringar du vill använda och välj att endast tillämpa markerade korrigeringar.
- Granska alltid egennamn, facktermer, siffror och citat manuellt efteråt.

Integritet: Gemma behandlar texten lokalt. Ingen API-nyckel eller betald kredit behövs. Modellhämtningen kräver internet, men användningen därefter kan ske utan nätanslutning.

Exportera AI-resultat

- Sammanfattningar och protokoll kan exporteras som Word (.docx), renderad PDF eller komplett LaTeX (.tex) för exempelvis Overleaf.
- Spara dina manuella ändringar före export så att den slutliga versionen används.

4. Exportera text och ljud

Transkription och undertexter

Format	Användning
Text (.txt)	Ren, enkel text.
RTF (.rtf)	Formaterad text för Pages, Word och TextEdit.
Word (.docx)	Läsning, redigering och delning i Word eller Pages.
SRT (.srt)	Vanligt undertextformat med tidskoder.
WebVTT (.vtt)	Undertexter för webben och många videoplattformar.
JSON (.json)	Strukturerad data med segment och metadata.
Copy All Text	Kopierar hela den korrigerade texten till urklipp.

Originalt ljud

- Välj **Original Audio as M4A...** för att exportera inspelningen eller ljudet från den importerade mediefilen som M4A.
- Detta ändrar inte talet. Funktionen är bra för arkivering och kontroll mot transkriberingen.

Korrigerad text som tal

- Välj **Corrected Text as Speech....** Det är den aktuella, redigerade texten som läses upp och exporteras som M4A.
- Vid flera talare kan varje namngiven talare få en egen röst.

Talsyntes	Bäst för	Inställningar
Apple voices	Enkel och snabb export på många språk.	Installerad macOS-röst per talare samt talhastighet.
Kokoro Swedish	Naturligare svenska med färdiga röster.	10 svenska röster och hastighet 0.75-1.25.
Chatterbox V3	Röstkloning, främst engelska.	Språk, sparat röstprov och röst per talare. Svenska är experimentell.

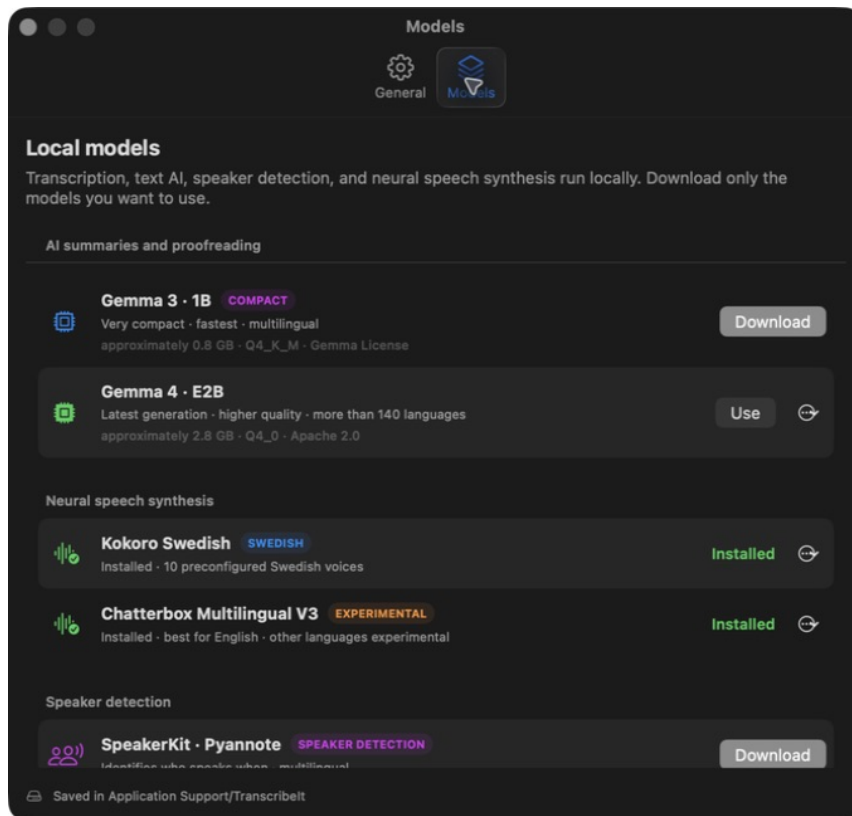
Röstprov i Chatterbox: använd ungefär 5-20 sekunder tydligt tal från en enda person. Provet kan sparas i röstbiblioteket och återanvändas. Använd endast andra personers röster med deras tillstånd.

Kvalitetsval

- Välj Kokoro för svenska. Chatterbox är bäst för engelska; andra språk kan få brytning eller tysta partier.
- Förhandslyssna på rösterna innan du skapar en lång fil. Längre synteser kan ta flera minuter.

5. Inställningar och lokala modeller

Öppna **Settings** och välj fliken **Models**. Hämta bara modellerna du behöver. **Download** hämtar, **Use** väljer, **Cancel** avbryter och modellens meny kan ta bort den.



Modell	Storlek / profil	Användning
Gemma 3 - 1B	ca 0.8 GB, kompakt, snabb	Lokala sammanfattningar och korrektur.
Gemma 4 - E2B	ca 2.8 GB, högre kvalitet, 140+ språk	Bättre lokala AI-resultat, kräver mer minne och tid.
SpeakerKit - Pyannote	Talarmodell	Identifierar vem som talar när.
Kokoro Swedish	Neural svensk talsyntes	10 färdiga svenska röster.
Chatterbox V3	Flera GB inklusive lokal motor	Röstkloning; bäst för engelska.

Whisper-modeller

Modell	Ungefärlig storlek	Profil
Tiny / Tiny English	75 MB	Snabbast; English är endast engelska.
Base / Base English	145 MB	Snabb; bra för prov och enklare material.
Small / Small English	500 MB	Bra balans mellan fart och kvalitet.
Large v3 Turbo	626 MB	Rekommenderad, flerspråkig och hög noggrannhet.
Large v3 Turbo (full)	1.5 GB	Maximal kvalitet för kraftfulla Macar.

6. Lagring, integritet och felsökning

Var sparas modeller och arbetsfiler?

Öppna Finder, välj **Gå - Gå till mapp...** och skriv:

```
~/Library/Application Support/TranscribeIt
```

- Här finns nedladdade modeller, lokala mediekopior och appens arbetsdata. Storleken kan bli flera GB.
- Ta helst bort modeller via **Settings - Models** så att appen samtidigt uppdaterar sin modellstatus.

Integritet

- Whisper, Gemma, SpeakerKit, Kokoro och Chatterbox körs lokalt efter nedladdning och använder inga API-krediter.
- Apple Speech kan köras på enheten när valt språk stöds av macOS. Appen visar och använder tillgängliga systemfunktioner.
- Importerade medier och liveinspelningar sparas lokalt för uppspelning och export.

Vanliga problem

Problem	Kontrollera
Ingen text i live-läge	Mikrofonbehörighet i Systeminställningar, rätt mikrofon och att Stop Live följts av slutbehandling.
Importen verkar stå still	Vänta på kopiering och slutanalys. Testa en mindre modell. Långa filmer och talardetektering kräver mer tid.
Endast början transkriberas	Använd senaste appversionen, låt filen kopieras helt och kontrollera att det finns ledigt lagringsutrymme.
Fel språk eller märkliga ord	Välj källspråk uttryckligen eller aktivera automatisk språkdetektion för Whisper.
Extra text efter tystnad	Ta bort den felaktiga raden. Tydligt språkval och mindre bakgrundsljud minskar Whisper-hallucinationer.
Modell går inte att använda	Kontrollera att hämtningen är klar, välj Use och starta om appen vid behov.
Media saknas vid uppspelning	Välj Choose Media File... och peka ut originalet igen.
Ingen exportdialog	Se efter om dialogen ligger bakom fönstret och kontrollera behörighet till vald målmapp.

Snabb rekommendation: välj Large v3 Turbo + automatisk talardetektering för viktiga möten, Gemma 4 för bästa lokala sammanfattning och Kokoro Swedish för svensk ljudexport.

Version 1.0 - manual för den fulla DMG-versionen. Funktionernas tillgänglighet kan påverkas av macOS-version, datorns minne och installerade modeller.