

Test och jämförelse av appar för tal-till-text

Uppdatering 2024



Innehåll

Inledning.....	2
Process.....	2
Material	3
Resultat.....	4
Slutsatser	7
Diskussion	7

Inledning

Automatisk omvandling av tal till skriven text blir alltmer relevant inom området elektronisk kommunikation och det finns en mängd av programvaror som erbjuder funktionen tal-till-text. I och med mängden av programvaror som finns, både förskrivningsbara och konsumentprodukter, kan det vara svårt för förskrivare att veta vilka programvaror som finns och hur väl de fungerar. Att individuellt undersöka en mängd olika produkter är ett stort arbete. Centrum för Elektronisk Kommunikation (CFEK) har under 2023 gjort ett test av ett antal vanligt använda programvaror för tal-till-text. CFEK har sedan upprepat detta test 2024 med detta års relevanta programvaror. Syftet med testet är att underlätta för landets förskrivare i deras urval av programvaror att informera om, ge råd kring eller förskriva.

CFEK tar inte ställning för eller mot någon av programvarorna som förekommer i testet. Vi har inte heller för avsikt att främja eller avråda från användning av någon specifik produkt. Vi uppmuntrar er att betrakta detta test som en utgångspunkt för reflektion och diskussion i ert fortsatta arbete kring tal-till-text.

Process

Programvaror för tal-till-text kan fungera på olika sätt och kallas till exempel diktering, transkribering eller automatisk textning. De programvaror som användes i testet och hädanefter benämns ”tal-till-text” är programvaror som automatiskt omvandlar tal till skriven text i realtid. För att avgränsa testet har vi valt att i vår poängsättning endast jämföra textningens kvalitet, ej andra aspekter.

För ytterligare information om tal-till-text, urval och process, vänligen se 2023 års rapport.

Exempel på andra typer av manuell eller automatisk textning som inte tagits med i detta test men kan vara ett alternativ som, i kombination med andra lösningar, kan underlätta för användaren är:

- Skrivtolkning, på plats eller på distans
- Text-TV
- Automatisk textning i PowerPoint
- Diktering i Word
- Liveundertext i Teams

Vi har i årets test använt oss av samma förinspelade ljudklipp som förra året. Vi har även läst upp samma manus i ”live”-testerna, men av logistiska skäl av andra personer. Vi har använt oss av samma rum och samma hårdvaror som vid föregående test. Anledningen till detta var att vi önskade få så liknande förutsättningar som möjligt, för att på ett rättvist sätt kunna jämföra 2023 och 2024 års textningar.

Group Transcribe har under året införlivats i Microsoft Translator och har samma textningsfunktionalitet men använder det nya namnet. Vi har valt att endast testa TERA i form av TERA Local, då det är samma textningsfunktion både i lokala samtal och telefonsamtal. Vi har även valt att utesluta T-talk från årets test då förskrivare nationellt har indikerat att T-talk inte förskrivs.

Testet genomfördes i två faser. I fas ett användes förinspelade ljudklipp, inspelade vid ett tidigare tillfälle med testets mikrofon. Ljudklippen spelades upp från en dator med hjälp av högtalare, till de olika programvarorna inkopplade med mikrofon. I fas två lästes manus upp i realtid av olika personer, med hög och klar röst och med samma avstånd till mikrofonen.

Textningarna har förts över till dator och styckeindelats på samma sätt, utan att någon text har förändrats. För att kunna jämföra textningarna har vi delat upp originaltexterna mening för mening. Textningarna har delats upp på samma sätt, mening för mening.

Varje mening i originaltexten har poängsatts. Poäng har getts för de ord och fraser som vi bedömt som ”viktiga”, de meningsbärande ord och fraser som gör att samtalet går att följa och förstå, totalt 30 poäng per samtal. Poäng har sedan satts på textningarna, om dessa viktiga ord och fraser har textats korrekt eller ej.

Bedömning om poäng har gjorts enligt följande: Om de poängsatta orden och fraserna har textats med samma innebörd ges poäng. Om bortfall skett men innebörden kvarstår ges poäng. Exempelvis om originaltexten anger ”Ja, det vill jag” kan både ”Ja, det vill jag”, ”Ja” eller ”Det vill jag” ge poäng.

Stavfel har godkänts, om innebörden blir densamma. Stavfel som ger ordet eller frasen en annan innebörd har inte godkänts.

Material

Programvaror:

- Anteckningar, dikteringsfunktion på Apples tangentbord - Apple
- OneNote, dikteringsfunktion - Microsoft Corporation
- Transkribering och aviseringar - Research at Google
- Microsoft Translator - Microsoft Corporation
- Rogervoice - RogerVoice
- TERA-Touch - Europea i Malmö AB

Hårdvaror:

- Apple iPhone 11
- Apple iPad 10,2 (9th gen)
- Samsung Galaxy A52
- Google Pixel 6
- Sennheiser SP 20 ML
- Logitech Z313

Manus:

- Biobesök
- Telefonförsäljare
- Biluthyrning
- Samtal med chefen
- Mobilförsäljare

Resultat

Nedan följer resultat i poäng från fas 1, fas 2 och totala poäng. Resultatet från 2024 finns i text- och diagramform och jämförande resultat 2023/2024 finns i diagramform. Full originaltext, textningar och poäng kan läsas i bilagan.

Fas 1 – inspelat ljud, 2024-10-17

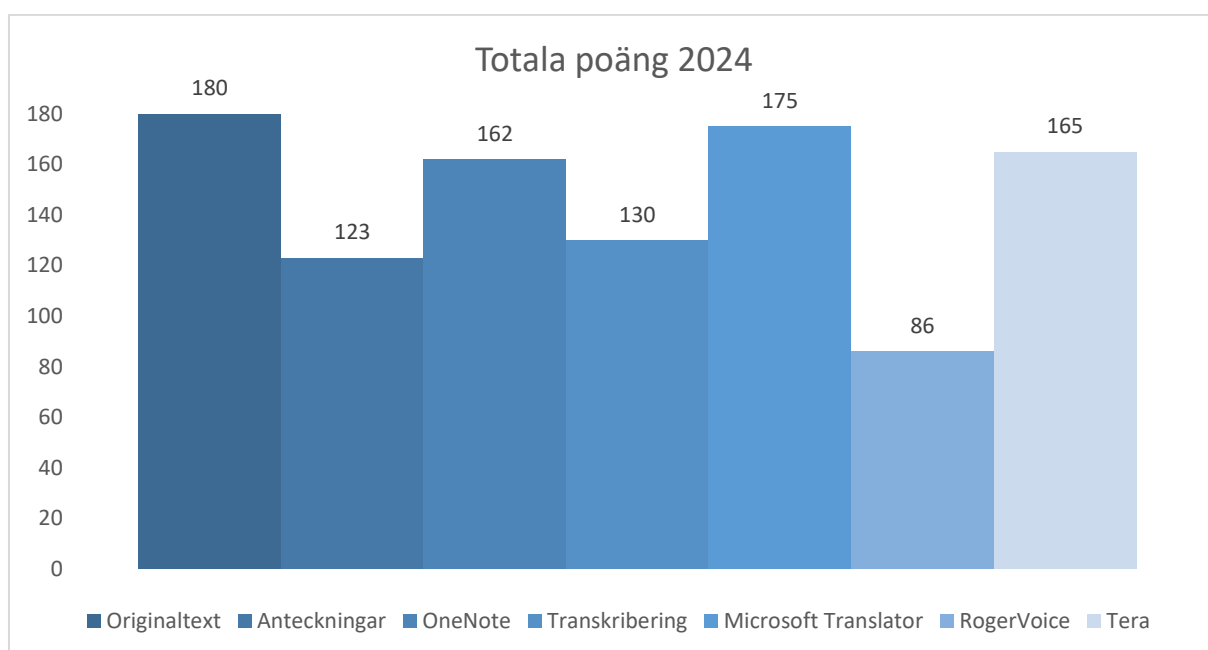
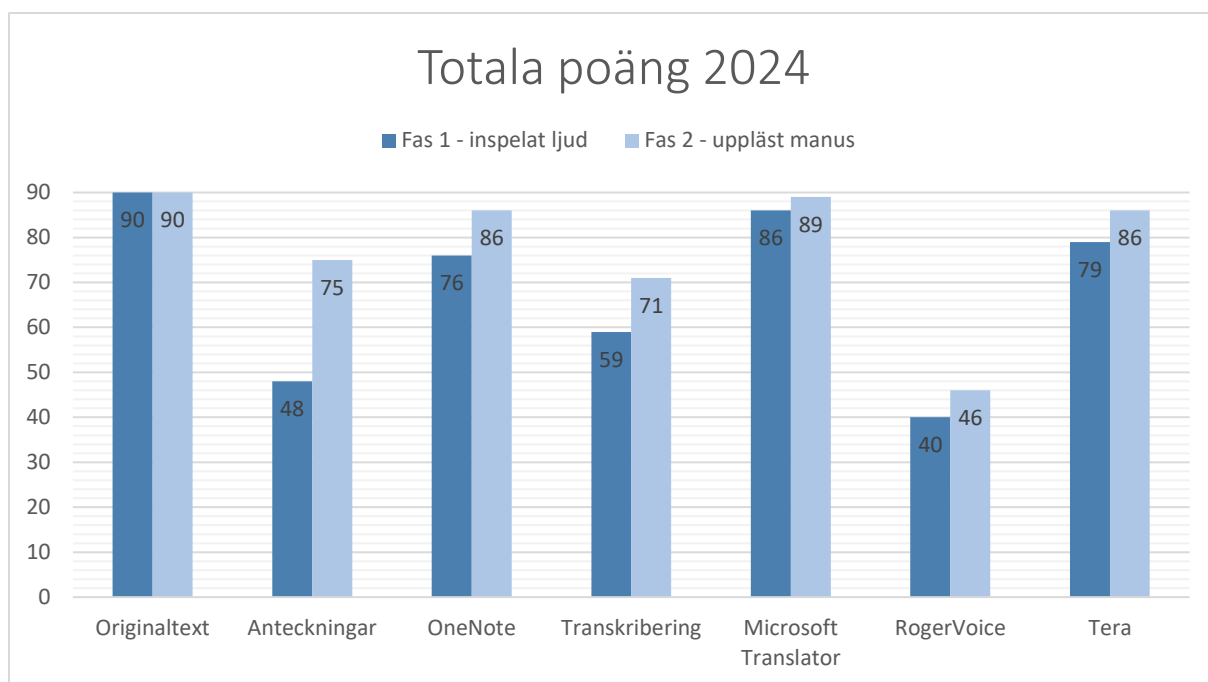
	<u>Biobesök</u>	<u>Telefonförs.</u>	<u>Biluthyrning</u>	<u>Totalt</u>
Originaltext	30	30	30	90
Anteckningar	11	14	23	48
OneNote	26	25	25	76
Transkribering	15	20	24	59
Microsoft Translator	28	30	28	86
RogerVoice	13	15	12	40
TERA	25	28	26	79

Fas 2 – uppläst manus, 2024-10-17

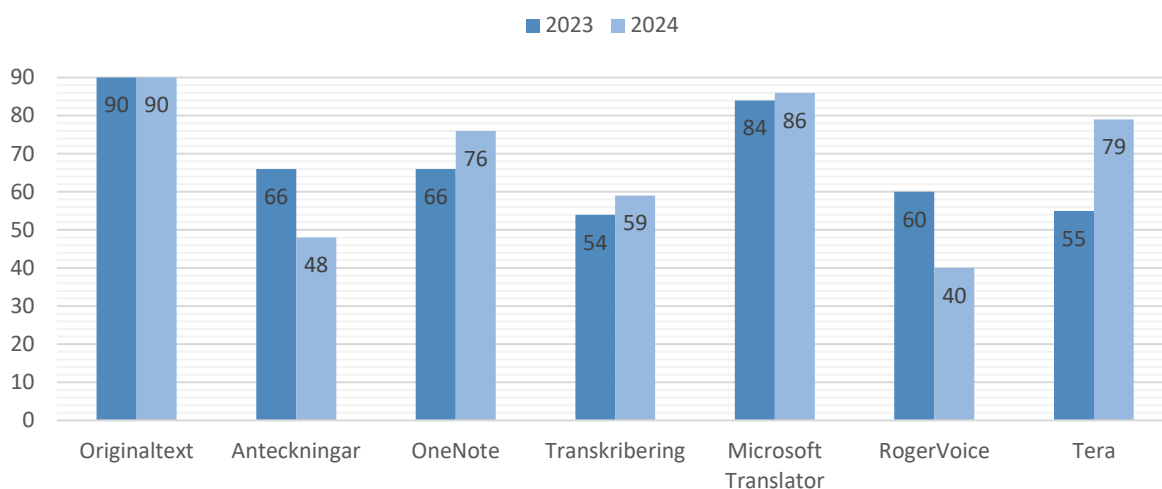
	<u>Med chefen</u>	<u>Mobilförs.</u>	<u>Biluthyrning</u>	<u>Totalt</u>
Originaltext	30	30	30	90
Anteckningar	28	21	26	75
OneNote	27	29	30	86
Transkribering	28	15	28	71
Microsoft Translator	30	30	29	89
RogerVoice	10	22	14	46
TERA	28	29	29	86

Sammanställning av poäng

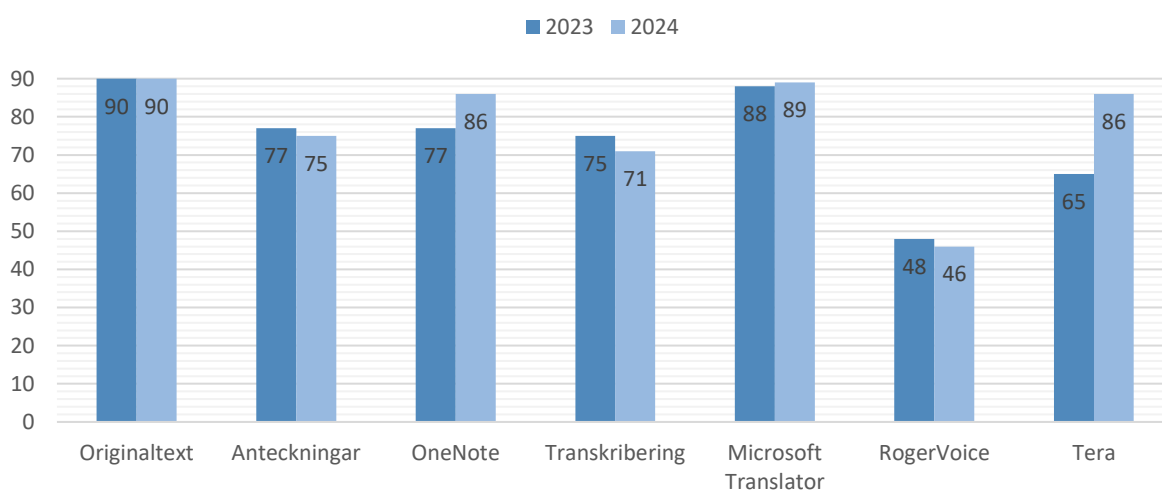
	<u>Fas 1</u>	<u>Fas 2</u>	<u>Totalt</u>
Originaltext	90	90	180
Anteckningar	48	75	123
OneNote	76	86	162
Transkribering	59	71	130
Microsoft Translator	86	89	175
RogerVoice	40	46	86
TERA	79	86	165



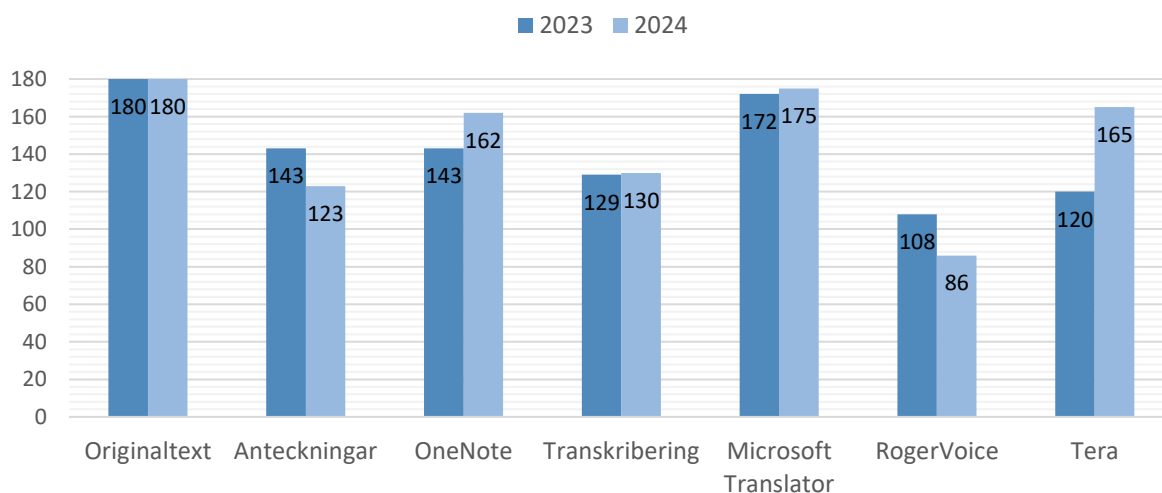
Inspelat ljud - jämförelse



Uppläst manus - jämförelse



Totalt resultat - jämförelse



Slutsatser

Vi har dragit några slutsatser utifrån poängsättningen som visas i resultatet.

Microsoft Translator har ökat sin pricksäkerhet från 95,5% till 97,2% och har därmed högst procent korrekta textningar i både 2023 och 2024 års test.

TERA har avsevärt högre pricksäkerhet än 2023 och har ökat från 66,6% till 91,6% korrekta textningar.

Lägst procent korrekta textningar har RogerVoice med 47,7%, vilket både är en försämring jämfört med 2023 samt en lägre procent än T-talk, som 2023 fick lägst procent (53,3%).

Diktering i Anteckningar och OneNote hade i fjolårets test lika många poäng. I årets test har Anteckningars resultat försämrats och OneNotes resultat förbättrats, vilket gör att deras resultat 2024 skiljer med 39 poäng från varandra.

Att Group Transcribe har införlivats i Microsoft Translator kan vara en av orsakerna till att OneNote, som också kommer från Microsoft, har förbättrat sitt resultat i årets test med 10,5%.

Samtliga appar har fler korrekta textningar med ett uppläst manus än med inspelat ljud, vilket är förväntat då naturligt talad röst har ett större omfång än digitalt uppspelat ljud.

I en jämförelse av 2023 och 2024 års inkorrekta textningar kan man se att det 2023 förekom många fler ”tokigheter”, där textningar blivit väldigt inkorrekta, felaktiga på ett komiskt sätt och omfattande feltextningar. 2024 års inkorrekta textningar bestod av färre ”tokigheter”, utan snarare feltextningar närmare originalet eller bortfall, där text helt enkelt saknas.

Diskussion

Utifrån mitt arbete som förskrivare av Alternativ Telefoni i Region Örebro län och som representant för Centrum för Elektronisk kommunikation kommer jag att väga in mina erfarenheter och tankar och dela med mig av några ytterligare tankar som förhoppningsvis kan vara till nytta för dig i fortsatta reflektioner, diskussioner och framtida arbete kring tal-till-text. Ha detta i bakhuvudet när du läser följande diskussion, och dra alltid egna slutsatser!

Obs! Fokus för denna diskussion är programvarornas dagsläge och förändringar som skett sedan föregående test. Se 2023 års rapport för en mer djupgående diskussion kring de olika programvarorna.

Microsoft Translator har fått högst procent korrekta textningar även i årets test, vilket går i linje med våra förväntningar. Under året har programvaran gått från att vara ett Microsoft Garage-projekt för iOS till att införlivas i Microsoft Translator, och finns därmed numer både för iOS och Android. Denna övergång hade troligtvis inte skett om programvaran inte fortsatt hålla en hög kvalitet. Under året har även flertalet användare och förskrivare uttryckt sig mycket positivt kring programvaran.

I dagsläget har programvaran olika utseende beroende på operativsystem. Det skulle vara en fördel om appen ser likadan ut för både iOS och Android. Microsoft Translator

för Android ser ut att inte vara lika utvecklad, vilket kan vara förståeligt då den först utvecklades till iOS. Eftersom flera enheter kan koppla upp sig till samma samtal, nu oavsett vilket operativsystem man använder, hoppas vi att appen får ett mer enhetligt utseende framöver.

Diktering i Anteckningar har fått ett sämre resultat i årets test. Detta var överraskande då Apples produkter generellt används i så stor utsträckning, vilket borde ge funktionen goda förutsättningar för utveckling.

Diktering i OneNote har å andra sidan förbättrat sitt resultat. Vi spekulerar i att det har att göra med att Group Transcribe har införlivats i Microsoft Translator vilket kan ha lett att OneNote, som också kommer från Microsoft, också har förbättrat sitt resultat.

Transkribering håller samma nivå sedan föregående år. Då flera av de andra programvarorna har förbättrat sitt resultat hoppas vi att Transkribering, som kommer från Google med alla dess resurser, också fortsätter att förbättras i framtiden för att inte halka efter. Transkribering är gratis och smidig, och är dessutom en tillgänglighetsapp. Textningen ger fortfarande ett gott resultat, men behöver fortsätta att utvecklas för att vara funktionell. Andra fördelar i form av smidighet och enkel användning skulle inte väga upp för en dålig textningskvalité.

TERA har avsevärt förbättrat sin textningskvalité sedan förra årets test. Detta är en positiv förbättring eftersom TERA fungerar både i lokala samtal och telefonsamtal, har bra möjligheter till synanpassningar, samt är en förskrivningsbar app. TERA har uppdaterats flertalet gånger sedan förra årets test med nya funktioner och förbättrad textning.

Rogervoice har tyvärr försämrat sitt redan låga resultat, vilket är tråkigt då det är den enda gratisapp i testet som textar telefonsamtal. Man bör kunna få bättre resultat av att använda ett headset, då appen är gjord för telefonsamtal, men om detta är gott nog är upp till användaren själv att avgöra.

Som summering så har alla appar sina fördelar och sina olika användningsområden. Oavsett poäng i detta test har alla appar något att ge. Olika appar kan behövas beroende på situation, och det gäller att hitta vad som fungerar bäst för varje person. Detta test kan användas som ett verktyg för detta, och vi på CFEK kan agera bollplank vid frågor och tankar.

Tekniken kring tal-till-text är relevant och utvecklas fort. Utseende, funktioner och tekniken innanför skalet kan förändras, förhoppningsvis till det bättre. Apparna i det här testet kan ha hunnit uppdateras från att jag skriver dessa ord, till att du läser dem. Ha i bakhuvudet att teknik är en färskvara och det är viktigt att göra en egen bedömning av appens nuvarande status innan en rekommendation eller förskrivning görs. Genom att använda dessa programvaror, jämföra dem och tycka till, kan vi leda utvecklingen i rätt riktning.

Att hålla sig uppdaterad kring all ny teknik är inte alltid en enkel sak på ett brett fält. Vi på CFEK tar gärna emot både frågeställningar och tips på nyheter inom området elektronisk kommunikation. Hör gärna av er till oss!