

OPTIMISING CONTACT CENTRE EFFICIENCY

Creating more accurate voice-to-text transcriptions with Jabra headsets

Transcription accuracy: an essential key to contact centre value

Contact centres are critical to many businesses, and the right technologies are essential to support staff in delivering outstanding customer care and for businesses to learn and grow.

In any one day, contact centres engage in hundreds of customer interactions. It's those interactions that determine the value of the company-customer relationship. To gain insight into customer satisfaction, how the market is changing, and how compliant employees are, analysis of those interactions is vital.

Customer Engagement Technologies that use artificial intelligence (AI) such as Interaction Analytics, Automated Quality Management, and AI-driven predictive modelling can leverage the unstructured data in these interactions. Unstructured data – data that exists in a raw and open form – may include call recordings, chat transcripts, survey comments and emails. When this data is accurately tagged and categorised according to specifications that are unique to the company, it can provide useful business and customer insights, increase business efficiencies, and transform the customer experience.

However, for accurate categorisation to occur – especially for call recordings – transcription accuracy is essential.

There are several factors that can impact transcription accuracy, from technology and environmental factors to language tuning and audio acuity (sharpness and clarity of sound). Just like human-to-human interactions, variances in background noise, pronunciations (e.g. different accents), ergonomics, voice clarity (e.g. mumbling or phone input quality) can assist or detract from the given baseline capability of an AI language model.

To be able to mine customer calls for business-critical insights at scale, it's necessary to have a system that accurately transcribes them, and technologies to compensate for less-than-ideal contact centre conditions.

Improving transcription accuracy requires a multi-pronged approach

Many contact centre managers aren't aware that their efforts to capture customer service interactions are being negatively impacted by factors outside their control.

Here's a simple example: Say a company wants to check their agents' compliance process around asking for a customer's address before providing sensitive financial information. If the transcription system transcribes "current address" as "can I address" or "calendar dress", it becomes challenging to use automation or analytics to understand how often agents are being compliant with this requirement.

The impact of poor-quality transcripts multiples across thousands of calls, undermining a company's efforts to support customers, train agents and drive business goals, especially in compliance-based investigations.

Enhancing transcription accuracy relies on quality voice data inputs – so how can we optimise these?

There are three key factors.



THE RIGHT EQUIPMENT

One piece of equipment that can make or break a contact centre's ability to obtain accurate transcriptions is the headset. For optimal sound quality, it's important to pay attention to the microphone, the speakers and the connection capability.

A headset with a noise-isolating and noise-cancelling microphone and a mute button will not only enhance the customer experience and dampen background noise for call centre agents, but it will support Automatic Speech Recognition (ASR) software to provide more accurate transcriptions. A quality microphone that maintains audio acuity is one of the single most important foundations for transcription accuracy.

Good-quality speakers will aid in sound isolation. This helps prevent external noise from polluting the contact centre agent's conversation, which ASR often tries to transcribe in competition to the person speaking. Limiting or even eliminating this background noise improves the likelihood of transcription accuracy.

A wireless connection will allow contact centre agents to be mobile, letting them move away from noisy areas, whereas a wired connection can provide maximum sound quality. It's also possible to combine a wireless connection with smart microphone technology, helping to preserve audio acuity while keeping contact centre agents mobile.

THE RIGHT ENVIRONMENT

With hybrid work being the norm, contact centres can't guarantee a controlled and quiet environment for their agent's audio, or a professional sound-proofed environment that is free of interruptions. This not only affects potential voice-to-text transcription, but other factors such as average handle time and customer satisfaction, as parties struggle to be understood.

For in-office calls, investing in sound dampening is a smart move. Consider acoustic treatments that help to absorb or diffuse sound, such as laying carpet on the floor, using cubicles or dividers between agents, and installing foam sound baffles on the walls and/or ceiling to prevent sound reverberation. Separating agents as much as possible is also advantageous.

While regulating an agent's home workspace isn't easy, contact centre managers can evaluate some aspects of their micro-environment. For example, are they placing their headset microphone at the right distance from their mouth? Is there anything in their workspace that creates unnecessary noise? Can any of this be mitigated by headset technology?

THE RIGHT SOFTWARE

Today's AI-powered ASR software can provide language services across multiple communication platforms that can help businesses gain insights into what their customers think and feel about their products and services. AI-driven sentiment analysis, opinion mining, predictive modelling, risk profiling and intent classification are some examples. But as beneficial as these services can be, they still rely on accurate transcriptions to deliver consistent, meaningful results.

The words that are missed in the gap between 85% and 95% individual word accuracy can be the ones that matter the most. They're likely to be the words that are specific to a business and can unlock value, such as:

- company and product names (for competitor and product issue insight)
- long numbers such as serial numbers (for highlighting risk such as recording of credit card information)
- dollar amounts and dates (for clues about trends and specific product/bundle issues).

When choosing AI-powered Analytics or ASR software to support accurate transcriptions, look for one that:





- is or can be custom tuned and continuously adapted to the lexicon of the business and the market (for example, interest rates, technology terms, etc.)
- has access to multiple language and accent models depending on call centre location and customer/agent demographics
- can analyse data and surface insights that create significant impact for the business
- is backwards-compatible with any surrounding technology, including best-of-breed headsets.

Understanding transcription accuracy

Human vs machine

There are multiple factors that influence how well a human can understand and interpret speech. From a machine perspective, similar factors apply, however the measure of success is determined by transcription accuracy.

Factors that affect human auditory processing vs machine audio processing

	HUMAN AUDITORY PROCESSING	MACHINE AUDIO PROCESSING
	Fluency: <ul style="list-style-type: none">In the languageAccent variation	Fluency: <ul style="list-style-type: none">Correct language model(s)Fine-tuning
	Background noise/ environment: <ul style="list-style-type: none">Of similar pitch e.g. noisy pub	Background noise/ environment: <ul style="list-style-type: none">Noisy vs quiet backgroundMicrophone acoustics
	Auditory acuity: <ul style="list-style-type: none">Clarity of the person speakingPhone line	Auditory acuity: <ul style="list-style-type: none">Clarity of the person speakingRecording qualityCompression rates
	Linguistic principles: <ul style="list-style-type: none">Suprasegmental features:<ul style="list-style-type: none">Stress (accent)ToneWord juncture (the transition between two successive syllables in speech. For example, “beer dripped” and “beard ripped” are distinguished by word juncture)Assumptions (for example are we expecting to hear “I scream” or “Ice cream” in the context of the sentence?)	Linguistic principles: <ul style="list-style-type: none">Semantic Indexing – the process of drawing meaning from textSuprasegmental features:<ul style="list-style-type: none">StressToneWord juncturePlosive consonants – the sounds that are generally associated with the letters p, t, k, b, d, g in English words. For example, the words ‘pat’, ‘kid’, ‘bag’.
SUCCESS MEASURE	Spoken words are heard and understood	Spoken words are accurately transcribed

Speech transcription error types

Transcription errors occur for reasons ranging from variability of the person speaking to hardware and software limitations. There are three main types of transcription errors:

- 1. Substitution** – when a spoken word is replaced with an incorrect guess.
- 2. Insertion** – when a spoken word is added that wasn’t said (usually from noise pollution).
- 3. Deletion** – when a spoken word is omitted altogether from the transcript, with no guess.

Within each error type there can be ‘relevant’ or ‘irrelevant’ errors.

- 1. A ‘relevant’ transcription error** is where the term is potentially useful for searching or algorithm discovery and can affect transcription data, potentially impacting the search returns or macro-level statistics.
- 2. An ‘irrelevant’ transcription error** refers to terms that are not useful for searching or algorithms, or could not reasonably be expected to be accurately transcribed (such as the name of a person or place) and is less consequential.

The goal is to minimise the number of ‘relevant’ transcription errors to gain a full picture of the call, and enable meaningful cross-correlation of search terms and diagnostic algorithms.

Speech transcription errors

Substitution

IRRELEVANT (Example)

Said:

“my name is Belinda at this time I'd like to let you know that this call may be ..”

Transcribed:

“my name is Melinda at this time I'd like to let you know that this call may be ..”

RELEVANT (Example)

Said:

“there may be a wait of up to fourteen days”

Transcribed:

“there may be a wait of up to forty days”

Insertion

IRRELEVANT (Example)

Said:

“not a problem _ do you happen to have the order number so I can ..”

Transcribed:

“not a problem so er do you happen to have the order number so I can ..”

RELEVANT (Example)

Said:

“you can contact the airline directly”

Transcribed:

“you can't contact the airline directly”

Deletion

IRRELEVANT (Example)

Transcribed:

“mobile signal lights on the back ___ on”

Omitted:

“are”

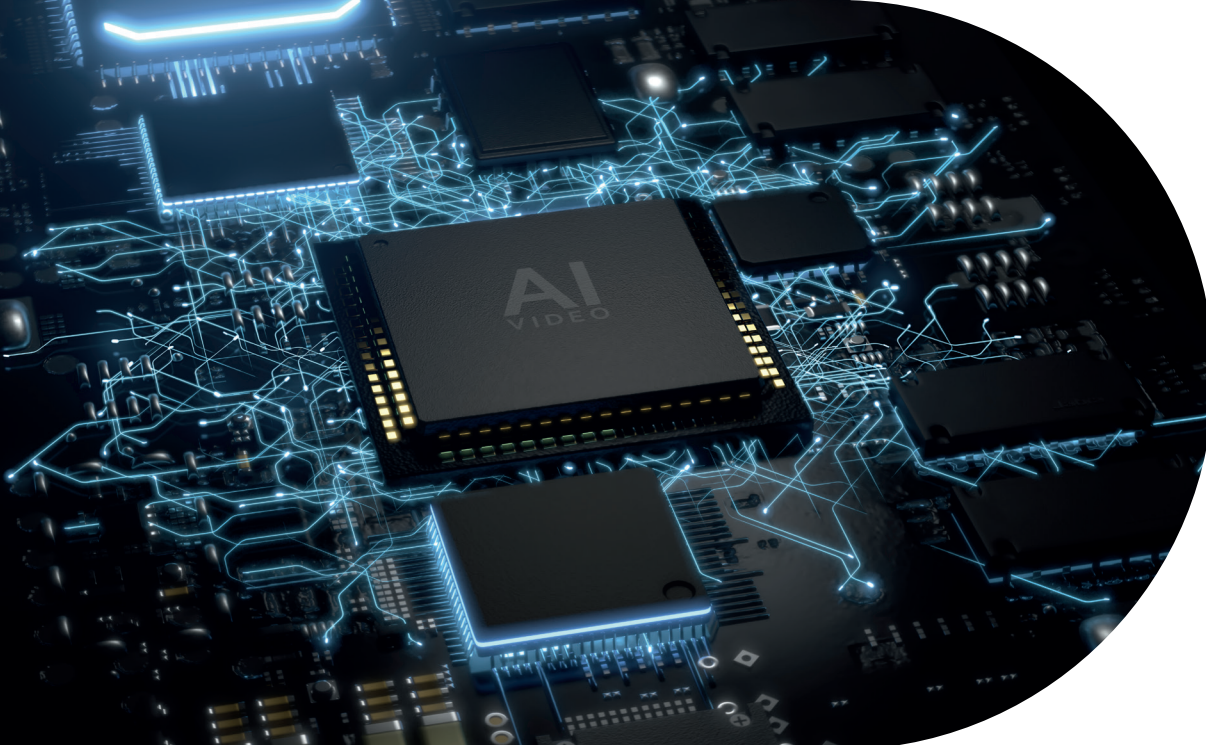
RELEVANT (Example)

Transcribed:

“___ still happening”

Omitted:

“it's (still happening)” (useful repeat call term)



The Jabra + NICE headset project

Test and learn

Jabra's mission is to improve people's lives through the power of sound. We regularly put our products to the test, pushing boundaries to create the most advanced sound solutions that support contact centre agents in delivering quality customer service.

In collaboration with NICE, Jabra conducted an experiment to establish the impact of Jabra headsets on speech analytics transcription output and accuracy vs non-Jabra headsets and earbuds in multiple environments, ranging from quiet to very noisy. The experiment was designed to simulate different contact centre scenarios.

Project aim

To determine to what extent the quality of the headset/microphone impacts the accuracy of voice-to-text transcription, and to what extent different background noise pollution impacts the transcription accuracy via the selected headset/microphone.

Method

A 1400-word test script (of which 831 words are considered 'relevant' to search/categorisation/algorithms) was read aloud by a contact centre agent a total of 15 times. Each time, one of five different headsets/microphones were used. The script was recorded across three different environments per headset – quiet background, mid-noise background and very noisy background.

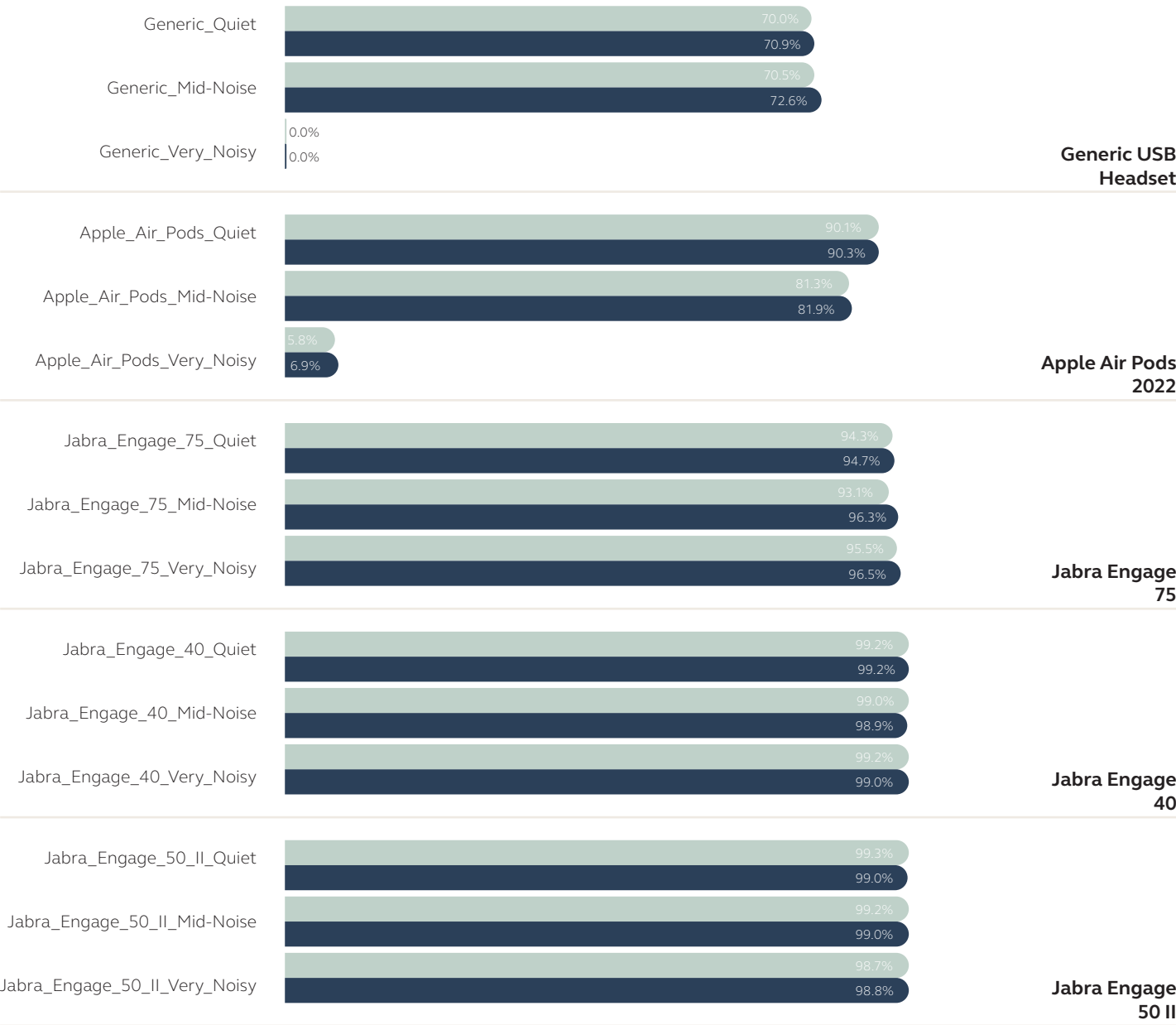
All recordings were transcribed using ASR software, NICE ElevateAI (AI-driven ASR Technology with sentiment scoring and behavioural analysis).

AUDIO PRODUCTS TESTED	ENVIRONMENTS TESTED	MEASURES
Jabra Engage 40 – wired Jabra Engage 50 II - wired Jabra Engage 75 – wireless Generic headset – wireless Apple AirPods (2022 edition) – wireless	<ol style="list-style-type: none"> 1. Quiet – no intentional noise interference. 2. Mid-noise - noisy contact centre. Voice input was high volume and close to microphone. 3. Very noisy - noisy café in close range + audio book played on high volume less than 0.5m from microphone + radio played mid volume less than 1m from microphone. 	<p>Transcriptions were measured against actual spoken words using the following measures:</p> <ol style="list-style-type: none"> 1. Accuracy Rate (%) = measure of accuracy of transcription against what was actually said. 2. Word Error Rate (WER) = errors or omissions of individual words across total words spoken. It is the inverse of the accuracy rate. 3. Word Relevancy Rate = percentage of words accurately transcribed that would be considered relevant in a speech analytics search or algorithm. This means ignoring proper nouns, agent names, and filler terms such as 'um' that are not required to materially impact the quality of speech analysis.

Results: Jabra headsets support the most accurate transcriptions

Comparison of Total Transcription Accuracy

	QUIET (no intentional noise interference)	MID-NOISE (noisy contact centre)	VERY NOISY (noisy café + audio book + radio)
Generic headset – wireless	70%	71%	0%
Apple Air Pods (2022 edition) – wireless	90%	81%	6%
Jabra Engage 75 – wireless	94%	93%	96%
Jabra Engage 40 – wired	99%	99%	99%
Jabra Engage 50 II – wired	99%	99%	99%



ACCURACY

All three Jabra headsets supported the most accurate transcriptions in a quiet environment, with >94% word accuracy and word relevancy accuracy. This equates to a Word Error Rate (WER) of <6%, which by industry standards is considered gold standard.

The generic headset failed to reach the required acuity to transcribe at all in a very noisy environment. In a controlled environment it performed adequately, but it did not shine in any environment.

CONSISTENCY

All three Jabra headsets performed consistently across environments irrespective of the amount of noise pollution.

The Apple AirPods performed well in a quiet environment but were highly impacted in a noisy environment.

NOISE CANCELLATION

All three Jabra headsets successfully eliminated background noise and maintained speaker acuity across multiple background interferences.

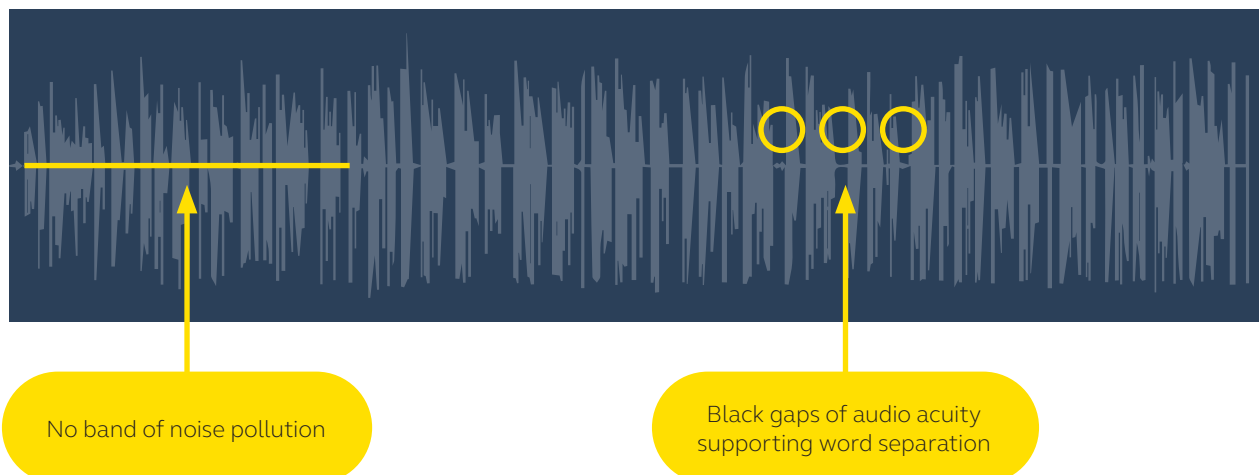
>94% individual word accuracy is an extremely high transcription score.

A 'good' transcription score for high-performance ASR software is >80%. Across a noisy environment this is exceptional.

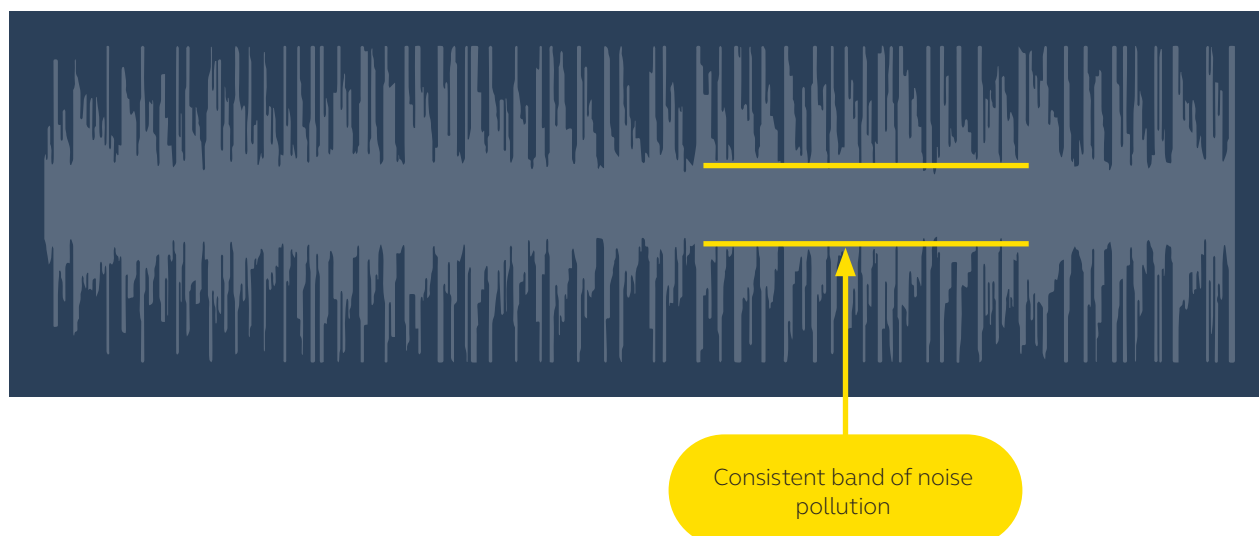
Above 99% individual word accuracy, which the Jabra Engage 40 and the Jabra Engage 50 II achieved, is better than human-level transcription accuracy.

Visualising audio acuity

Jabra Engage 75 in noisy environment



Generic headset in noisy environment



Jabra supports transcription accuracy to boost strategic value in contact centres

While it's amazing how many factors can contribute to poor transcription accuracy, the good news is that highly accurate transcriptions are now possible with Jabra headsets. This experiment established that the smart microphone of the Jabra headset shrugged off background interference and maintained audio acuity to allow for more accurate transcriptions.

It's time to Engage with Jabra.

JABRA ENGAGE 40



- **Professional quality calls** – Boost your call experience and don't miss a word with 2-mic technology and BalancedVoice which improves the incoming voice to be clearer and crisper. Transcription accuracy is enhanced as a result.
- **All-day comfort for the ears** – The ultra-lightweight angled ear cushions with maze-shaped pattern relieve pressure and provide an adaptive fit. There is also noise protection for agents' ears with Jabra Safetone.
- **Enhanced customer experience** – Free downloadable Engage+ software provides real-time intelligent call insights such as background noise, silence and interruptions to ensure client satisfaction.
- **Seamless collaboration** – Engage 40 is fully compatible with Unified Communications (UC) platforms such as Zoom, Google Meet, Microsoft Teams and more. Answer, mute, or adjust the volume using the inline control unit.

JABRA ENGAGE 50 II



- **Highest call quality** – This headset has 3-mic technology for exceptional sound quality, and noise-isolating earcups for awesome background noise-cancellation. These technologies combined allow customers to hear every single word said, enabling accurate speech-to-text call transcription. It also uses BalancedVoice technology to help the incoming voice sound clearer and crisper.
- **Enhanced customer experience** – Free downloadable Engage+ software provides real-time intelligent call insights such as background noise, silence and interruptions to ensure client satisfaction.
- **Healthy, happy ears** – With Jabra Safetone, agents can focus on your calls, safe in the knowledge that their hearing is protected. The lightweight headset and angled ear cushions make for a comfortable fit.
- **Never miss a call** – The SmartRinger technology uses sensors in the Engage 50 II headset to determine when it is being worn and when it is lying on the desk. If it is on the desk and a call comes in, the SmartRinger will ring to alert the agent to a call.
- **Seamless collaboration** – Engage 50 II plays well with all leading contact centre and Unified Communications platforms and even has integrated call control for the likes of Amazon Connect, Genesys Cloud CX, and NICE CXone, when used with the control unit accessory.* With a UC variant certified for platforms like Google Meet and Zoom and another for Microsoft Teams, agents can enjoy seamless and stress-free customer interactions across the board.
- **Protect agents' focus** – The integrated busylight helps prevent agent interruptions so they can stay focused on their customers and allow colleagues to see when they are free.

*Software may be needed for configuration.

JABRA ENGAGE 75



- **Superior wireless connectivity** – With a range of up to 150 metres, agents can work away from their desk and still maintain quality sound and service.
- **Increase agent focus** – Overcome interruptions with an integrated busylight that acts as a do-not-disturb sign for colleagues. Plus, the advanced noise cancelling microphone and enhanced speakers allows for crystal clear calls and improved focus.
- **All-day battery** – Get up to 13 hours of talk time. When charging is needed, the fast charge feature powers it to 40% in just 30 minutes and 100% in 90 minutes.
- **Enhanced hearing protection** – This is key for heavy users. With more innovations to safeguard intensive use, agents can comfortably use this headset all day.
- **Premium usability** – Easily answer/end calls, mute and adjust volume using the intuitive controls on the headset. Connect up to 5 devices at once to increase availability for customer calls.

Australian Signal Directorate (ASD) certification – Meets the strictest security requirements in Australia, helping high-risk industries such as government, military, intelligence and financial to avoid sensitive information being compromised by cyber-attack.



Find out more

If you have any questions about Jabra products, please contact your Jabra representative or visit **Jabra.com**

ABOUT JABRA

We engineer technology that makes life look and sound better. Our world-leading headsets, intelligent video technology and advanced earbuds make sure life and work stay wonderfully in tune. As part of GN, we're the only company in the world combining consumer, professional and medical grade sound solutions, all together under one roof.

ABOUT NICE ELEVATE AI

NICE ElevateAI leverages cutting-edge research in artificial intelligence (AI), automatic speech recognition (ASR), and natural language processing (NLP) to transcribe audio conversations into precise text that can be fed into any data lake or application.

With ElevateAI, individuals and organisations can access key data needed to understand conversations at scale. CX AI provides objective insights into audio, transcription, and chat interactions, such as voice activity and out-of-the-box, industry-agnostic behavioural models (including sentiment analysis, customer satisfaction behaviours, vulnerable customer behaviours, high risk transactions and more). For more information, please visit **elevateai.com**

ABOUT NICE

With NICE, it's never been easier for organisations of all sizes around the globe to create extraordinary customer experiences while meeting key business metrics. Featuring the world's #1 cloud native customer experience platform, CXone, NICE is a worldwide leader in AI-powered

self-service and agent-assisted CX software for the contact centre – and beyond. Over 25,000 organisations in more than 150 countries, including over 85 of the Fortune 100 companies, partner with NICE to transform – and elevate – every customer interaction. For more information, please visit

nice.com