
BRIEFING NOTE

Voice AI for knowledge extraction

Validating the case for voice as a capture mechanism

The core case

Speed

Average conversational speech runs at **150 to 160 words per minute** against typing at around **40 wpm**, a ratio of roughly 3.75:1. Stanford HCI research recorded **161 wpm for speech against 53 wpm** for keyboard entry on a phone. Under active cognitive load, where someone is thinking and composing at the same time, typing drops to as low as **19 wpm** and the gap widens to around **6x in favour of speech**. ^{[1][2][3][4]}

Richness

A 2025 Glaut study of 252 AI-moderated interviews found voice responses contained **236% more words**, were **138% more lexically diverse**, used **213% more content words**, and surfaced an average of **27 distinct themes per participant against 21** for text. Sector research confirms spoken responses are typically **4 to 5x longer** than typed equivalents. Voice also carries prosody (pitch, tempo, pauses) that transmits **over an order of magnitude more information** about emotional state than text alone. ^{[5][6][7][8]}

Reduced self-censorship

Speaking encourages stream-of-consciousness output with less real-time editing. Baddeley's phonological loop model explains why: typing forces **serial processing** (thinking plus motor coordination), while speaking allows **parallel processing**. Harvard Business School research on implicit voice theories shows employees routinely self-censor pro-organisational insights in written settings, and a conversational voice interface can lower that threshold. ^{[10][11][12][13][14]}

Lower barrier to entry

Typed capture needs a keyboard, screen, and seated posture. Voice can happen on a commute, on a walk, or immediately post-meeting, which matters because **expert recall is most vivid before cognitive residue decays**. AWS has built an enterprise reference architecture around exactly this pattern (Transcribe plus generative AI for structuring), and it is increasingly seen as the only viable mechanism for capturing tacit knowledge before senior staff leave. ^{[16][17][18]}

Additional benefits

- **Cognitive load reduction.** Writing demands simultaneous management of grammar, structure, motor coordination, and content generation. Speaking separates thinking from transcription. ^{[11][23]}
- **Tacit knowledge at scale.** A Kenyan Railways case study using AI voice detection to capture engineer expertise demonstrated that voice-based capture could surface knowledge that had never been codified, with the model improving as more voice data was ingested. ^[24]
- **Emotional signal extraction.** For qualitative work, voice AI can layer emotion detection alongside content extraction, a layer of intelligence categorically unavailable from typed surveys. ^[6]
- **Reduced survey fatigue and gaming.** Around 46% of typed survey respondents had to be removed due to bad data (incoherent or AI-generated). Voice responses are harder to fake convincingly. ^[6]

Sources

- [1] tap2talk.app/blog/typing-vs-dictation-speed
- [2] hci.stanford.edu/research/speech/paper/speech_paper.pdf
- [3] weesperneonflow.ai/en/blog/2025-10-16-voice-dictation-vs-typing-speed-productivity
- [4] aiche.app/blog/future-of-typing
- [5] glaut.com/glaut-research/voice-vs-text-in-ai-moderated-interviews
- [6] greenbook.org/insights/focus-on-apac/voice-ai-and-the-future-of-qual
- [7] voiceform.com/blog-posts/voice-data-for-academic-research
- [8] arxiv.org/html/2512.16832v1
- [9] aaltodoc.aalto.fi (2025 GenAI voice market feedback prototype study)
- [10] x.com/avysotsky/status/2040256075115724800
- [11] nesslabs.com/thinking-out-loud
- [12] bryon.io/how-voice-interaction-with-ai-boosts-productivity-and-creativity
- [13] hbs.edu/faculty/Pages/item.aspx?num=38244
- [14] journals.aom.org/doi/10.5465/amj.2011.61967925
- [15] alexandria.unisg.ch (University of St Gallen, 2025, 122-participant study)
- [16] cogzest.com/2012/03/some-benefits-of-dictation-for-cognitive-productivity
- [17] squirro.com/squirro-blog/ai-tacit-knowledge-capture
- [18] aws.amazon.com/blogs/machine-learning/unlock-organizational-wisdom-using-voice-driven-knowledge-capture
- [19] heydata.eu/en/magazine/a-deep-dive-into-data-privacy-in-voice-ai-technology
- [20] blog.milvus.io/ai-quick-reference/what-are-the-privacy-concerns-with-speech-recognition
- [21] brasstranscripts.com/blog/ai-transcription-accuracy-claims-investigation
- [23] researched.org.uk/2019/06/24/writing-and-cognitive-load-theory
- [24] su-plus.strathmore.edu (Kenyan Railways case study)
- [25] speechtonote.com/blog/the-challenges-of-accurate-transcription