

Voice Provenance

When sound, belief and disclosure diverge, provenance becomes a separate evidentiary question.

Peter van Poelgeest
RN-02 | Version 1.0 | Published | 29 September 2026

RESEARCH QUESTION

When actual voice origin, perceived origin and disclosed origin can diverge, what can provenance legitimately establish?

ABSTRACT

Recent studies on voice cloning indicate that human origin cannot always be inferred reliably from sound alone, that human and synthetic voices can differ on specific perceptual attributes without supporting a universal “human is better” conclusion, and that source labels can influence evaluation independently of the recording itself.

This note separates actual origin, perceived origin, disclosed origin and provenance. Its claim is deliberately narrow: provenance can make a source claim more inspectable; the cited research does not establish that verified provenance creates trust, comprehension, persuasion or communication effectiveness.

KEY FINDINGS

- 1. Audible origin is becoming less dependable.** In the cited contemporary voice-cloning studies, listeners did not always identify AI-generated recordings reliably.
- 2. Human origin is not a general quality certificate.** Observed differences depend on attributes, prosody, synthesis and context.
- 3. Source information can affect evaluation.** Human/AI labels can change evaluation even when the recording itself is unchanged.[4]

EVIDENCE AND ANALYSIS

1. Human origin is becoming harder to infer from sound alone.

Barrington, Cooper and Farid found that listeners identified AI-generated recordings correctly only around 60% of the time in their experiment.[1] Lavan and colleagues likewise found that some contemporary voice clones could sound approximately as real as human recordings.[2] Neither result means synthetic voices are universally indistinguishable from human ones. Detection depends on the voice, system, material and task. But audible origin can no longer be assumed.

2. Human does not automatically mean better.

Research comparing human and synthetic voices does not support a simple human-good / AI-bad division. Human voices can score higher on qualities such as human-likeness, animateness and emotional richness.[3] At the same time, synthetic voices can respond strongly to prosodic differences[3], and under some conditions have been rated as more trustworthy than neutrally delivered human voices.[5] Performance, synthesis, prosody and context all matter.

3. What listeners are told can change what they hear in the same recording.

Kipp and colleagues compared recordings spoken by humans and highly natural AI voice clones while independently varying whether participants were told the voice was human or AI-generated. The actual human recordings still received higher ratings. But the label also changed evaluation. Recordings labelled human were rated more positively than those labelled AI-generated, independently of the actual source. In that study, AI voices labelled human were even evaluated more positively than human voices labelled AI. The recording had not changed. The source information had. [4]

Where provenance enters

This creates an important distinction between:

DISTINCTION	QUESTION
Actual origin	What produced the recording?
Perceived origin	What does the listener believe produced it?
Disclosed origin	What is the listener told produced it?
Provenance	What evidence supports the claim about where the recording came from?

Those are not necessarily the same thing. Provenance therefore asks a different evidentiary question from whether a voice sounds human or whether a label says it is human.

What provenance does not establish

Verified provenance is not:

- a quality certificate;
- proof that a human voice will be preferred;
- proof of trust;
- proof of understanding;
- proof of persuasion;
- proof that a message landed.

The existing research shows that source information can influence evaluation. It does not yet establish that verified provenance affects evaluation differently from an ordinary source label.

Why this matters

As audible origin becomes less reliable, claims about origin increasingly depend on evidence outside the sound itself. Analytic inference: provenance is relevant here because it can make a source claim more inspectable by providing evidence about origin. That does not establish that receivers notice, value or respond differently to verified provenance.

Does verified provenance change evaluation differently from simply telling someone that a voice is human or synthetic?

WHAT THIS NOTE ESTABLISHES

Evidence level: synthesis of five published studies; no original empirical data.

The cited studies support three bounded conclusions: audible origin is not consistently reliable in the tested contemporary voice-cloning conditions; human and synthetic voices can differ on evaluated attributes without a universal human-better result; and human/AI labels can alter evaluation independently of actual source. The proposition that provenance can make a source claim more inspectable is an analytic inference, not an experimentally established reception effect.

CLAIM BOUNDARY

This Research Note does not claim that listeners can no longer distinguish human from synthetic voices, that human voices are superior, or that verified human provenance creates trust, understanding, persuasion or communication effectiveness.

It also does not establish that verified provenance changes evaluation differently from an ordinary source label. That remains the next research question.

SOURCES & NOTES

1. Barrington, S., Cooper, E. A., & Farid, H. (2025). People are poorly equipped to detect AI-powered voice clones. *Scientific Reports*, 15, 11004. DOI: 10.1038/s41598-025-94170-3. Used for the bounded claim that listeners in the experiment correctly identified AI-generated recordings only around 60% of the time.
2. Lavan, N., Irvine, M., Rosi, V., & McGettigan, C. (2025). Voice clones sound realistic but not (yet) hyperrealistic. *PLOS ONE*, 20(9), e0332692. DOI: 10.1371/journal.pone.0332692. Used for the bounded claim that some contemporary voice clones can sound approximately as real as human recordings.
3. Chen, W., Pell, M. D., & Jiang, X. (2026). Does speech prosody shape social perception equally for AI and human voices? A 16-attribute rating study. *Computers in Human Behavior Reports*, 22, 101142. DOI: 10.1016/j.chbr.2026.101142. Used for the comparison of human and AI voices on attributes including human-likeness, animateness and emotional richness, and for prosody-dependent differences.
4. Kipp, K. H., Hanenberg, J., Klinar, K., Tersek, E., Klemm, A., & Wehrum, H. (2026). The label effect in speech perception: Comparing AI-generated and human voices. *Journal of Interdisciplinary Voice Studies*, 11(1), 9–27. DOI: 10.1386/jivs_00118_1. Used for the label-effect claim: human/AI source information altered evaluation independently of actual source.
5. Maltezou-Papastilianou, C., Scherer, R., & Paulmann, S. (2025). Evaluating trustworthiness across ethnically diverse human and commercial synthesised voices: A comparative study. *Computers in Human Behavior Reports*, 19, 100762. DOI: 10.1016/j.chbr.2025.100762. Used for the bounded trustworthiness comparison in which synthesized voices could outperform neutrally delivered human voices under some conditions.

PUBLICATION INFORMATION

AUTHOR Peter van Poelgeest

PUBLISHER COMMODYN Research

EVIDENCE TYPE Literature-based synthesis; no original empirical data

VERSION 1.0

COMPETING INTERESTS The author also works professionally as a human voice-over artist through DutchVibezzz. This professional activity is relevant to the topic of human and synthetic voice provenance.

AI ASSISTANCE AI tools were used for research support, editorial development and document production. Source selection, interpretation, claim boundaries and publication decisions remain the author's responsibility.

RECOMMENDED CITATION van Poelgeest, P. (2026). Voice Provenance. COMMODYN Research Note RN-02, version 1.0.