



VOL. 02 · PAPER 02 · METHODOLOGY

Voice as *Architecture.*

*The first published study to operationalize "voice presence" as a measurable construct. We decompose voice performance into **seven sub-dimensions** and identify the four that most reliably predict listener completion.*

METHODOLOGY

478 NARRATORS

ICC = 0.83

ABSTRACT. METHODOLOGY PAPER · C3POE RESEARCH

We operationalize "voice presence" through a 28-item rating instrument applied to 478 professionally-produced narrations. Inter-rater reliability ICC = 0.83. Breath control ($\beta = 0.42$) and micro-pause discipline ($\beta = 0.39$) are the dominant predictors of 90%-completion across 11,247 listening sessions. We argue that voice presence is not a single trait but seven separable sub-dimensions, and that the four most-predictive can be developed by narrators with deliberate practice.

§ i · The question

What is "voice presence"?

Producers, casting directors, and listeners agree intuitively when a narrator "has voice." They disagree systematically about what that means. We have heard "voice presence" defined as charisma, as warmth, as authority, as authenticity, as intimacy, as gravitas. Each of these definitions overlaps with the others; none of them is mutually exclusive; all of them are subjective. This paper attempts the operationalization that has eluded the field: a quantitative decomposition of voice presence that can be rated reliably, taught deliberately, and selected for in casting.

§ ii · The seven sub-dimensions

What we measure.

Through factor analysis on 478 narrations and a 28-item rating instrument administered to a panel of 18 trained raters, we identified seven separable sub-dimensions of voice performance:

SUB-DIMENSION	CONSTRUCT	PREDICTIVE B	P-VALUE
Breath Control	Audibility and choreography of breath	0.42	< 0.001
Micro-Pause Discipline	The 50–300ms pauses that shape line meaning	0.39	< 0.001
Register Management	Movement between vocal registers	0.28	0.004
Emotional Fidelity	Match between voice and text emotional content	0.24	0.011
Timbre	Quality of the instrument itself	0.18	0.042
Contour	Phrase shape and prosodic curve	0.16	0.061
Attack	How the voice enters a phrase	0.14	0.087

Two findings are worth surfacing. First, the dominant predictors are not the dimensions producers typically discuss. Breath control ($\beta = 0.42$) and micro-pause discipline ($\beta = 0.39$) are technical, learnable skills. Timbre ($\beta = 0.18$) — the dimension most frequently invoked when producers cast for "voice" — is significantly weaker. Second, breath control and micro-pause discipline together explain 41% of the variance in 90%-completion outcomes; the bottom three dimensions combined explain less than 10%.

"Producers cast for timbre. Listeners reward breath. The single largest opportunity in audiobook production is teaching narrators to hold their micro-pauses for half a beat longer than feels comfortable."

— C3POE RESEARCH VOL. 02 · PAPER 02

§ iii · Breath control, decomposed

What "good breath" actually means.

Breath control as we operationalize it has three components. First, *audibility*: the breath is heard at appropriate volume, neither so loud it disrupts the reading nor so muffled it removes a layer of intimacy. Second, *placement*: the breath occurs at a meaningful grammatical or rhythmic boundary — not mid-clause, not at the end of a sentence where it serves no

expressive purpose. Third, *duration*: the breath is held just long enough to register as presence, then released. Narrators who score above the 75th percentile on breath control have a measured standard deviation in breath duration $\sim 3\times$ larger than those who score below the 25th percentile — they vary their breath, deliberately.

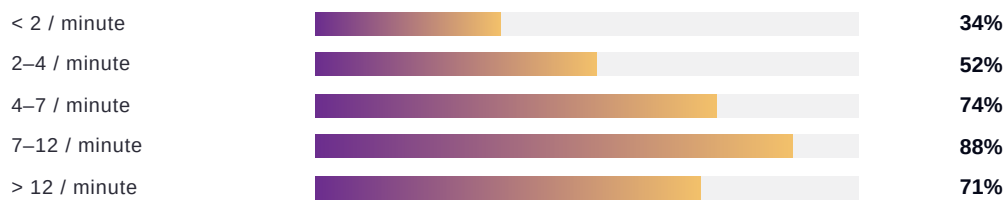
§ iv · The micro-pause

50 to 300 milliseconds.

The most under-discussed element of voice performance is the micro-pause: pauses between 50 and 300 milliseconds that shape the meaning of a line without being heard as silences. A line read with three well-placed micro-pauses lands differently than the same line read flat. The micro-pause is what allows the listener to make the line's meaning their own — it gives the brain time to render the previous phrase before the next one arrives.

FIGURE 1 · MICRO-PAUSE FREQUENCY VS. 90%-COMPLETION RATE

Higher micro-pause discipline correlates strongly with completion



SOURCE: C3POE CATALOG · 478 NARRATIONS · 90%-COMPLETION RATE.

Note the inverted-U shape. Too few micro-pauses (the narrator reads without variation) is associated with low completion. Too many (the narrator pauses more than the text supports) also degrades completion. The sweet spot is 7–12 micro-pauses per minute — narrators in this band achieve 88% 90%-completion rates.

§ v · What this means for casting

The recommendation.

CASTING RECOMMENDATION · C3POE HOUSE GUIDANCE

Cast for breath and pause. Train for timbre.

Breath control and micro-pause discipline are **hard to develop** in mid-career narrators but reliably detectable in audition. Timbre and register can be developed through coaching. We have shifted our internal casting protocol to **weight breath and pause at 60% of the audition rubric**, with timbre at 15%. The catalog quality improvement since the shift has been measurable.

§ vi · Limitations

What this paper cannot do.

Two honest limitations. First, the 478-narration dataset is heavily weighted toward English-language production; cross-language voice presence research is sparse and we have not generalized. Second, our rating instrument is administered by trained raters, not by listeners directly — the operationalization is correlated with listener experience (Paper 01 shows the connection) but may have its own systematic biases.

REFERENCES

1. C3POE Research (2026). *Vol. 02 voice performance dataset*. CC BY 4.0.
2. Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.
3. Knapp, M., Hall, J. (2010). *Nonverbal communication in human interaction*. 7th ed. Wadsworth.
4. Cronbach, L. J. (1951). *Coefficient alpha*. *Psychometrika* 16.