

AI Voice Product Test QA

Synthetic multilingual listening evidence separated from a seductive average score.

12	3	4	5
TEST RUNS	PRODUCT STATES	ISSUES LOGGED	QUALITY DIMENSIONS

Executive result

Three synthetic product states are compared across English and Japanese, single and multi-speaker scenarios, latency, cost, retries and continuity—without averaging away a severe Japanese failure.

What this sample proves

Voice quality can be treated as evidence: audio references, language-specific defects, retry burden, cost and qualitative residue stay linked to the decision.

Use boundary

All product names, audio references, scores and observations are synthetic demonstrations. No named commercial voice system was tested.

State transformation

01 Chaos

Naturalness, pronunciation, speaker identity, emotion, continuity, latency and cost pull in different directions and hide behind one subjective impression.

02 Operation

Scenario matrix → bounded metric definitions → row-level evidence → weighted projection → failure gate → issue and recommendation register.

03 Usable state

A reviewer can choose a bounded pilot, exclude an unsafe use case and identify which pronunciation or continuity test must be repeated.

Uncompressed residual

Human listening, cultural appropriateness, rights clearance and production context cannot be compressed into one score.

Observable proof

- 12 synthetic scenario rows
- English/Japanese separation
- Pass/fail gates before averages
- Severe issue preserved as evidence

Portfolio integrity

Independently constructed self-directed technical work using public or synthetic inputs. Not a reproduced marketplace brief, confidential client engagement, paid result or claim of deployment.