Real people.
Real accents.
Real voice-agent failures.

Real-human validation for voice AI: Southeast Asian second-language English speakers, recruited and managed by one accountable operator.

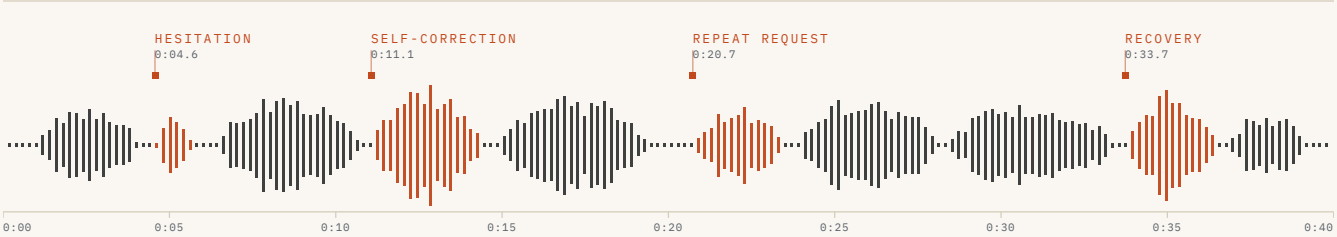

FIG. 01 · ANNOTATED UTTERANCE TIMELINE SECOND-LANGUAGE ENGLISH · LIVE MOBILE CALL

Synthetic testing scales. Real people still surprise the system.

Voice AI teams now test at machine scale. Simulation platforms generate thousands of calls per night, sweep prompts and edge cases, and catch regressions before release. That layer is essential, and BKU does not replace it.

But synthetic callers are built from assumptions about how people speak. Second-language speakers on real phones, in real rooms, are the population most likely to break a voice agent. They are also the population that assumption-based testing reproduces least faithfully.

An agent that passes ten thousand synthetic calls can still fail the humans it was deployed to serve. Those failures surface late, in production, where they are most expensive to find:

a. **Abandoned calls**
Callers give up mid-task, and the work quietly moves to a human channel.

b. **Repeated escalations**
Unrecognized accents and entities route straight to your most expensive resource.

c. **Support cost**
Every misheard name, date, or number becomes a correction loop somewhere downstream.

d. **Quiet churn**
Users who are consistently misunderstood do not file tickets. They stop calling.

BKU complements automated simulation platforms. It does not replace them.

The failure patterns that only surface with real humans.

FIG. 02 · WHAT THAT SOUNDS LIKE

“It’s on— sorry, I mean Thursday, not Tuesday.”

SELF-CORRECTION

“My name is Thiri. T-H-I-R-I.”

SPELLING SEQUENCE

“Zero nine, four five... can you repeat? The line is not so clear.”

REPEAT REQUEST · CHANNEL NOISE

“Okay... okay, never mind. I will just go to the branch.”

ABANDONMENT

Illustrative fragments. Pilot transcripts are delivered with consent, metadata, and structured labels.

a. **Hesitation & self-correction**
Mid-utterance restarts that break intent parsing.

b. **Pronunciation & pacing variance**
Stress, speed, and vowel quality shift within a single call.

c. **Imperfect grammar & code-switching**
Meaning carried by structures the ASR was not tuned for.

d. **Entity restatement**
Names, numbers, dates, and addresses corrected on the fly.

e. **Repeat requests & recovery**
How a caller actually behaves after the agent mishears.

f. **Escalation & abandonment**
The moment a human gives up on the machine.

g. **Device, network & environment**
Real handsets, real networks, real rooms.

A real-human validation layer for the voice AI stack.

BKU offers four services built on one capability: verified second-language speakers, managed end to end.

01

Real-human voice-agent testing

Verified speakers call your voice agent, in staging or against a production-safe line, and work through buyer-defined scenarios. Measurements cover task completion, repeated prompts, entity errors, turn-taking, escalation, and abandonment.

TYPICAL USE · PRE-LAUNCH VALIDATION; REGRESSION CHECKS AFTER MAJOR PROMPT OR MODEL CHANGES.

02

Custom speech-data collection

Scripted reading, spontaneous speech, guided conversation, and call audio, collected to your channel, device, environment, and metadata specification.

TYPICAL USE · ACCENT-SPECIFIC ASR EVALUATION AND TRAINING SETS.

03

Human ground-truth validation

Automated tests generate failure hypotheses. We check which of them reproduce with real speakers before your customers run into them.

TYPICAL USE · SEPARATING THEORETICAL FAILURES FROM PRODUCTION RISK BEFORE ROADMAP DECISIONS.

04

Regional speaker operations

Recruitment, identity verification, briefing, consent, scheduling, compensation, and re-recording, run by one accountable operator rather than an anonymous crowd.

TYPICAL USE · A STANDING CAPABILITY BEHIND RECURRING RELEASES.

Start small. Test something real.

Every engagement starts from the buyer's specification. The configuration below shows shape and scale for one example pilot.

ILLUSTRATIVE CONFIGURATION	EXAMPLE ONLY
SPEAKERS	10–20, verified
SCENARIOS	Buyer-defined
CALLS	Live or staged voice-agent calls
LABELS	Structured error taxonomy
METADATA	Speaker · device · environment
REPORT	Executive summary of confirmed risks

HOW IT RUNS

01 **You define**
Target population, workflow, channel, and what failure costs you.

02 **We propose**
Speaker profile, scenario set, deliverables, timeline, and terms.

03 **We run**
Verified speakers execute; BKU manages briefing, quality, and re-records.

04 **You decide**
Findings arrive as data you can act on, whether the next step is a fix, a re-test, or a launch.

Pricing follows scope. A pilot is quoted as a fixed engagement, with no platform fee and no long-term commitment.

Deliverables designed for engineering and procurement alike.

01	Authorized audio recordings
02	Ground-truth transcripts
03	Scenario outcomes (success and failure)
04	Repeated-prompt counts
05	Name, date, number, and address errors
06	Interruption and turn-taking failures
07	Escalation and abandonment behavior
08	Participant and environment metadata
09	Structured failure taxonomy
10	Executive summary

SAMPLE FAILURE TAXONOMY · FORMAT ILLUSTRATION

CATEGORY	EXAMPLE LABEL	WHAT IT TELLS YOU
ENTITY CAPTURE	Date restated twice before confirmation	Booking flows misread rescheduling requests.
TURN-TAKING	Agent interrupts during hesitation pause	Barge-in threshold tuned only for native pacing.
RECOVERY	Caller abandons after second repeat request	Error loops drive real-world drop-off.
CHANNEL	Task failure on mobile with street noise	Production channel missing from the test set.

Sample format only. The taxonomy is agreed with the buyer before collection begins.

Formats and schemas are agreed before collection begins, including audio, transcript, and label conventions.

Built around managing real people across borders.

Before voice AI, BKU's core work was Myanmar talent operations: recruitment, identity verification, multilingual briefing, documentation, field coordination, and compensation management.

Real-human voice validation runs on the same muscles, applied to a new kind of deliverable.

OPERATING CAPABILITIES

- PARTICIPANT RECRUITMENT
- IDENTITY VERIFICATION
- MULTILINGUAL BRIEFING
- CONSENT & DOCUMENTATION
- SCHEDULING & FIELD COORDINATION
- COMPENSATION MANAGEMENT
- QUALITY CONTROL & RE-RECORDING

INITIAL SPEAKER COVERAGE

ACCENT	Myanmar-accented English
PROFICIENCY	A2–B2 (CEFR)
BASE	Thailand · Myanmar
CHANNELS	Mobile and VoIP
ENVIRONMENTS	Quiet and ordinary settings
SPEECH	Scripted and spontaneous

Coverage follows buyer demand, and the model extends across Southeast Asia's second-language English populations.

The long-term aim: real-human voice validation infrastructure for emerging Asia.

Consent and permitted use are part of the product.

BKU is built around responsible collection principles. Terms are project-specific, in writing, and agreed before the first recording.

i. **Explicit, project-specific consent**
Participants know the buyer's use case before recording starts.

ii. **Buyer-defined permitted use**
The license states what the data may train, test, or evaluate.

iii. **No voice cloning without separate authorization**
Cloning never rides in on a general consent form.

iv. **No reuse beyond the agreed scope**
One project's audio is not another project's training set.

v. **Withdrawal procedures**
Participants can withdraw; affected data is flagged and handled per terms.

vi. **Secure transfer, retention, and deletion**
Handled to the buyer's retention schedule and transfer requirements.

vii. **No employment linkage**
Participation never affects recruitment or job opportunities.

When we cannot promise something cleanly, we say so.

Five ways teams work with BKU.

01 ## Regional collection partner

Speech data collected in Southeast Asia to your program's specification, for data companies running multi-region collection programs.

02 ## Human QA partner

A recurring real-human validation layer behind your automated testing platform. Your simulations find candidate failures and BKU confirms them with real speakers.

03 ## Accent coverage partner

Myanmar-accented English now, with Southeast Asian expansion to follow, for your accent and locale coverage matrix.

04 ## Custom data supplier

One-off datasets built to a fixed specification and schema, delivered with consent documentation and metadata.

05 ## Release-validation partner

A standing, verified speaker pool that re-tests your agent at every major release, giving you regression testing with real humans.

All five start the same way: a small, buyer-defined pilot.

Tell us who your voice agent needs to understand.

Describe the workflow, the target speakers, and what failure would cost you.
BKU will propose the smallest real-human pilot that answers the question.

WHAT TO SEND US	WHAT COMES BACK
The workflow you are testing	The smallest credible pilot proposal
The population your agent serves	Speaker profile and scenario set
Channel and environment requirements	Deliverables, timeline, and terms
What failure costs you	A reply from the person who runs the operation

Yuhei Yamada

BKU · OPERATIONS

yamada@bku1.com

bku1.com

Bangkok · Myanmar operations

Start with a buyer-defined pilot. Scale only after the workflow is proven.