rubrex.

HUMAN EVALUATION

Audio evaluation

How a human reviewer decides whether a spoken instruction is ready to use.

PREPARED FOR APEX PROGRAMMERS

Purpose of this draft

Please review the proposed criteria, evidence fields and decision rules.

READING GUIDE

02-03	Scope and decision rubric
04	Results by work type
05-06	Two worked findings
07	Open issues and owners
08-09	Complete sample register

REVIEW DESIGN

What the review covers

A voice assistant reads short instructions aloud. A listener must be able to act on the words they hear, even when one small word changes the task.

Unit of review

The review unit is one clip paired with one script. Each of the four task groups contains ten clip records.

Task group	What to inspect
Device instructions	Listen for action verbs, units and negation.
Appointments	Check dates, times and pauses between options.
Account prompts	Check codes, amounts and the order of actions.
Navigation	Check destinations, turn direction and distance.

Review procedure

The reviewer compares the instruction with the script, then listens again at the disputed timestamp. Harmless pacing and accent differences are acceptable when the meaning stays clear.

Coverage boundary

The sample covers English instructions read in isolation. Conversation timing, background noise and speaker preference would need separate test material.

REVIEW DESIGN

How verdicts are assigned

Use the same decision rules for every item. Record the evidence that led to the verdict so a second reviewer can follow the reasoning.

Meaning

Compare actions, negations, quantities and units.

Revise when the listener would carry out a different action.

Intelligibility

Listen for missing syllables, clipped endings and joined options.

Revise when a required word cannot be recovered by ear.

Reference clarity

Check if the written script has one defensible reading.

Escalate when several readings are plausible.

Acceptance

Listen once more in context after checking the script.

Accept when all required words are intelligible and meaning is preserved.

When reviewers disagree

Compare the observations against the original reference or acceptance criterion. Record the disputed point and the final rationale before changing the verdict.

SAMPLE RESULTS

Distribution of verdicts

This example contains 40 clips. Each item contributes exactly one verdict. Counts below reconcile to the item register on pages 8 and 9.

30
Accept / 75%

8
Revise / 20%

2
Escalate / 5%

Task group	Accept	Revise	Escalate
Device instructions	7	3	0
Appointments	7	2	1
Account prompts	8	2	0
Navigation	8	1	1
Total	30	8	2

Reading the results

Ten clips remain open. Eight need a corrected recording. Two need a decision about the script before a reviewer can judge the audio.

What to inspect first

Device instructions account for three revisions. A single missing word can reverse a command, so prioritize meaning-changing cases ahead of cosmetic rerecording.

WORKED FINDING / AUD-007

One missing word reverses the instruction

A listener following the spoken version would restart a device that should remain on.

DEVICE INSTRUCTIONS / REVISE

Reference script

During setup, do not restart the device.

Heard at 00:04-00:06

During setup, restart the device.

Review decision

Revise. The missing negation changes the requested action even though the audio is clear.

Correction and retest

Re-record the full sentence. Listen across the words immediately before and after the correction and confirm that a fresh listener hears the prohibition.

> **Reviewer note**
>
> A useful note quotes both readings and identifies the changed action. A comment such as 'sounds wrong' gives the recording team too little to work with.

WORKED FINDING / AUD-019

An ambiguous date needs an owner

A reviewer cannot determine the intended reading from the supplied script alone.

APPOINTMENTS / ESCALATE

Reference script

Your appointment is on 4/15.

Heard at 00:11-00:12

The token is spoken as 'four fifteen.'

Review decision

Escalate. The project needs a rule for rendering numeric date tokens aloud. The script gives no target pronunciation.

Resolution and retest

Ask the script owner for the preferred spoken date. Store that reference with the clip, then compare a new reading against it.

> **Reviewer note**
>
> Keep a reference question separate from a recording defect. Recording again before deciding the preferred reading could produce the same disagreement.

FOLLOW-UP

Open issues and owners

This queue lists every unresolved item in the sample. The owner shown is a proposed role for the next action.

ID	Issue	Verdict	Proposed owner
AUD-005	Clipped final word	Revise	Recording editor
AUD-007	Negation omitted	Revise	Recording editor
AUD-010	Duration changed	Revise	Recording editor
AUD-012	Time misread	Revise	Recording editor
AUD-017	Pause joins choices	Revise	Recording editor
AUD-019	Date ambiguous	Escalate	Script owner
AUD-025	Code digit lost	Revise	Recording editor
AUD-026	Amount misread	Revise	Recording editor
AUD-032	Name unspecified	Escalate	Script owner
AUD-036	Distance unit lost	Revise	Recording editor

Routing order

Start with the missing negation, duration, time, code, amount and distance. For the date and street name, ask the script owner for the intended reading before requesting a recording.

Close an item

Keep the original verdict, add the corrected asset or build reference, and record the retest result. Recalculate the summary only after the reviewer confirms the change.

SAMPLE REGISTER

Records AUD-001 to AUD-020

This half of the register covers device instructions and appointments. Every row is a first-pass decision.

ID	Task	Verdict	Review note
AUD-001	Start the timer	Accept	Clear reading
AUD-002	Pause playback	Accept	Clear reading
AUD-003	Set volume to five	Accept	Clear reading
AUD-004	Restart the app	Accept	Clear reading
AUD-005	Mute the speaker	Revise	Clipped final word
AUD-006	Select input two	Accept	Clear reading
AUD-007	Keep the device on	Revise	Negation omitted
AUD-008	Open the cover	Accept	Clear reading
AUD-009	Reset pairing	Accept	Clear reading
AUD-010	Hold for ten seconds	Revise	Duration changed
AUD-011	Confirm Monday	Accept	Clear reading
AUD-012	Read the time slot	Revise	Time misread
AUD-013	Choose an afternoon	Accept	Clear reading
AUD-014	Cancel the visit	Accept	Clear reading
AUD-015	Read back phone number	Accept	Clear reading
AUD-016	Move to Thursday	Accept	Clear reading
AUD-017	Choose one of two dates	Revise	Pause joins choices
AUD-018	Confirm the clinic	Accept	Clear reading
AUD-019	Read the date token	Escalate	Date ambiguous
AUD-020	Check arrival time	Accept	Clear reading

Register subtotal

Accept: 14 / Revise: 5 / Escalate: 1 / Total: 20

SAMPLE REGISTER

Records AUD-021 to AUD-040

This half of the register covers account prompts and navigation. Every row is a first-pass decision.

ID	Task	Verdict	Review note
AUD-021	Read the balance	Accept	Clear reading
AUD-022	Confirm last four digits	Accept	Clear reading
AUD-023	Choose a language	Accept	Clear reading
AUD-024	Enter the PIN	Accept	Clear reading
AUD-025	Read the support code	Revise	Code digit lost
AUD-026	Repeat the amount	Revise	Amount misread
AUD-027	Confirm sign-out	Accept	Clear reading
AUD-028	Update the address	Accept	Clear reading
AUD-029	Read the account label	Accept	Clear reading
AUD-030	Confirm the new email	Accept	Clear reading
AUD-031	Take the next left	Accept	Clear reading
AUD-032	Read the street name	Escalate	Name unspecified
AUD-033	Continue straight	Accept	Clear reading
AUD-034	Exit at gate three	Accept	Clear reading
AUD-035	Announce arrival	Accept	Clear reading
AUD-036	Read the distance	Revise	Distance unit lost
AUD-037	Repeat the landmark	Accept	Clear reading
AUD-038	Cross at the signal	Accept	Clear reading
AUD-039	Name the terminal	Accept	Clear reading
AUD-040	Choose the north exit	Accept	Clear reading

Register subtotal

Accept: 16 / Revise: 3 / Escalate: 1 / Total: 20