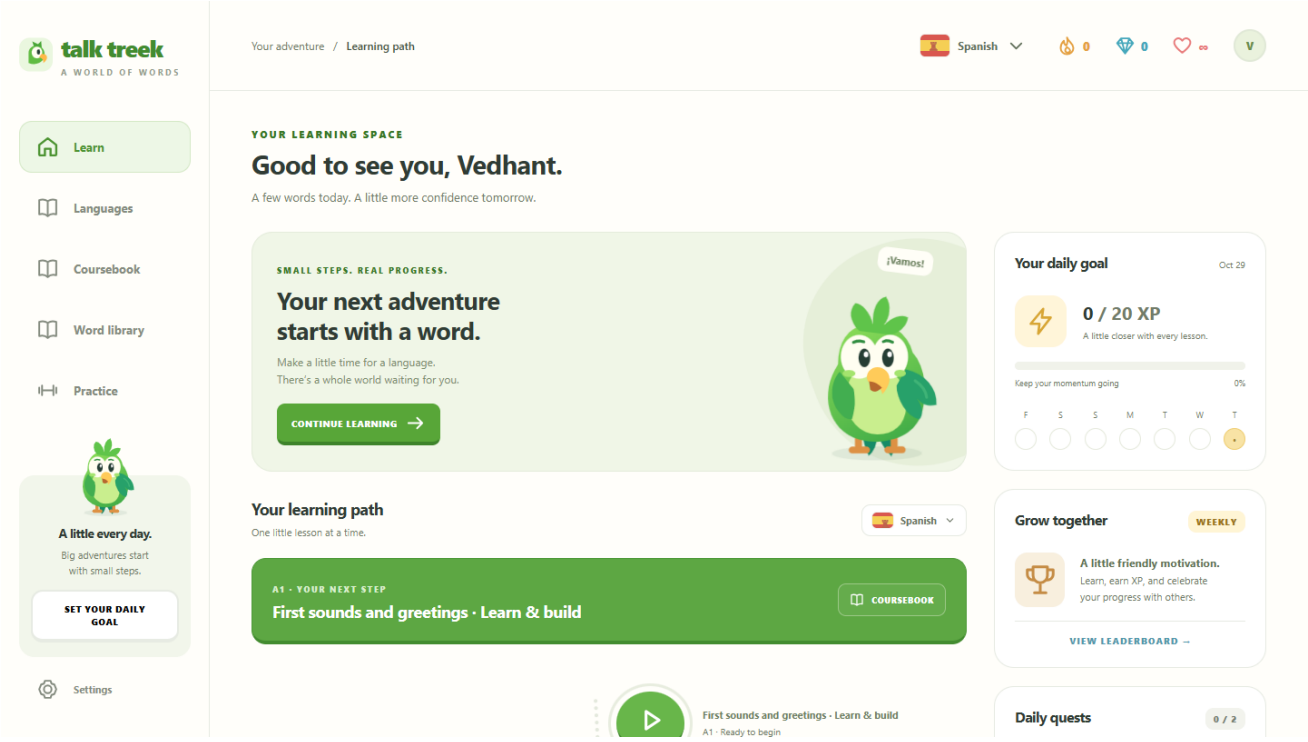
Talk Treek
Language learning, connected.

Project documentation and technical overview

Created by Vedhant Khajuria

29 October 2026



A parrot-themed language-learning platform combining progressive lessons, vocabulary reference, practice activities and personal progress tracking. Built for PHP and MySQL hosting through cPanel.

Project preview: https://talktreek.vedhant.in

Main website: https://vedhant.in

LinkedIn: https://www.linkedin.com/in/vedhantkhajuriaofficial

GitHub: https://github.com/vedhant-khajuria

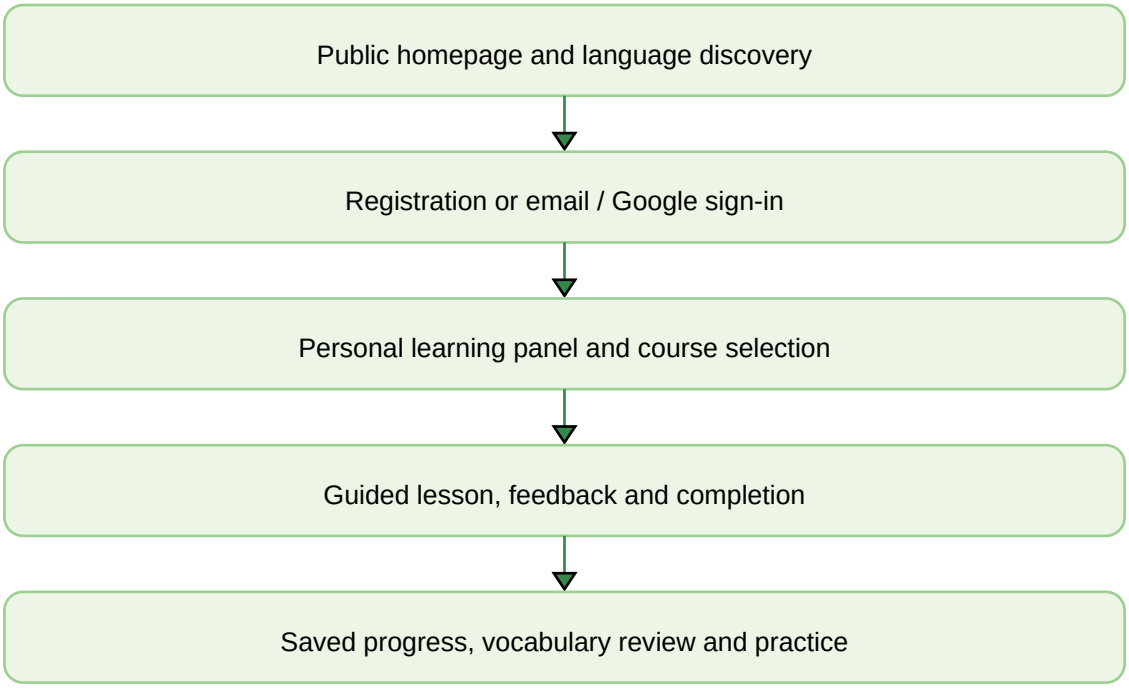
Contact: hello@vedhant.in

Project scope and experience

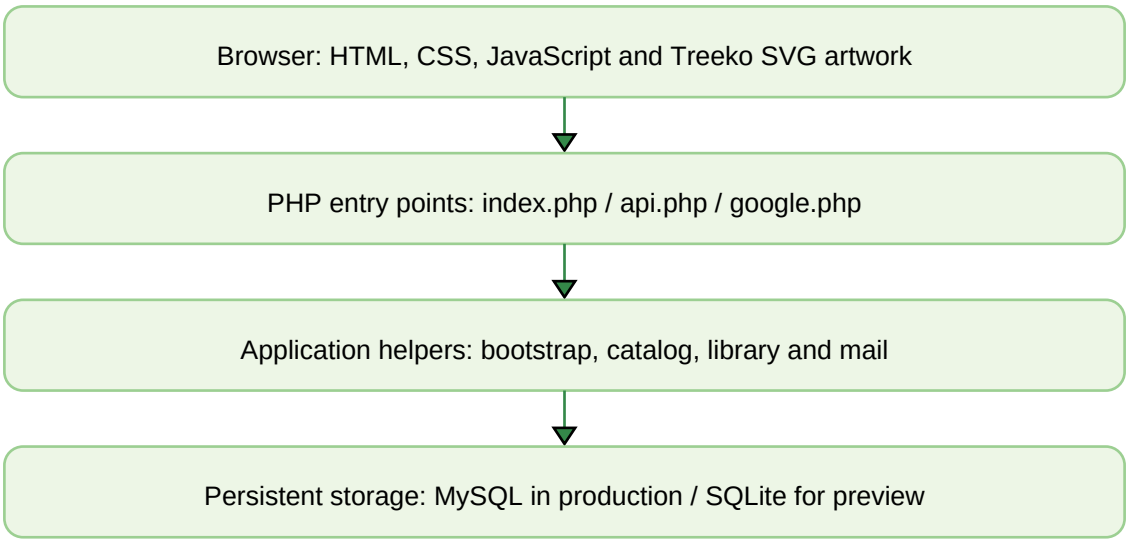
Talk Treek supports English, German, French, Italian, Spanish, Japanese and Hindi. English is the instruction language. Treeko, the illustrated parrot, provides a consistent visual identity throughout the learning experience.

Area	Current implementation
Course structure	24 chapters and 48 new modules per language, with six retained supplementary lessons per language.
Content totals	378 lessons and 3,024 exercises; repeated activities reinforce expressions rather than representing unique phrases.
Word library	3,000 dictionary entries per language: 21,000 in total. Search, save, mark familiar and practise.
Activities	Multiple choice, writing, listening, speaking transcription and word ordering.
Motivation	Daily goals, XP, streaks, lesson prerequisites, quests and weekly leaderboard.
Administration	Learner management, activity statistics and lesson/exercise editing with publishing controls.

How a learner moves through the product



System architecture



External connections

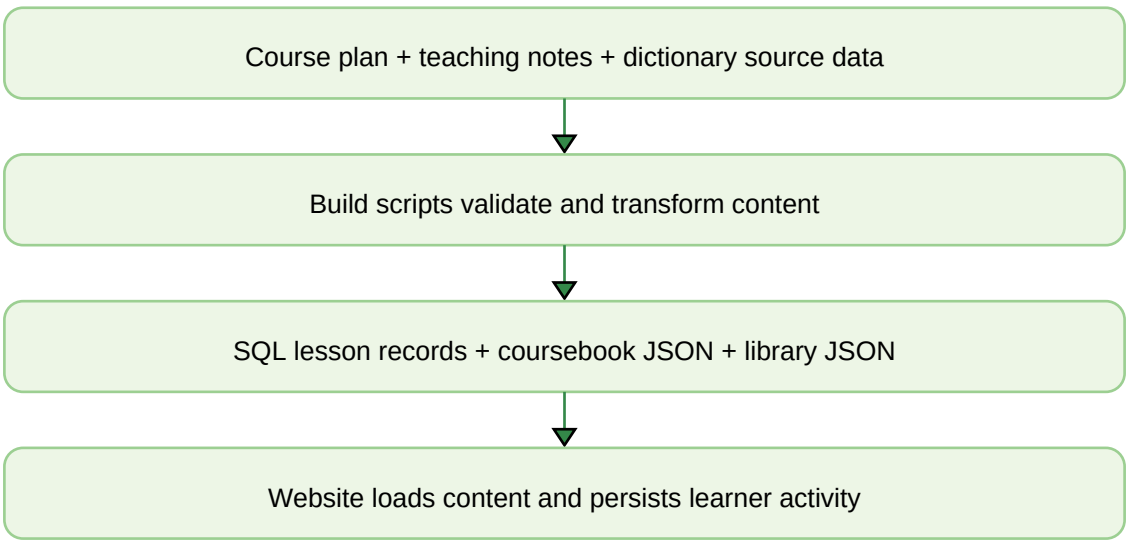
Connection	Purpose and boundary
Google OAuth	google.php performs the server-side authorization-code flow. Client ID, secret and exact callback URL are configured by the site owner.
SMTP server	app/mail.php delivers password-reset messages through the configured mailbox and SMTP provider.
Browser speech	Web Speech APIs provide text-to-speech and optional speech recognition. Browser support and permissions affect availability.
Static JSON assets	Coursebook explanations and vocabulary dictionaries are loaded by the browser from assets/coursebooks and assets/library.

The browser requests JSON from api.php. PHP validates the session and input, executes prepared database queries and returns application results. Production hosting does not require Node.js, Composer or a server-side frontend build.

Technology and source map

Layer / file	Responsibility
PHP 8.2+	Routing, authentication, answer checks, progression and integration endpoints.
MySQL / MariaDB	Users, lessons, progress, active runs and operational records. PDO provides database access.
HTML / CSS / JavaScript	Responsive screens, navigation, activities, animation and speech controls.
index.php / assets/app.js	Page metadata, shared UI, homepage, learner panel and account screens.
assets/lesson-player.js	Question rendering, browser speech, answer feedback and result screens.
assets/library.js / app/library.php	Coursebook rendering, search, vocabulary workouts and saved-word persistence.
assets/admin.js	Admin statistics, learner controls and content editing.
app/bootstrap.php	Sessions, validation, database helpers, authentication and rate limits.
tools/build_curriculum.py	Builds coursebooks, lesson content and SQL output.
tools/build_library.py	Transforms dictionary sources into searchable language-specific JSON.
tools/package.py	Builds an allowlisted cPanel ZIP excluding local tools and preview credentials.

Content pipeline



Database relationship map

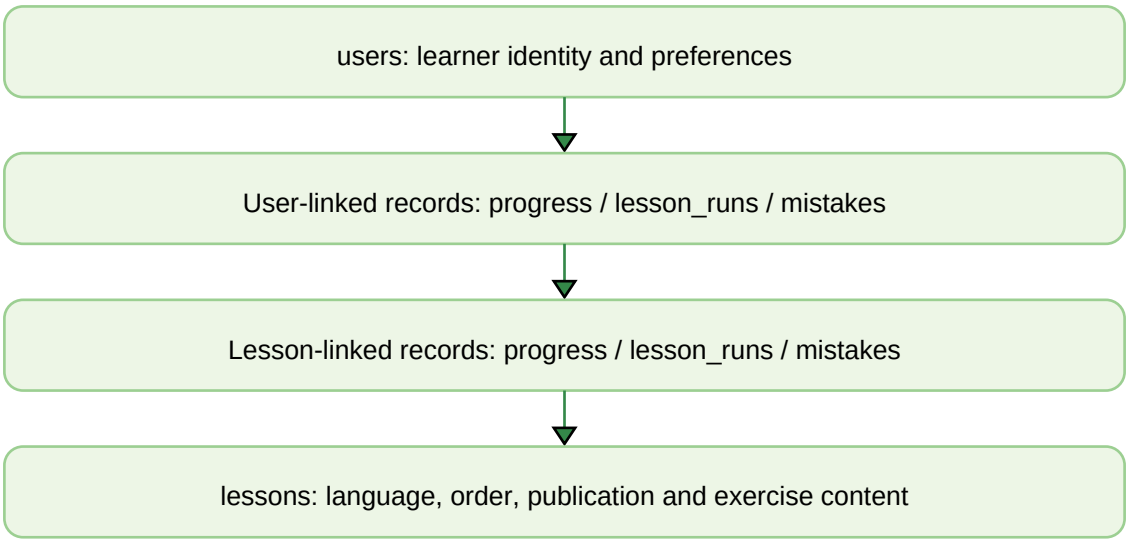
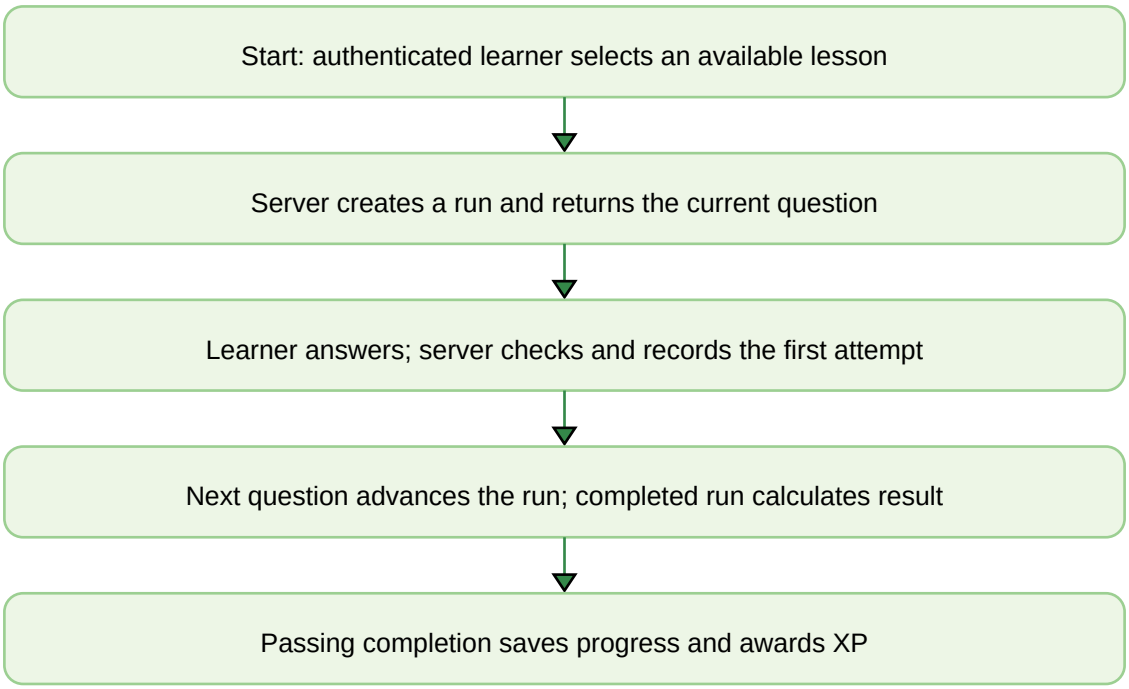
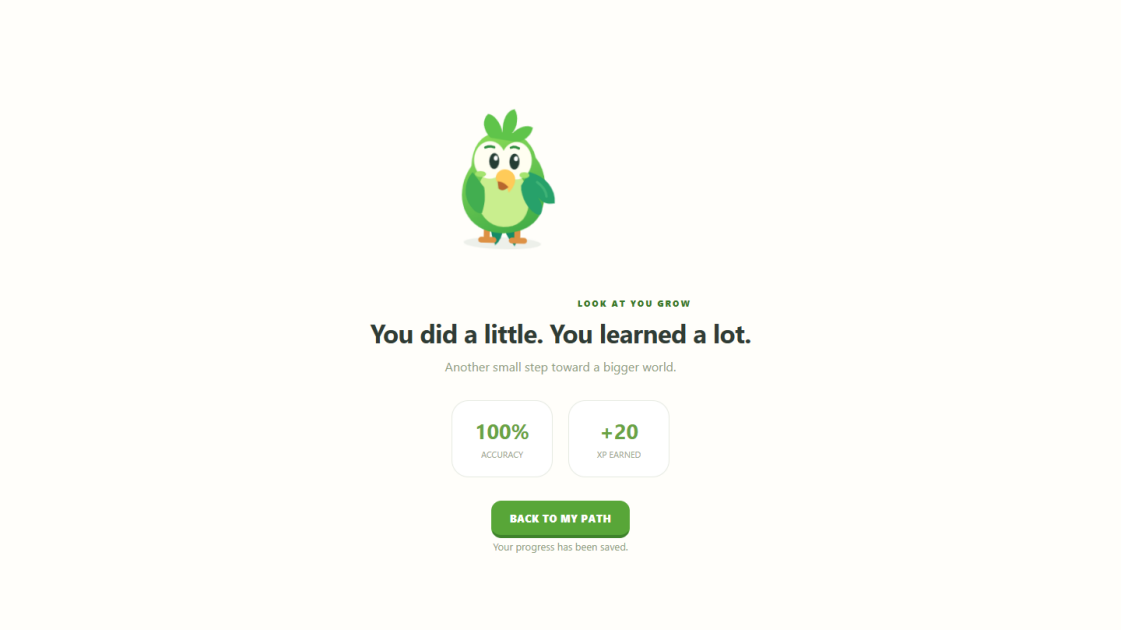


Table	Role and relationship
users	Parent learner record; password hash, language and learning preferences.
lessons	Published or draft learning modules and their exercise data.
progress	Connects learners to lesson completion and performance.
lesson_runs	Tracks an active learner attempt and question position.
mistakes	Associates learner mistakes with lessons for later practice.
password_resets	User-linked, expiring reset-token hashes.
vocabulary_progress	User-linked status for a language and dictionary word ID; dictionary entries themselves live in JSON files.
rate_limits / audit_log	Operational records for request throttling and audit events.

This is a conceptual relationship map, not a column-by-column schema. database.sql is the authoritative fresh-install schema. Foreign-key cascades remove dependent learning records when a learner account is deleted.

Working lesson and progress flow

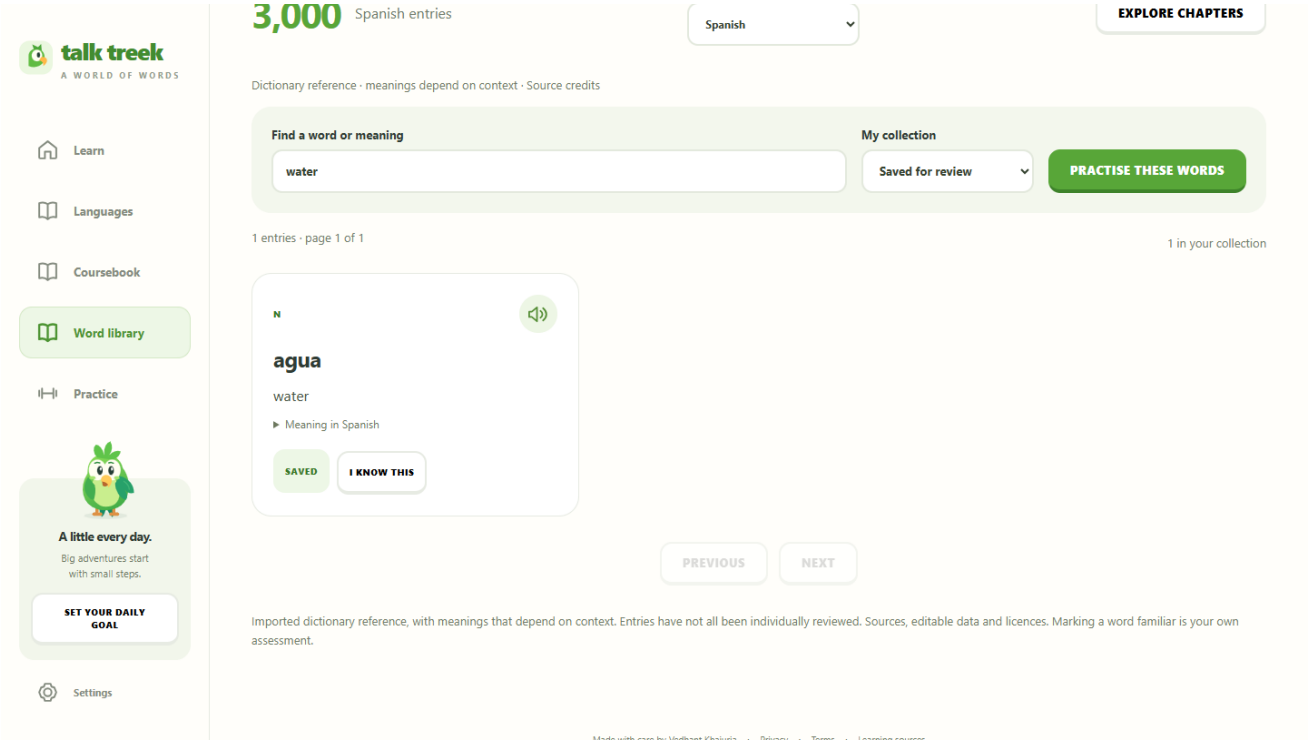




The recorded local demonstration completed all eight questions with 100% accuracy and earned 20 XP. A first-time pass awards 20 XP; a successful review awards 5 XP. The passing threshold is 70%, with a daily earned-XP cap of 500. Streaks and weekly activity use UTC.

Choice questions respond immediately when selected. A learner can explore another option after an error, while the first answer remains the scored attempt. Typed answers use normalization and approximate matching; this is not a guarantee of phonetic equivalence.

Vocabulary, speech and learning design



Reference and review

Learners search native words, meanings or readings, then save useful entries or mark them familiar. Saved status is stored per learner. Vocabulary workouts are unscored and do not award XP. Each course also includes a 40-item beginner toolkit for numbers, weekdays, colours and pronouns.

From first words to advanced topics

The chapter sequence introduces sounds, greetings, everyday vocabulary, questions, routines and travel before moving into narration, comparison, opinions, conditions, negotiation and presentations. Each chapter combines explanations, model expressions and independent speaking missions.

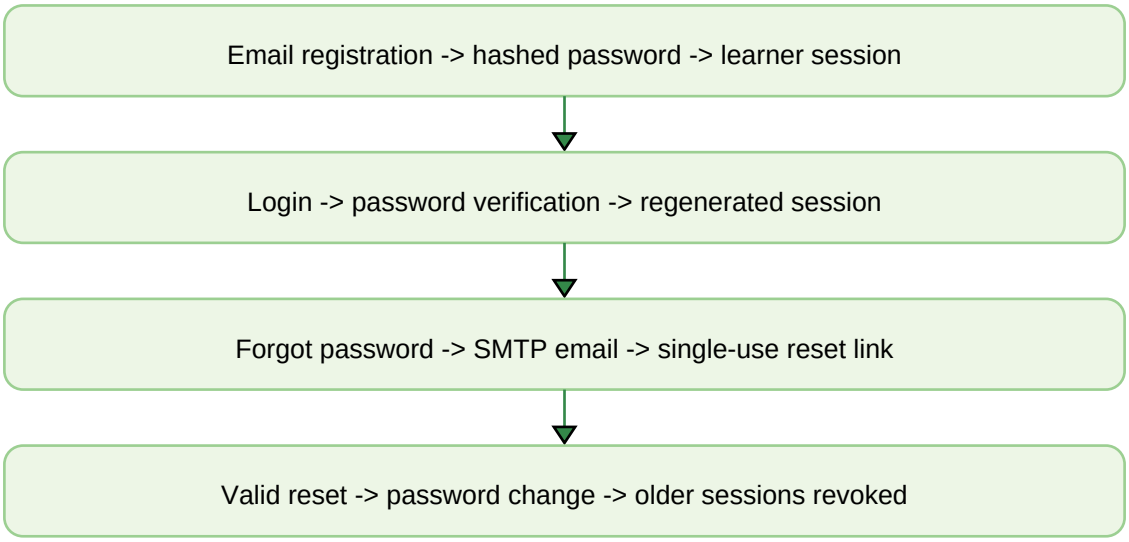
Speech implementation

Speech synthesis reads words through device/browser voices. Optional recognition transcribes microphone input for answer checking; it needs a compatible browser, permission and sometimes network access. Hindi is displayed in Roman letters while native text supports Hindi speech. The platform does not store microphone recordings, although a browser speech provider may process audio.

Stage labels are indicative. Imported dictionary entries still need expert review, and the course is not a validated complete fluency programme. Open speaking missions are self-assessed; speech recognition checks recognized words rather than phoneme-level pronunciation.

Authentication, security and administration

Account access



Google login is an optional separate path using OpenID scopes openid, email and profile. The callback checks OAuth state and requires a verified email. Existing password identities are not silently merged with Google identities.

Implemented safeguards

Prepared SQL statements, output escaping, CSRF tokens, session regeneration, production secure cookies, request throttling, a Content Security Policy and server-enforced progression are included. Reset links expire after 30 minutes and only token hashes are stored. These controls still require production testing and independent security review.

Administrator workflow

The admin signs in separately at ?page=admin using admin@vedhant.in and the credential configured in config.php. The panel provides overview statistics, learner search, disable/enable and deletion controls, and lesson editing with draft/publish management. Administrator sessions expire after two hours and configured password changes revoke them.

Coursebook guide prose is maintained in content files. Lesson editing and dictionary import are distinct workflows: importing generated content can replace edits to generated lesson IDs, so back up and review migrations before applying them.

Deployment and integration guide

Step	Action
1. Hosting	Create talktreek.vedhant.in, enable SSL and HTTPS, and select PHP 8.2+ with pdo_mysql, mbstring, curl and openssl.
2. Database	Create the MySQL database/user in cPanel. Import database.sql into an empty database through phpMyAdmin.
3. Existing installation	Back up first; use curriculum-update.sql for the original existing schema rather than importing the fresh-install file.
4. Upload	Extract the cPanel ZIP directly into the subdomain document root. Include .htaccess and remove the uploaded archive after extraction.
5. Configuration	Set site_url, database settings, production environment and an administrator password of at least 12 characters. Keep secrets private.
6. Google	Create a Web application OAuth client; set redirect URI to https://talktreek.vedhant.in/google.php and configure its ID and secret.
7. Email	Create noreply@vedhant.in; copy the hosting provider SMTP host, username, password, port and encryption into config.php.
8. Acceptance	Test registration, login, reset email, lesson completion, saved vocabulary and administrator controls on the actual host.

Operations and discoverability

Back up database and configuration, monitor PHP error logs and keep hosting software current. Verify access restrictions for internal files. Canonical URLs, page metadata, robots.txt and an XML sitemap support discovery; they do not guarantee ranking.

Local preview uses a separate SQLite database and the PHP development server on port 8089. Screenshots in this report are from that working local build. The requested preview URL is a project reference; this report does not certify its current public availability.

Validation, next steps and credits

Evidence from this project

Content validation checked dictionary counts, chapter and lesson references, exercise choices, Hindi audio mappings and SQLite foreign keys. API checks covered all seven course catalogs, initial lesson answers, repeat-answer rejection, vocabulary save/update/remove and invalid word IDs. PHP and JavaScript syntax checks were run for changed files.

The UI demonstration verified email login, the learner panel, a complete eight-question lesson, 100% result with 20 XP, saved vocabulary and profile screens. Live SMTP delivery, production Google OAuth, microphone recognition on every device and production database behaviour are separate deployment checks.

Recommended next development stages

1. Native-speaker review of dictionaries, lesson translations and accepted variants.
2. More authentic listening and reading materials and broader grammar practice.
3. Cross-device speech, accessibility and responsive-layout testing.
4. Production security review, backup recovery drills and integration monitoring.
5. Learner feedback and outcome-based course refinement.

Sources and ownership

Technical descriptions are based on this repository: README.md, database.sql, api.php, application helpers, UI scripts and content builders. Dictionary provenance is bundled in assets/licenses/sources.json and exposed through the Learning sources page. Sources include WordNet, WikDict and FreeDict; retain their included attribution and licence requirements when distributing the project.

Vedhant Khajuria

https://vedhant.in

https://talktreek.vedhant.in

https://www.linkedin.com/in/vedhantkhajuriaofficial

https://github.com/vedhant-khajuria

hello@vedhant.in