

User Guide - IntonTrainer 2-Ch

Table of Contents

User Guide - IntonTrainer 2-Ch	2
1 What IntonTrainer 2-Ch is	2
2 Main capabilities	3
2.1 Practice with reference patterns	3
2.2 Record your voice directly from the microphone	3
2.3 Automatic speech detection	3
2.4 Three different training modes	3
2.5 Compare audio and visual details	4
2.6 Review saved attempts	4
3 Typical workflow	4
4 Main screens and what they are used for	4
4.1 Home screen	4
4.2 Template Categories screen	5
4.3 Template Files screen	6
4.4 Training screen	7
4.5 Advanced Analysis screen	8
4.6 Records screen	9
4.7 Settings screen	10
5 Training modes available in the app	11
5.1 Auto mode	11
5.2 Guided mode	12
5.3 Manual mode	12
5.4 Difference between the modes	13
5.5 Which mode to use in different cases	13
5.6 Practical recommendation	14
6 Metrics and progress indicators on each screen	14
6.1 Home screen metrics	14
6.2 Template category and template file list metrics	15
6.3 Training screen metrics	15
6.4 Records screen metrics	16
6.5 Visual curves are diagnostic metrics, not only decoration	16
6.6 How training results are fixed	17
7 Recommended training routine	18
8 Settings you can safely change	18
8.1 Language	19
8.2 Theme	19

8.3 Primary Color	19
8.4 Show Navigation Menu	19
8.5 Font Size	19
8.6 Guided mode	19
8.7 Auto Stop Recording	19
8.8 Autocalibrate before recording	19
8.9 Play a beep to start recording	19
8.10 Play the ready beep signal	20
9 Practical tips for new users	20
9.1 Start simple	20
9.2 Use a high-quality microphone	20
9.3 Prefer calibration after environment changes	20
9.4 Use guided mode when learning the pattern	20
9.5 Use advanced analysis only after the score is already visible	20
9.6 Keep the same room and mic setup when comparing progress	20
10 Troubleshooting checklist	20
11 Best practice summary	21
12 Final recommendation	21
13 License	21
14 Third-Party Libraries	21
15 Authors	22

User Guide - IntonTrainer 2-Ch

This guide explains how to use IntonTrainer 2-Ch in a practical, task-oriented way. It focuses on what the application does, how the main workflow works, which settings are safe to change, and which settings are better left alone unless you are troubleshooting a specific audio issue.

1 What IntonTrainer 2-Ch is

IntonTrainer 2-Ch is a desktop application for practicing Chinese tones using reference audio templates (patterns).

In Chinese, there are four basic tones:

- First tone (high and level): represented by the symbol $\bar{}$ (e.g., mā).
- Second tone (rising): represented by the symbol $\acute{}$ (e.g., má).
- Third tone (falling-rising): represented by the symbol $\check{}$ (e.g., mǎ).
- Fourth tone (sharply falling): represented by the symbol $\grave{}$ (e.g., mà).

The meaning of a word can change depending on the tone used, for example:

- 妈 (mā) - mother (first tone).
- 麻 (má) - hemp (second tone).
- 马 (mǎ) - horse (third tone).
- 骂 (mà) - to scold (fourth tone).

The application provides the set of reference audio templates:

- 12 monosyllabic Chinese words with their semantic meaning for each of the four tones,
- sets of phrases made up of various combinations of these words,
- 40 audio examples of words and phrases frequently used in communication practice,
- 9 audio examples of simple everyday multi-word phrases.

The application helps you:

- listen to a reference phrase,
- record your own attempt,
- compare your pitch contour and timing with the reference,
- review a similarity score,
- revisit previous audio attempts for analysis.

In short, the app is designed for guided practice and audio-based self-correction.

2 Main capabilities

IntonTrainer 2-Ch includes the following core capabilities:

2.1 Practice with reference patterns

You can open a template phrase, listen to the reference recording, and try to match its melody (tone movement) with your pronunciation.

2.2 Record your voice directly from the microphone

The app can capture live microphone input and analyze it against the reference wave file.

For reliable results, use a high-quality microphone — a dedicated external mic or a good built-in laptop microphone. Cheap or noisy mics distort the recorded pitch contour and make similarity scores less trustworthy.

2.3 Automatic speech detection

If automatic start-stop is enabled, the application can detect when you start speaking and end the recording automatically.

2.4 Three different training modes

Available training modes:

- Automatic mode
- Guided mode
- Manual mode

The user selects a training mode in accordance with the recommendations in Section 5.4.

2.5 Compare audio and visual details

The app provides:

- waveform display,
- pitch contour display,
- amplitude and spectral diagnostics,
- DTW (Dynamic Time Warping) alignment views,
- a final similarity score.

2.6 Review saved attempts

You can open previous records and compare your result with the template you are interested in.

3 Typical workflow

A normal training session usually follows this sequence:

1. Open the app.
2. Go to the Home screen.
3. Start training.
4. Select a category and then a template file.
5. Listen to the reference sample.
6. Record your own attempt.
7. View the comparative tone movement curves and the result of the similarity of the curves in percentages.
8. Repeat the exercise if needed.

This is the recommended way to use the app for regular practice.

4 Main screens and what they are used for

4.1 Home screen

The Home screen is the starting point.

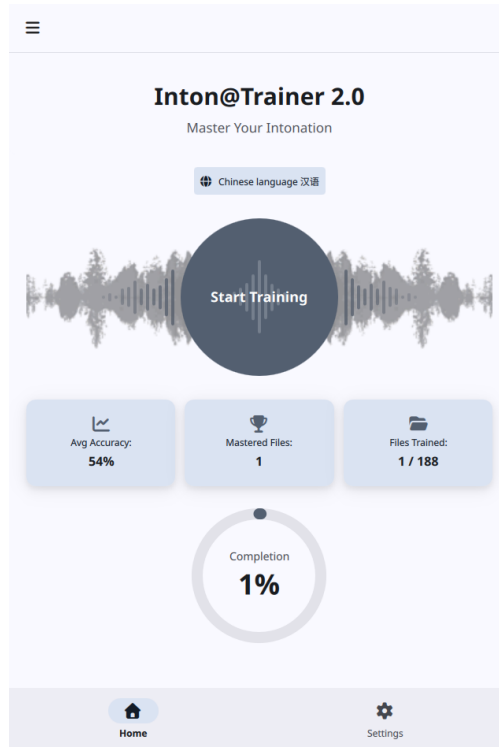


Figure 1: Home screen

It provides you with information:

- about your overall status,
- about the language you're learning,
- about the entry point for starting training,
- a quick overview of your training progress.

Use this screen when you are ready to begin a new training session.

4.2 Template Categories screen

This screen lets you browse the available pattern folders.

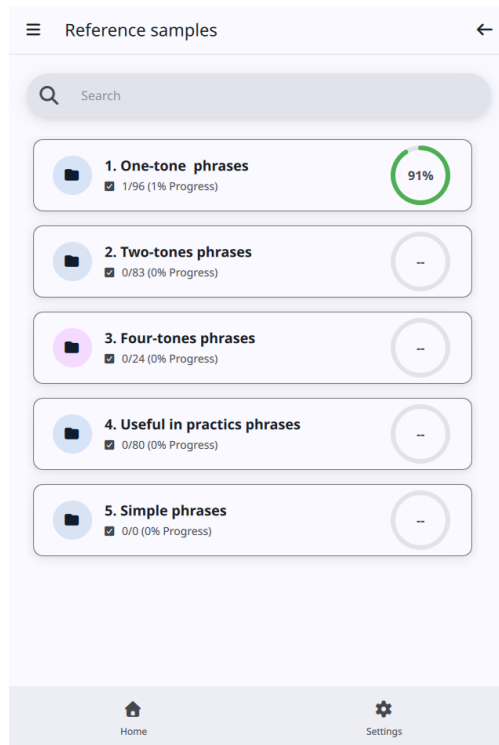


Figure 2: Template Categories screen

Use it to:

- search for a category,
- open a folder of template recordings,
- continue to the template file list.

4.3 Template Files screen

This screen lists the actual reference .wav files inside the selected category.

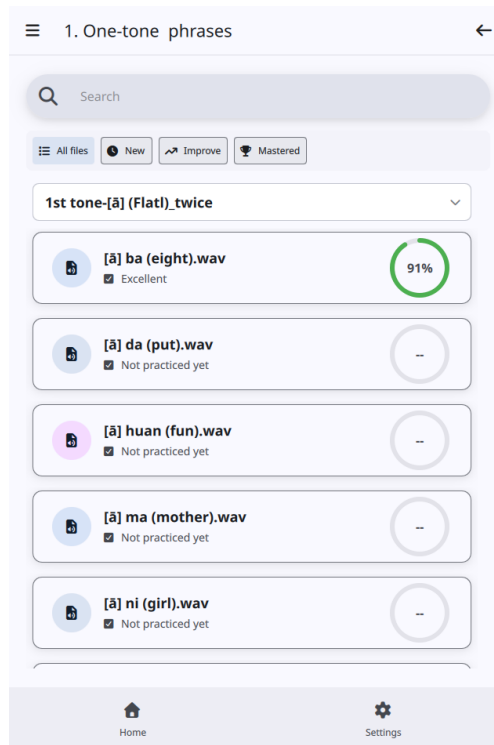


Figure 3: Template Files screen

You can:

- search for a specific file,
- filter by status,
- expand or collapse audio groups,
- open a template for training.

4.4 Training screen

This is the main practice screen.

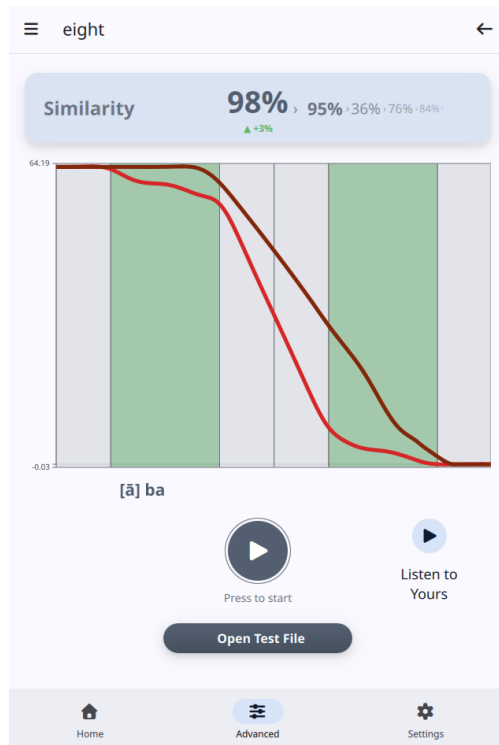


Figure 4: Training screen

On this page you can:

- play the reference,
- record your attempt,
- analyze the result,
- compare your audio with the target sample,
- replay your latest recording.

This is the most important screen in the application.

4.5 Advanced Analysis screen

This screen is for deeper inspection.

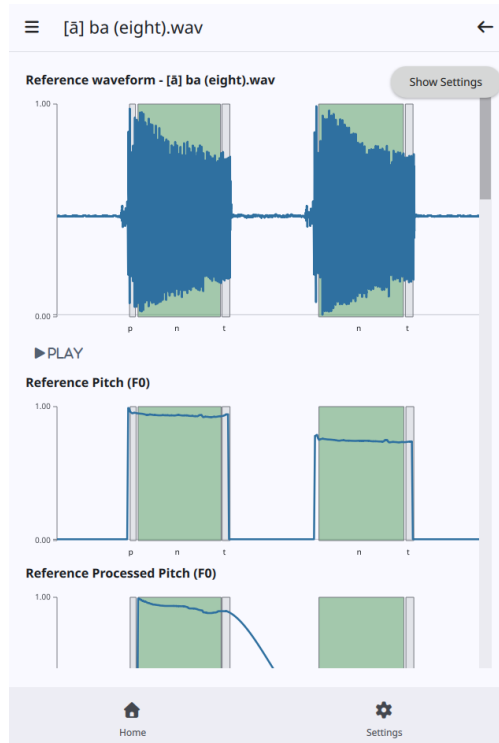


Figure 5: Advanced Analysis screen

This screen is designed for more in-depth analysis. These settings directly affect the performance of the signal analyzer and the quality of the spoken phrase's matching with the reference. These settings are very effective, but if unjustified, they can lead to unpredictable application behavior.

This is useful when you want to see:

- waveform differences,
- pitch curves,
- DTW matching behavior,
- DTW traces,
- spectrum/cepstrum diagnostics.

It is recommended to change these advanced settings **only after consulting** with the developer. Use them only after you have already obtained a result and want to understand why the score was too high or low.

4.6 Records screen

This page stores your saved attempts.

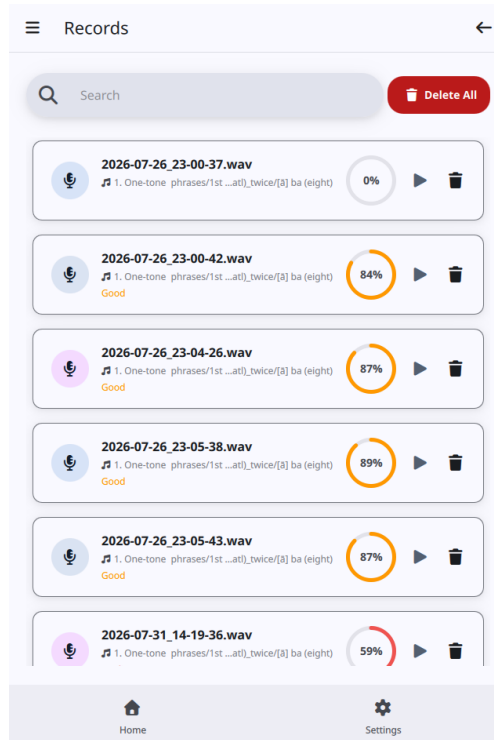


Figure 6: Records screen

It allows you to:

- review historical recordings,
- inspect previous results,
- reopen a record in the analysis view,
- delete individual items if needed.

4.7 Settings screen

The Settings page controls how the app behaves.

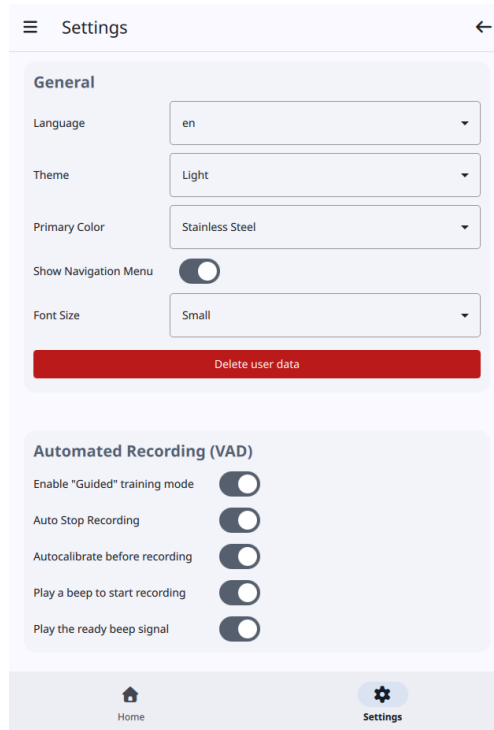


Figure 7: Settings screen

It includes:

- appearance settings,
- deleting user data if necessary,
- navigation settings,
- recording and VAD (Voice Activity Detector) settings,
- detailed DSP (Digital Speech Processor) tuning settings.

Most users only need a small subset of these controls.

5 Training modes available in the app

IntonTrainer 2-Ch supports three practical training styles. The difference is mainly in how the app starts and stops recording during the practice cycle (SeeL Screenshot: Settings screen).

5.1 Auto mode

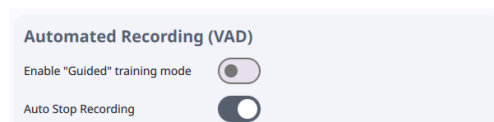


Figure 8: Auto mode (VAD)

Auto mode is the simplest way to train.

How it works:

- the app starts listening as soon as the training screen is active,
- the speech detector detects when you start speaking and records your speech,
- the recording stops automatically after silence is detected,
- the result is processed immediately.

Best for:

- fast repetition,
- continuous practice,
- users who are comfortable with a free-flow speaking style,
- users who want the app to handle the recording lifecycle automatically.

5.2 Guided mode

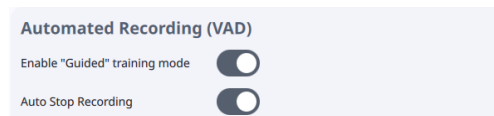


Figure 9: Guided mode (VAD)

Guided mode is the structured training mode.

How it works:

1. The app plays the reference audio.
2. The user listens to it.
3. After a short pause, the app opens a listen window.
4. The user speaks when ready.
5. The system records the response and processes it.

Best for:

- beginners,
- users who want more control over the timing of the response,
- practice sessions where the reference should be heard first every time,
- situations where the user wants a clear listen → prepare → speak sequence.

You can tune the guided-mode timing (listen timeout and post-playback delay) in the Settings page under the Recording / VAD section - look for the Guided mode options.

5.3 Manual mode

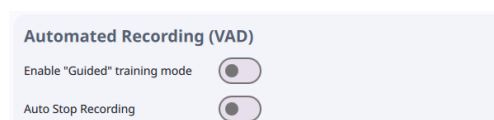


Figure 10: Manual mode (VAD)

Manual mode is used when VAD / automatic speech detection is effectively disabled.

How it works:

- the app does not try to stop the recording automatically,
- the user controls the recording manually,
- the recording is saved after the user finishes it,
- the result is then analyzed normally.

This mode is useful when:

- you want full control over when the recording starts and ends,
- the environment is noisy and automatic voice detection is unreliable,
- you prefer a deliberate, hands-on practice style.

5.4 Difference between the modes

Mode	Behavior	Best use case
Auto	Starts recording on speech activity and lets the system drive the full loop	Fast, continuous practice
Guided	Plays the reference first, then waits for your response in a controlled sequence	Structured practice and beginner-friendly repetition
Manual	Recording is controlled by the user when VAD is disabled	Deliberate, hands-on practice and noisy environments

5.5 Which mode to use in different cases

Use Auto mode when:

- you want a fast and uninterrupted practice loop,
- you are comfortable speaking immediately,
- you want minimal interaction with the recording flow.

Use Guided mode when:

- you are learning the pattern step by step,
- you want to listen to the reference before responding,
- you want a more consistent practice rhythm,
- you want better control over the response timing.

Use Manual mode when:

- you want to control the recording manually,
- VAD is not working well in the current environment,
- you want to practice with a stable, explicit start and stop.

5.6 Practical recommendation

- Start with Auto mode if you already know the workflow and want speed.
- Start with Guided mode if you are learning a new pattern or want a more deliberate practice cycle.
- Use Manual mode when you want full control and auto-detection is not reliable.

6 Metrics and progress indicators on each screen

The app gives you feedback in several different ways. Some numbers describe your overall progress, while others show how a single attempt compares to the reference pattern.

6.1 Home screen metrics

The Home screen is your overall progress summary.

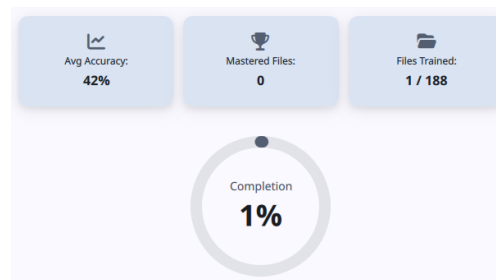


Figure 11: Home screen metrics

Metric	Meaning	How it changes during progress
Avg Accuracy	The average of your best results across the files you have already practiced	Increases as your repeated attempts become more accurate
Mastered Files	How many files have reached a strong result level, currently around 90% or higher	Grows as you consistently improve a file to the mastery range
Files Trained	How many files have been practiced at least once	Increases whenever you complete a valid training attempt
Completion	The overall progress level for the full training set	Rises as more files are practiced and improved

In practice:

- a higher Avg Accuracy means you are consistently matching the reference better,
- more Mastered Files means more patterns are now under control,

- a growing Files Trained count shows that your practice routine is moving forward.

6.2 Template category and template file list metrics

On the Template Files and category views, the app shows how ready each item is.

For a single file card:

- the card shows a score when the file has already been practiced,
- the status label may read Excellent, Good, Medium, Bad, Failure, or Not practiced yet,
- the circular progress indicator reflects your best score for that file.

For a folder or category card:

- the app shows how many files inside it have already been practiced,
- the folder completeness percentage is calculated from the results inside that category.

As you improve:

- a file that has not been practiced yet starts without a score,
- after a successful attempt, the score and label update,
- better results gradually move the card into the higher-quality ranges and increase folder completion.

6.3 Training screen metrics

The Training screen is where you see the immediate result of the current attempt.

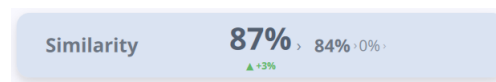


Figure 12: Training screen metrics

Metric	Meaning	How it changes during progress
Current Similarity Score	Your current result as a percentage	Changes every time a new recording is compared to the template
Trend Indicator	Whether the new attempt is better or worse than your previous one	An upward arrow means improvement, a downward arrow means a weaker result, and a flat dot means no change
History Trail	A short sequence of your recent scores	Grows as you practice more, with the newest result nearest the center

How to read this feedback:

- a score close to 100% means your attempt is very close to the reference,
- the trend tells you whether you are improving from one try to the next,
- the history trail helps you notice if your performance is becoming steadier over time.

A good rule is to treat the Training screen as the “live feedback” page, and the Home screen as the “long-term progress” page.

6.4 Records screen metrics

The Records screen stores your completed attempts for later review.



Figure 13: Records screen metrics

Metric	Meaning	How it changes during progress
Record Score	The score attached to that saved attempt	Stays with that recording in the archive
Reference Pattern	The template the record was compared against	Does not change after saving
Date / Time	When the recording was made	New attempts appear as new entries in the history

This page is mainly a review and archive page. It keeps the results of your practice so you can revisit older attempts and compare them later.

6.5 Visual curves are diagnostic metrics, not only decoration

The graphs on the training view are also part of the feedback system:



Figure 14: Pitch and waveform comparison curves

- the user curve and reference curve show how your intonation shape lines up with the model,
- large mismatches in the graph often correspond to lower similarity scores,
- as your accuracy improves, the curves usually become more visually consistent with the reference.

So, the app combines numerical progress metrics with visual comparison feedback.

6.6 How training results are fixed

The app preserves your training progress in two ways:

- the best score for each template file is remembered and shown on the file card,
- completed attempts are stored as records so you can review the exact result later.

How the result is fixed:

- each time you finish a training attempt, the app calculates a similarity score,
- if this score is the best one so far for that file, it updates the file's progress marker,
- the attempt itself is optionally saved in the Records screen for later review,
- the Home screen and category cards then use those fixed best results to update averages, completion, and quality labels.

What this means in practice:

- a better score on a repeat attempt replaces the previous best result for that file,
- old attempts still remain available in the Records archive,
- the app uses the fixed best score, not the latest score, to show your overall progress.

How this affects the rest of the app:

- the Home screen updates its averages and completion numbers using the fixed best scores from each file,
- folder and category progress values are recomputed from the updated file results,

- the per-file card label may change from Good or Medium to Excellent when your best score improves,
- the total number of Mastered Files can grow if your new result crosses the mastery threshold.

This method helps you track both your current performance and your long-term improvement.

7 Recommended training routine

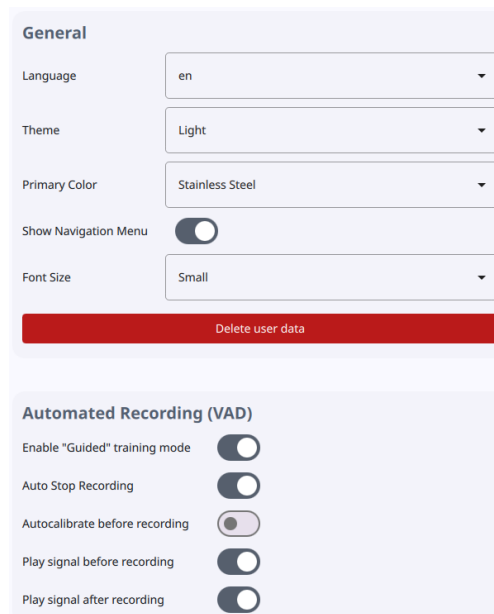
For the best experience, use the app in a quiet environment, with a high-quality microphone, and keep the practice flow simple:

1. Start with a familiar template.
2. Listen to the reference recording carefully.
3. Speak naturally and clearly into a high-quality microphone (external or a good built-in laptop mic). Microphone quality has a direct effect on pitch analysis.
4. Let auto-stop finish the recording when it is enabled.
5. Review the similarity score after each attempt.
6. Repeat the same phrase several times before moving on.
7. Use the advanced analysis view only when you need to inspect the mismatch in more detail.

A good habit is to stay with a small set of templates and practice them repeatedly, rather than switching too often between many different patterns.

8 Settings you can safely change

These are the settings that are typically safe and useful for most users.



The screenshot displays the settings interface of the application, organized into two main sections: General and Automated Recording (VAD).

General Settings:

- Language:** A dropdown menu currently set to "en".
- Theme:** A dropdown menu currently set to "Light".
- Primary Color:** A dropdown menu currently set to "Stainless Steel".
- Show Navigation Menu:** A toggle switch that is currently turned on.
- Font Size:** A dropdown menu currently set to "Small".

Below these settings is a prominent red button labeled "Delete user data".

Automated Recording (VAD) Settings:

- Enable "Guided" training mode:** A toggle switch that is currently turned on.
- Auto Stop Recording:** A toggle switch that is currently turned on.
- Autocalibrate before recording:** A toggle switch that is currently turned off.
- Play signal before recording:** A toggle switch that is currently turned on.
- Play signal after recording:** A toggle switch that is currently turned on.

Figure 15: Basic settings section

8.1 Language

Change this setting if you want the inscriptions in the application to be displayed in English or Russian.

8.2 Theme

You can switch between light, dark, or system theme.

8.3 Primary Color

A cosmetic setting that changes the visual accent color of the interface.

8.4 Show Navigation Menu

Use this if you prefer either a standard navigation layout or a more detailed desktop-style presentation layout.

8.5 Font Size

Increase or decrease the UI text size if the default feels too small or too large.

8.6 Guided mode

If you want a more controlled training cycle, enable guided mode.

This is typically a good option for beginners because it separates:

- listening to the reference,
- preparing to answer,
- speaking the response.

8.7 Auto Stop Recording

This is often useful because it reduces the need to manually stop recording.

Recommended for most users unless you intentionally want to control the recording length yourself.

8.8 Autocalibrate before recording

This is a practical setting for microphone noise management.

If you change room, headset, or microphone, run calibration again.

8.9 Play a beep to start recording

This is a notification signal to begin pronouncing the training phrase immediately after the sample being studied has been voiced.

8.10 Play the ready beep signal

This signal notifies that all analysis processes have been completed and the corresponding tone curve of the spoken phrase has been displayed on the screen.

9 Practical tips for new users

9.1 Start simple

Do not change many settings at once. If a result seems odd, first test one variable at a time.

9.2 Use a high-quality microphone

Pitch analysis depends on a clean recording. Prefer a dedicated external microphone or a good built-in laptop microphone over a cheap or noisy one. Poor microphone quality often looks like a pronunciation problem in the score.

9.3 Prefer calibration after environment changes

If you move the microphone, change rooms, or switch headphones, run calibration again.

9.4 Use guided mode when learning the pattern

Guided mode is a helpful structure for beginners because it creates a clear cycle: listen, prepare, speak, evaluate.

9.5 Use advanced analysis only after the score is already visible

Do not start deep tuning unless you already know which part of the result you want to inspect.

9.6 Keep the same room and mic setup when comparing progress

Consistency makes score changes more meaningful.

10 Troubleshooting checklist

If the app does not behave as expected, check these items first:

1. Confirm the microphone is connected and active.
2. Re-run calibration if room noise changed.
3. Make sure Auto Stop Recording matches your preferred work style.
4. If the app detects speech too often or too rarely, adjust the VAD method and thresholds carefully.
5. If the score looks unstable, keep the default advanced DSP settings unless you are intentionally investigating alignment.

6. If guided mode fails, verify that auto-stop is enabled and that the listen timeout is reasonable.
7. If a recording is silently skipped: check `build/application.log` (or the runtime log in the application directory) for a DTW-rejection message, and consider raising the DTW Distance Limit in Settings → DP Calculation. Also confirm where recordings are stored (the app's data/records directory) if you want to inspect saved audio files directly.

11 Best practice summary

For most users, the best approach is:

- use the default settings,
- keep audio conditions consistent and use a high-quality microphone (external or a good built-in laptop mic),
- use guided mode for structured practice,
- only enter advanced analysis when there is a specific reason,
- avoid modifying DTW and DSP tuning unless you are intentionally diagnosing a technical issue.

This keeps the app easy to use while still giving experienced users full access to deeper analysis controls.

12 Final recommendation

If you are new to IntonTrainer 2-Ch, use the app in its default configuration first. Focus on learning the workflow, listening to reference files, recording your own speech, and reviewing score improvements over time.

Only after that should you begin exploring the advanced controls in the settings page.

13 License

The application's own source files are offered under the [MIT License](#). The built executable also incorporates GPL-licensed ALGLIB, and Qt is dynamically linked under LGPLv3. Complete notices, source availability, and Qt relinking information are available from **Open-source licences** in the side menu.

14 Third-Party Libraries

Library	Purpose
SPTK (Speech Signal Processing Toolkit 4.3)	F0 extraction (RAPT algorithm), audio feature processing
ALGLIB 4.06.0	Spline smoothing for pitch and UMP profiles
Font Awesome (Free 7.2.0)	Icon font used throughout the UI
Ten-vad	Vendored for evaluation; not linked into the current app

15 Authors

- **Boris Lobanov** - Scientific - [LinkedIn](#)
- **Vladimir Zhitko** - Development - [LinkedIn](#)