JAPANESE PITCH ACCENT TRAINER

PROJECT RESEARCH

HAFZA ABDULLAHI
4TH YEAR SOFTWARE DEV STUDENT
SUBMISSION DATE: 23/10/2025

CONTENTS

LINGUISTIC FOUNDATIONS OF THE JAPANESE PITCH ACCENTS

Japanese pitch accent is often described as significant acquisition barrier for second language learners, especially those from non-pitch accent languages. English has a stress accent which uses the loudness and duration where as Japanese uses pitch changes, according to Kanshudo [1] Japanese is “rhythmic” and jumpy with abrupt pitch changes between words, there are 4 main types of pitch changes in Standard Tokyo Japanese: *Heiban, Atamadaka, Nakadaka and odaka*. Each one is uniquely different and some words that are similar, can have different meanings depending on how it was pronounced. These patterns are phonemic, for example, pairs like "hashi" (箸 chopsticks: HIGH-LOW) versus "hashi" (橋 bridge: LOW-HIGH). Usually using the context of a sentence, you can assume what someone meant even if they used the wrong pitch accent but that’s not always the case.

The Four Pitch Patterns

- Heiban (平板): Low-High pattern (e.g., "iku" 行く *meaning to go*)
- Atamadaka (頭高): High-Low pattern (e.g., "hashi" 箸 meaning chopsticks)
- Nakadaka (中高): Low-High-Low pattern (e.g., "otoko" 男 meaning boy/male)
- Odaka (尾高): Low-High pattern with final drop (e.g., "sakura" 桜 meaning cherry blossom)

VISUAL FEEDBACK AND LEARNING

The study By Offerman and Olsen in 2016 provides crucial insight into the effectiveness of visual feedback in second language acquisition, their research shows that using speech analysis tools, learners can grasp what they are saying and get better overtime at more accurate pronunciation.

SPACED REPETITION LEARNING

Spaced repetition is a learning technique that demonstrates information is better retained when study sessions are distributed overtime, rather than concentrated into a few single sessions sporadically. The flashcard Anki is a great example that implements this well. As a Japanese learner and Anki user

I'm passionate about this learning technique and its effectiveness. This concept originates from Herman Ebbinghaus's 1885 research on memory, focusing on the "forgetting curve" which shows how information retention declines exponentially overtime

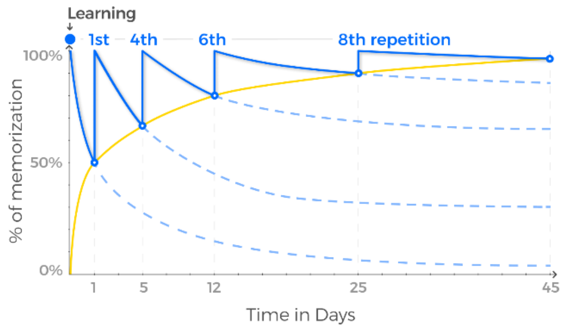


Figure 1 The forgetting curve by Hermann Ebbinghaus

Spaced repletion counteracts this by strategically timing review sessions of study material like flashcards just before they would naturally fade, this strengthens neural pathways and enhances long-term retention. Language learning benefits greatly from this, as learning and remembering vocab and kanji long term for Japanese takes a few review sessions and a lot of immersion.

Figure 2 "100 of the most common kanji characters"

FREQUENCY EXTRACTION USING PARSELMOUTH PYTHON

Parlselmouth is a python library for pitch and frequency extraction made for the praat software [4] based on the autocorrelation methods [5] . It provides tools for detailed speech analysis from recording or real time. This is done in multiple ways

ACOUSTIC PARAMETER OPTIMIZATION

1. Pitch detection range: 75-600 Hz, accommodating gender and age variations in vocal frequency

2. Time resolution: 10ms intervals, sufficient for capturing mora-level transitions

3. Voicing threshold: Optimized for Japanese vowel characteristics and duration patterns

Example graphs using parselmouth python library I made using the word "iku" said by me, compared to a native female speaker pitch accent dictionary website [6]

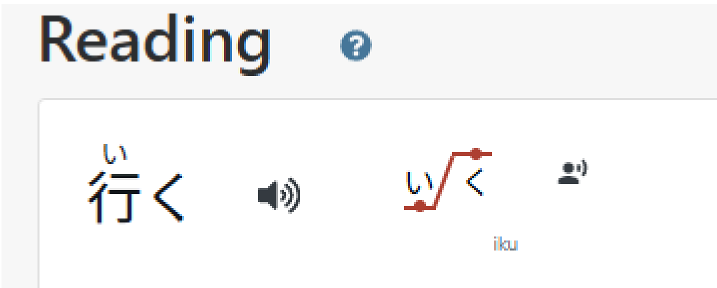
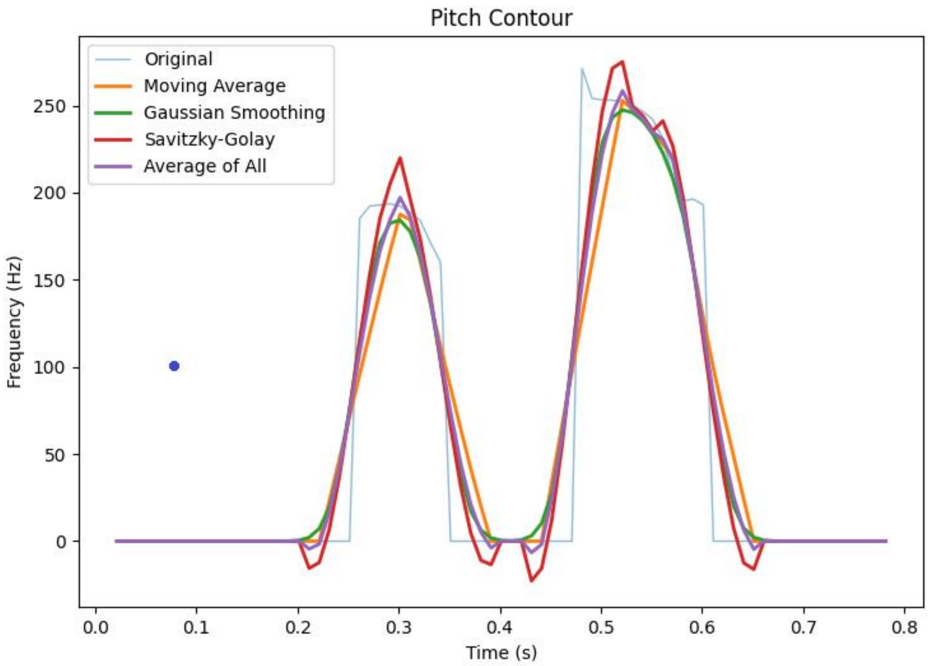
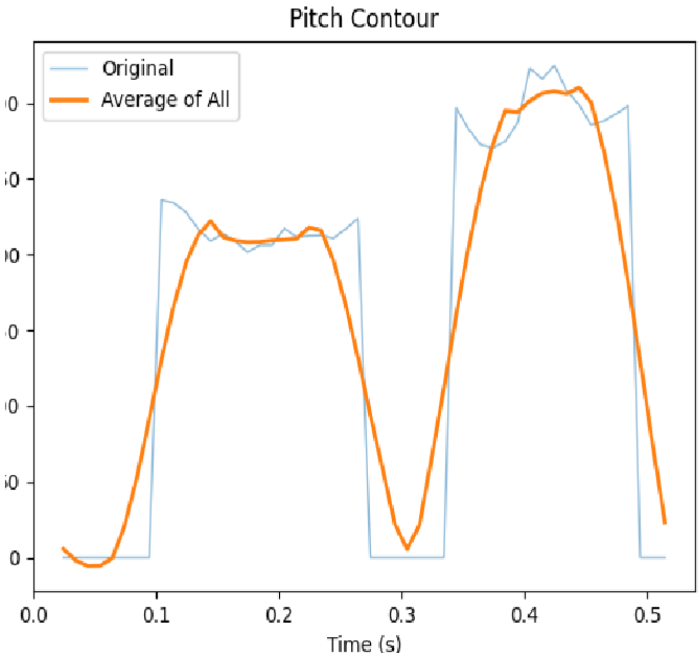


Figure [3] showing the pitch change in "iku" from low to high



graph showing the pitch change in the word iku, speaker: native female [6] img source: my code

These graph shows a clear change in pitch from low to high, the change from the first "i" (low pitch) + "ku" (High pitch)



graph showing pitch change in *in the word iku, speaker is me img source: my code*

GRAPH SMOOTHING AND SIGNAL PROCESSING

Raw pitch extraction contains noise and vocal jitter, recording artefacts. To eliminate this, smoothing of the graph ensures the audio is cleaner more average, so that we can see clear changes in pitch overall

Moving Average Filtering: Provides temporal averaging to reduce high-frequency fluctuations while preserving overall contour shape **(Smith, 1997) [8]**

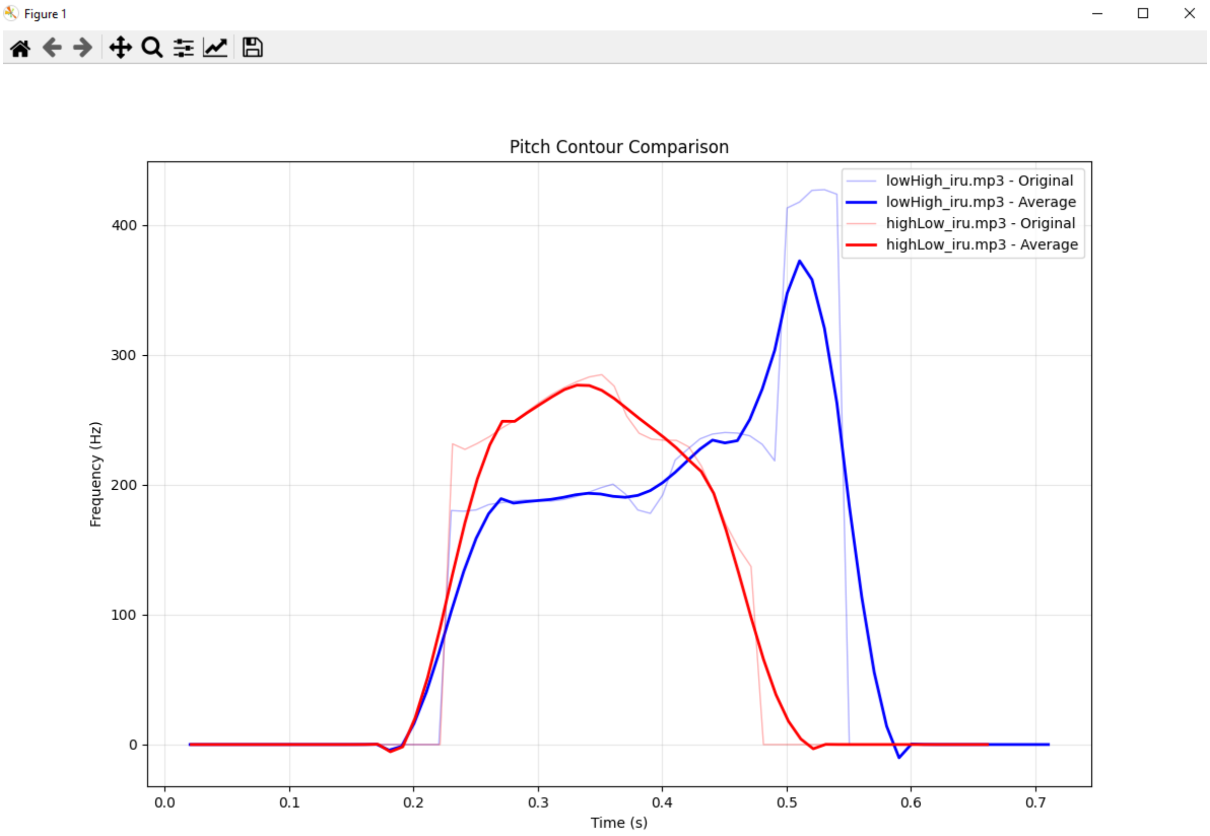
Gaussian Smoothing: Applies kernel-based smoothing that maintains natural pitch movement curvature, essential for preserving accent pattern characteristics.

Savitzky-Golay Filtering: Utilizes local polynomial regression to smooth signals while preserving critical features such as peak locations and transition points **[9]**

The consensus-based approach, averaging results from all three methods, shows clearer and smoother changes across varying recording conditions and speaker characteristics.

Example using the word: “iru” , in this graph both pronunciation (high- low atamadaka) and (low-high heiban) pitch patterns can be seen using a native female speakers recording. The average graphs shows a clearer change in pitch

iru (いる) pronounced using a high low pitch means to need where as, “iru (いる)” pronounced with low high pitch means to be.



graph showing different pitch changes in *in the word iru, speaker is me img source: my code*

ERROR DETECTION IN PITCH FOR USERS

Error classes:

- early_drop - Pitch drops too soon
- late_drop - Pitch drops too late
- no_drop - Missing pitch drop entirely
- wrong_pattern - Used completely wrong accent pattern

correct - Proper pronunciation

pitch accent by x amount of hz to be considered correct alongside Dynamic Time warping

DYNAMIC TIME WARPING FOR PATTERN COMPARISON

Dynamic Time warping (DTW) is a technique for comparing two time series that may vary in speed or timing, by warping the axis to find the best possible alignment. It solves the problem of comparing sequences that are similar in shape but not perfectly aligned. **[9]** It allows one of the sequences to stretch or compress to match the other, hence the name. This is useful for pattern recognition in words to analyse where the syllable and pitch changes happened, and if it matches up.

Dynamic time warping distances are then converted to percentage similarity scores using derived thresholds, this gives a "correctness scores" of a pronunciation of the given word compared to native speaker/s

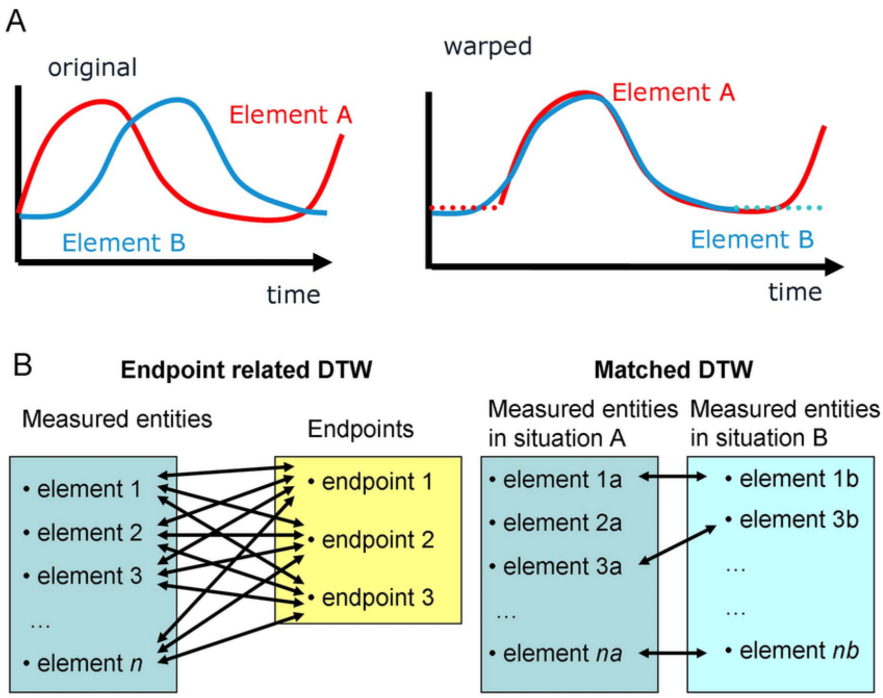


Figure [4] showing Dynamic Time warping

DEVELOPMENT

FLUTTER CROSS-PLATFORM DEVELOPMENT

Flutter provides a good framework for developing cross-platform mobile applications using Dart programming language. Since I used Flutter/dart and android studio in my last project I am comfortable on the workspace. Flutter is a widget-based architecture that enables consistent UI layouts and rendering.

Key advantages:

- Hot reload feature for rapid development iteration and checks
- Native plugin ecosystem for audio recording and playback
- Single codebase deployment to multiple platforms
- Massive widget library for UI and graphs etc

Feature Implementations:

- **Audio Recording**: Utilizes audiorecorder or flutter_sound plugins for capture
- **Data Visualization**: Employs charts_flutter for pitch contour graphing

BACKEND API ARCHITECTURE WITH PYTHON/FLASK

Flask is a popular lightweight framework for building RESTful APIs in Python. Flasks supports rapid prototyping and scalability

API Endpoint Architecture:
The API follows REST principles with JSON payloads, providing stateless communication layer between the flutter mobile client and python backend

FIREBASE AUTHENTICATION & SECURITY

Firebase provides a NoSQL cloud database with real-time synchronization capabilities." *Firebase Realtime Database and Cloud Firestore provide database services. I listed them both as "realtime, cloud hosted, NoSQL databases".* **[10]**

Firebase has easier and safer authentication setup with the options of google sign **[11]** in as well

RELEVENT OVERALL RESEARCH ABOUT WAVES AND FREQUENCY

In wave mechanics, wave speed doesn’t change when frequency changes. It is often misunderstood that changing the frequency of a wave also changes its speed. What changes frequency are oscillations
Wave Speed = Frequency × Wavelength
Wave speed is determined by the medium, the thing the wave is travelling

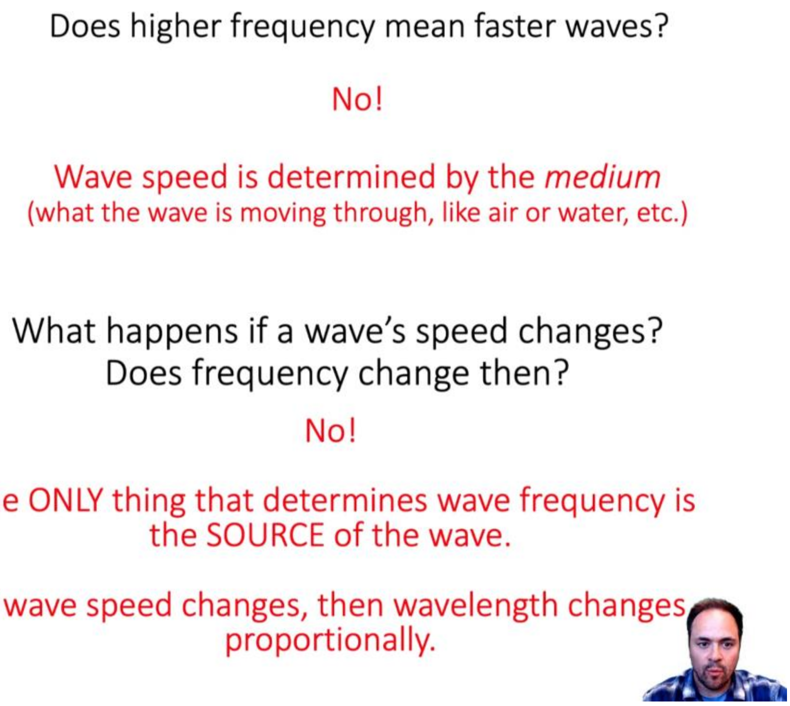


Figure 4 Waves - Frequency, Speed, and Wavelength YouTube video by Jon white

How are frequency and wavelength related?

- High Frequency results in a Short Wavelength.
- Low Frequency results in a Long Wavelength. **[12]**

FLUTTER DOCUMENTATION

ANKI CONNECT API

The **AnkiConnect** API plays an important role in enabling the Japanese Pitch Accent Trainer to integrate spaced-repetition vocabulary data directly from existing Anki decks. Because the final Flutter application must run on both Android and Web platforms or environments where direct parsing of. apkg (SQLite-based) Anki files is impossible. Instead Anki's API retrieves the card data and media and saves them locally

The system uses the following sequence:

1. **Discover Notes**
   The script issues a findNotes action with a query of the form:
   "deck:Kaishi 1.5k".
   This returns the full collection of note IDs belonging to the deck. Limiting the set to the first 50 entries ensures faster preprocessing and avoids memory overhead.

2. **Retrieve Full Note Metadata**
   Using notesInfo, each note ID is expanded into its complete data structure, including fields such as:

   - **Word**
   - **Word Reading**
   - **Word Meaning**
   - **Sentence**
   - **Word Audio** (stored as [sound:filename.mp3])

3. **Audio File Retrieval**
   The retrieveMediaFile endpoint returns base64-encoded audio data for each note's pronunciation file.
   These are decoded and written to:

   assets/audio/<filename>

This makes all native speaker recordings available offline inside the Flutter app.

4. **Normalization and Serialization**
   Card fields are cleaned of residual HTML and exported to a structured JSON file:

REFERENCES

LINKS

- **[1] Kanshudo Pitch accent site** https://www.kanshudo.com/howto/pitch **accessed 18/10/2025**

- **[2] Science Direct**, July 2016 https://www.sciencedirect.com/science/article/abs/pii/S0346251X16000373 **accessed 18/10/2025**

- **[3] The forgetting curve: Hermann Ebbinghaus, 1885** https://www.taalhammer.com/the-forgetting-curve-hermann-ebbinghaus/ **accessed 18/10/2025**

- **[4] Parselmouth praat ,** https://parselmouth.readthedocs.io/en/stable/ **accessed 19/10/2025**

- **[5] Auto correlation methods, Boersma, 1993** https://www.fon.hum.uva.nl/paul/papers/Proceedings_1993.pdf **accessed 19/10/2025**

- **[6] Pitch accent dictionary,** https://www.gavo.t.u-tokyo.ac.jp/ojad/search **accessed 20/10/2025**

- **[7] smoothing average, Steven W. Smith** https://www.dspguide.com/ch15/5.htm **accessed 20/10/2025**

- **[8] Savitzky-Golay Filter,** https://medium.com/pythoneers/introduction-to-the-savitzky-golay-filter-a-comprehensive-guide-using-python-b2dd07a8e2ce **accessed 20/10/2025**

- **[9] Dynamic Time warming**, https://link.springer.com/chapter/10.1007/978-3-540-74048-3_4 **accessed 18/10/2025**

- **[10] The medium, what is firebase, Doug Stevenson** https://medium.com/firebase-developers/what-is-firebase-the-complete-story-abridged-bcc730c5f2c0 **accessed 18/10/2025**

- **[11] Firebase authentication,** https://firebase.google.com/docs/auth/where-to-start **accessed 23/10/2025**

- **[12] Waves - Frequency, Speed, and Wavelength, Jon White accessed** https://youtu.be/4yfXp1jNBn8?si=0_yGhUKXm-Bb019K **23/10/2025**

FIGURES

- **[1] The forgetting curve: Hermann Ebbinghaus, 1885** https://www.taalhammer.com/the-forgetting-curve-hermann-ebbinghaus/ **accessed 18/10/2025**

- **[2] ThoughtCo. 100 common kanji** https://www.thoughtco.com/the-most-frequently-used-kanji-2028155 **accessed 20/10/2025**

- **[3] the word iku , https://www.japandict.com/%E8%A1%8C%E3%81%8F accessed 19/10/2025**

- **[4] Dynamic Time Warping, Jos Kleinjans** https://www.researchgate.net/figure/Dynamic-time-warping-Panel-A-shows-how-time-series-can-be-aligned-using-dynamic-time_fig4_256102831 accessed **23/10/2025**

- **[5] Waves - Frequency, Speed, and Wavelength, Jon White accessed** https://youtu.be/4yfXp1jNBn8?si=0_yGhUKXm-Bb019K **18/10/2025**