CHI'E Vol. 14 (1) (2026)

Chi'e: Jurnal Pendidikan Bahasa Jepang

Terindeks Sinta 3

https://journal.unnes.ac.id/journals/chie/index

Autonomous Shadowing Practice and Prosodic Development in Intermediate Learners of Japanese: A Mixed-Methods Study

Jiaan Xu[1*]

Japanese Language and Literature Program, Faculty of Foreign Languages, Ningbo University, China

Email: jiaanxu3@gmail.com

*) Corresponding Author

Article Info

Article History:

Received: December 2025

Approved February 2026

Published: March 2026

Keywords:

shadowing; Japanese prosody; pronunciation development; acoustic–perceptual analysis

Abstract

Autonomous oral practice is widely encouraged in Japanese language learning, but its pronunciation benefits are not always explained in ways that link measurable speech change with native-speaker perception. This mixed-methods study examined how repeated shadowing practice supported prosodic development in three intermediate learners of Japanese. Over seven weeks, participants completed two shadowing sessions per week using textbook audio recordings as model input. The analysis combined two approaches: acoustic inspection of fundamental frequency (FO) contours and timing patterns using PRAAT, and perceptual evaluation by two native Japanese speakers focusing on accent, intonation, word rhythm and overall improvement. Across learners, second attempts on the same material generally showed smoother pitch movement, more stable rhythmic organization, and higher perceptual ratings. Indicating improvement in prosodic naturalness even within a short practice cycle. At the same time, development was not uniform. Some learners improved mainly through more coherent intonation, while others showed faster delivery that sometimes coincided with reduced segmental precision, such as shortened long vowels. These findings suggest that autonomous shadowing can effectively support prosodic development, while also highlighting the need for pacing and feedback to maintain accuracy as fluency increases.

✉ Alamat korespondensi:
Gedung B4 Lantai 2 FBS Unnes
Kampus Sekaran, Gunungpati, Semarang, 50229
E-mail: chie@unnes.ac.id

E-ISSN 2685-6662
P-ISSN 2252-6250

INTRODUCTION

Spoken Japanese competence is not exhausted by grammatical accuracy. Even when intermediate learners can produce structurally well-formed sentences, prosodic control—timing, rhythm, pitch movement, and accent placement—often remains unstable, and this instability can affect intelligibility and perceived naturalness. For classroom contexts with limited contact hours, a recurring practical question follows: what kind of oral practice can be sustained outside teacher-led instruction while still producing observable gains in pronunciation, particularly in prosody?

Shadowing, understood here as near-simultaneous oral reproduction of aural input, has been repeatedly treated as a pragmatic technique for developing oral performance. Long-term evidence suggests that its contribution is not limited to short-term fluency effects.

Takahashi, Fukuda, Iwashita, and Sakoda (2010), working with advanced learners over roughly one year, tracked changes in pronunciation errors while also examining learners' diary-based reflections. Their results indicate two points that matter for research design: first, error counts can decrease across tasks, yet task difficulty and cross-learner comparability constrain straightforward interpretation; second, learners' "noticing" tends to shift from vague impressions to more concrete, model-based comparisons, and attention may gradually expand from segmental issues toward suprasegmental features.

At the same time, they highlight a measurement problem: general pre/post tests may show ceiling effects for higher-level learners, implying that prosodic change often requires finer-grained observation.

Methodological work in Japanese pronunciation research has therefore increasingly paired subjective listening judgements with acoustic visualization. Yin (2015) framed shadowing as a compensatory option for contexts where dedicated phonetics courses and prosody training are insufficient, and tested a three-month training program using both teacher auditory ratings and PRAAT-based F0 contour analysis. A key implication is that perceptible improvement does not always map neatly onto a single acoustic parameter; in some cases, "more natural" impressions may be related to changes such as speech rate rather than full convergence on micro-features like segment duration.

In a similar vein, Okumura (2015) examined Vietnamese learners' shadowing over four months, focusing on word accent and sentence-final intonation patterns. While improvement was observed for certain targets (e.g., accent accuracy and question-final rises), other items—such as intonation associated with sentence-final particles—proved more resistant.

The study underscores that prosodic gains can be item-specific and may interact with L1 prosodic tendencies, which makes it risky to generalize "shadowing works" without specifying what improved, how it was assessed, and for whom.

Another line of evidence emphasizes feasibility and learner experience rather than phonetic outcomes alone. In a large-class university setting, Wang and Hu (2019) implemented shadowing by replacing part of in-class read-aloud time with shadowing practice and compared it with a read-aloud control condition over eight weeks. Their listening test results showed a clear advantage for the shadowing group, particularly on materials unrelated to the practiced textbook texts, and they argued that dense practice and enforced auditory focus (including periods without looking at the script) may strengthen learners' processing of spoken input. This classroom-embedded design is valuable for programs where extensive individualized speaking instruction is unrealistic.

Yet it also draws attention to a boundary: the conditions that produce gains—practice density, speed, and cognitive load—may simultaneously generate frustration or avoidance for some learners.

Indeed, learner perceptions and motivational dynamics repeatedly surface as practical constraints on sustained shadowing. Widyaningtyas (2017), surveying students in an Indonesian program where shadowing was

embedded in an integrated listening-reading course, reported that learners generally viewed shadowing as helpful not only for comprehension but also for speaking-related dimensions such as intonation, accent, and pronunciation.

At the same time, students did not necessarily feel that shadowing alone enabled rapid comprehension behaviors, often relying on rereading to achieve fuller understanding. Sumiyoshi and Svetanant (2017) likewise showed that learners' attitudes toward shadowing are not uniform: while many perceive benefits and value teacher feedback, the pace of materials can become a decisive factor, with some learners interpreting fast input as productive challenge and others reacting negatively when speed exceeds processing capacity.

Taken together, these studies suggest that shadowing's effectiveness cannot be separated from how learners experience the task, regulate difficulty, and persist when progress is uneven.

Against this background, two limitations in the current picture stand out. First, studies that demonstrate measurable change in accent and intonation often rely on either short interventions or narrowly specified targets; fewer designs capture prosodic development across repeated attempts on different texts while combining perceptual evaluation with acoustic comparison in a way that is feasible for ordinary teaching contexts.

Second, work on learner perspectives highlights motivation and perceived affordances, but these insights are not always linked to observable pronunciation development within the same dataset. A tighter connection between "what changed in speech" and "what learners reported as obstacles in practice" remains underdeveloped, especially for intermediate learners whose prosodic control is still malleable but whose progress may be subtle.

The present study addresses these gaps by examining self-directed shadowing practice conducted by three intermediate-level learners using texts from Comprehensive Japanese Course, Book 3. From a broader practice sequence conducted over seven weeks (two shadowing sessions per week), the analysis focuses on paired first- and second-attempt recordings drawn from selected texts in Lessons 2, 6, and 7 for each learner. Pronunciation development is evaluated through two complementary lenses: (a) acoustic comparison of F0 contours using PRAAT and (b) auditory judgements by native Japanese speakers. In addition, the study documents challenges that learners themselves reported in sustaining regular practice and motivation.

By bringing together machine-based acoustic evidence and listener-based perceptual judgement, the study also seeks to offer a more usable frame for Japanese language teachers, who often need to decide whether learner improvement is visible only in instrumental data or is also audible in classroom performance. In this sense, the comparison between acoustic traces and native-speaker evaluations is not only methodological; it also has practical value for pronunciation teaching, because it helps clarify which kinds of change may matter most in actual instructional settings.

The objective is to clarify how far learner-managed shadowing practice can support prosodic development in a realistic instructional setting, and what constraints should be anticipated when such practice is assigned beyond the classroom.

METHOD

Research Design

This study examined pronunciation development associated with repeated shadowing practice by intermediate learners of Japanese. The design combined two complementary evaluation routes: (a) acoustic inspection of fundamental frequency (F0) contours using PRAAT and (b) perceptual evaluation by native Japanese speakers. The acoustic route was used to trace measurable changes in pitch movement and temporal organization within an utterance, whereas the perceptual route was used to judge whether such changes were heard as improvements in communicatively relevant prosody. The two routes drew on partially overlapping, but not identical, speech segments

so that each method could be applied to the type of data it evaluates most reliably.

Participants

The training program was originally conducted with second-year Japanese-major students at Ningbo University. For the present analysis, three intermediate-level learners were selected for close examination, based on the availability of paired recordings (first and second attempts on the same passage). All participants provided informed consent and agreed to take part in sustained oral practice aimed at improving pronunciation.

Training Materials and Procedure

Training materials were taken from Comprehensive Japanese Course, Book 3, using texts and dialogues from Lessons 2, 6, and 7. The accompanying CD recordings by native speakers served as the model input. In each session, learners wore headphones and shadowed the model recording, with the instruction to follow the CD as closely as possible. A single session lasted approximately four to five minutes. Practice was scheduled twice per week across seven weeks, yielding fourteen sessions per learner. The procedure generated a series of learner recordings that allowed within-learner comparison between earlier and later attempts.

Data Collection and Data Selection

Two datasets were prepared in accordance with the two evaluation routes.

Acoustic dataset (short excerpts): for each of the three focal learners, short segments were extracted from paired recordings (first and second attempts of the same passage). These excerpts were selected because short, well-bounded stretches allow stable visualization and comparison of F0 contours and facilitate repeated listening during annotation.

Perceptual dataset (longer passages): for the auditory evaluation, longer versions of the Lesson 2/6/7 materials were used together with the corresponding learner recordings. Longer passages were chosen here to support holistic judgments of intonation flow and rhythm across an utterance, rather than only local pitch events.

Analytical Framework

Acoustic analysis (PRAAT)

PRAAT was used to compare learners' production across attempts and against the native-speaker model. Analysis focused on (a) overall F0 contour similarity, (b) pitch-accent realization where relevant, and (c) temporal aspects observable alongside pitch movement, including total utterance duration and mora-level timing. Learner recordings were inspected both at the utterance level and at smaller units (e.g., long vs. short segments) through repeated playback, and F0 traces were interpreted in relation to the model recording for the same passage.

Auditory evaluation (native speakers)

Two native Japanese speakers evaluated learner recordings using a structured questionnaire. They rated three dimensions: accent (pitch placement within words), intonation (naturalness of pitch movement across the utterance), and word rhythm, with rhythm anchored in the production of special morae (geminates, nasals, contracted sounds, and long vowels). In addition, evaluators compared the first and second attempts and rated the perceived degree of improvement on an ordinal scale. A brief calibration step with sample recordings was conducted so that the two raters aligned their interpretation of the rating categories; remaining minor disagreements were resolved through discussion to arrive at a shared judgement.

Data Analysis

Acoustic comparisons and perceptual ratings were first summarized within learner, emphasizing differences between first and second attempts. The two evaluation routes were then interpreted together to clarify where visualized F0 changes converged with (or diverged from) native-speaker impressions of improvement. This triangulated approach was adopted to describe not only whether change occurred, but also how that change appeared in acoustic traces and how it was heard by native listeners.

RESULTS AND DISCUSSION

This section reports how learners' pronunciation and prosodic patterns changed between the first and second attempts of the same shadowing material, using two complementary lenses: PRAAT-based F0 contour inspection and

native-speaker auditory evaluation. Across learners, the main scientific finding is consistent: a second attempt tends to stabilize prosodic organization, often sounding more coherent to native listeners even when improvement is uneven across specific sub-features (e.g., long vowels, local pitch events). This pattern matters because it suggests that shadowing may generate rapid, within-session adjustments that are not always captured by a single metric.

(1) Fundamental Frequency (F0) Analysis – Student A

The acoustic comparison for Student A used the sentence "Tsūshin gijutsu no shinpo ni tomonatte" ("Along with the progress of communication technology"), contrasting the native model and the learner's first and second attempts.

Figure 1. Fundamental Frequency Contour of the Model Recording by a Native Speaker

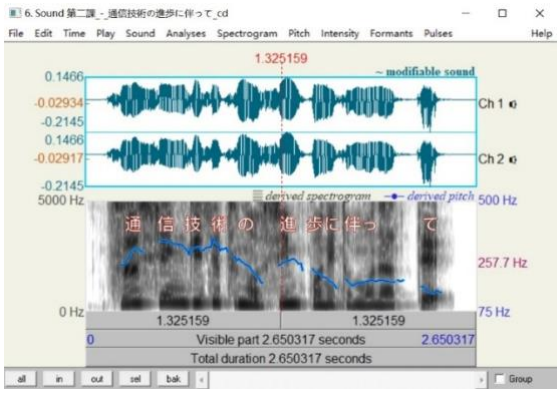


Figure 2. Fundamental Frequency Contour of Student A's First Shadowing Attempt

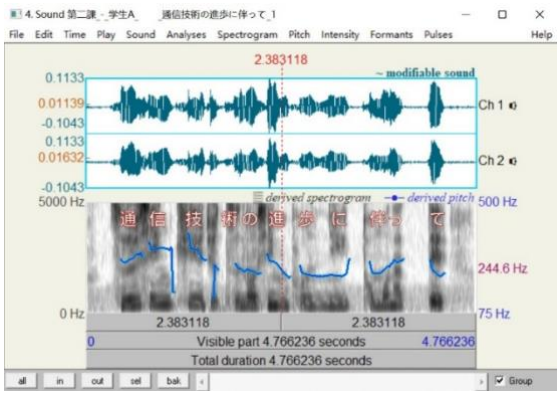
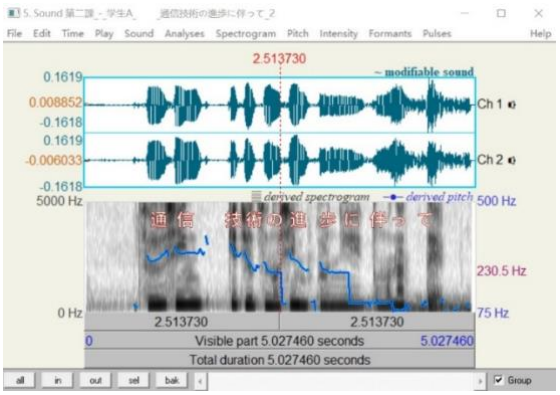


Figure 3. Fundamental Frequency Contour of Student A's Second Shadowing Attempt



In the model production, the utterance duration was 2.65 s and mean F0 was 257.7 Hz, with a smooth contour that declined naturally toward the end. Student A's first attempt was slower (4.77 s, mean F0 244.6 Hz) and showed discontinuous pitch movement, especially around "gijutsu no shinpo", where repeated sharp rises and falls created a segmented impression. Segmental instability was also present (e.g., "gi" realized as "ki"), accompanied by excessive vowel prolongation—features that jointly disrupt rhythmic cohesion.

In the second attempt (5.03 s, mean F0 230.5 Hz), duration did not shorten; however, the contour became visibly smoother and the sentence-final lowering more closely resembled the model. The reduction in mean F0, together with the less jagged contour, is consistent with a shift toward a more stable production mode. Put plainly, Student A's second attempt looked less "effortful" in pitch behavior even though speaking rate remained slow. This is one respect in which the present data differ from studies that equate improvement primarily with acceleration: here, the most salient gain was intonational coherence, not speed.

(2) Fundamental Frequency (F0) Analysis – Student B

For Student B, the target sentence was "Nihon ryōri ni tsukau chōmiryō de, jūyō na no wa shōyu de arō" ("Among the seasonings used in Japanese cuisine, soy sauce is the most

important"). Comparisons were made between the model and two learner attempts.

Figure 4. Fundamental Frequency Contour of the Model Recording by a Native Speaker

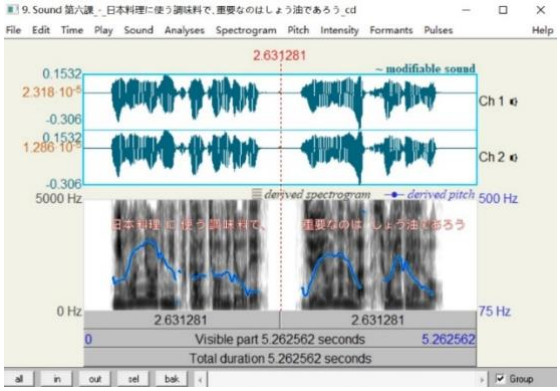


Figure 5. Fundamental Frequency Contour of Student B's First Shadowing Attempt

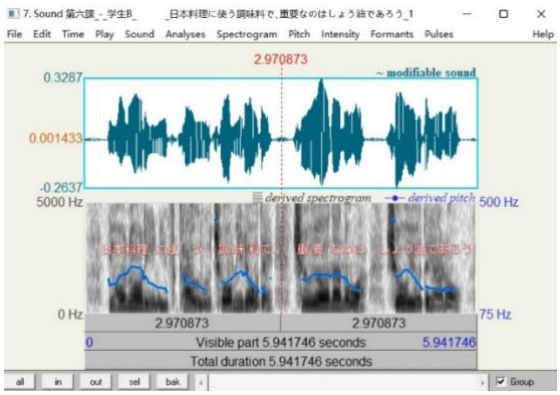
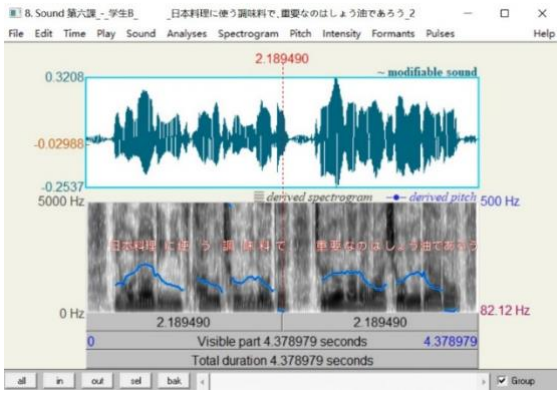


Figure 6. Fundamental Frequency Contour of Student B's Second Shadowing Attempt



The model lasted 5.26 s with mean F0 246.2 Hz, showing smooth phrase-final falling contours. Student B's first attempt was slightly longer (5.94 s, mean F0 232.0 Hz) and sounded less native-like due to disrupted phrasing and phonological errors ("jūyō" → "chūyō"; unnatural production of "shōyu"). The pitch range also appeared comparatively flattened, producing a monotonous impression.

In the second attempt, duration shortened markedly to 4.38 s, indicating a clear tempo increase. The contour displayed more distinct phrase-level rises and falls and "shōyu" approximated the model more closely. At the same time, the long vowel in "jū" tended to shorten, which risks a "rushed" quality. This combination—greater fluency and clearer prosodic phrasing, but occasional mora-length instability—matters when interpreting "improvement." Rather than a uniform upward shift, Student B shows a trade-off pattern: tempo gains can outpace segmental timing control. That nuance is often under-emphasized in classroom-intervention reports where pre/post outcomes are summarized without looking at how the learner reaches them.

(3) Fundamental Frequency (F0) Analysis – Student C

For Student C, the target sentence was "Kono shashin no yō na kanji de" ("In the manner shown in this picture").

Figure 7. Fundamental Frequency Contour of the Model Recording by a Native Speaker

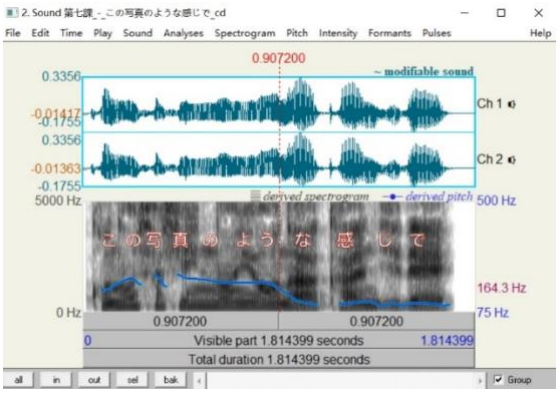


Figure 8. Fundamental Frequency Contour of Student C's First Shadowing Attempt

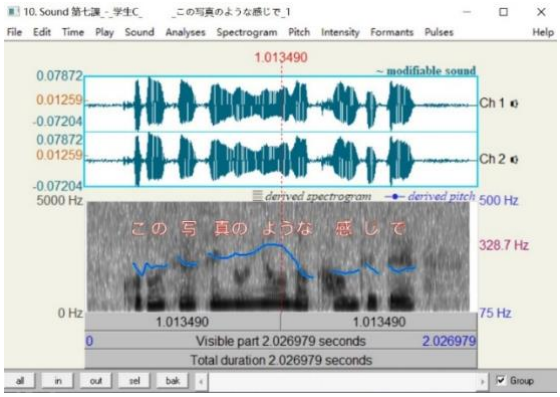
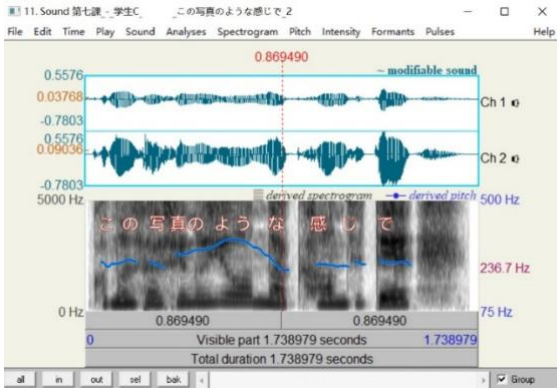


Figure 9. Fundamental Frequency Contour of Student C's Second Shadowing Attempt



The model lasted 1.81 s with mean F0 164.3 Hz, showing a gentle downward movement and stable final lowering. Student C's first attempt was longer (2.03 s) and mean F0 was 328.7 Hz, far higher than the model, with slight pauses between words and upward spikes. Despite clear articulation, this profile suggests heightened tension and insufficient integration of phrase-level prosody.

In the second attempt, duration shortened to 1.74 s, aligning more closely with the model's tempo, and mean F0 dropped to 236.7 Hz. The contour smoothed especially after "yō na", and the sentence-final region looked more native-like. Unlike Student A (slower but smoother), Student C shows both tempo alignment and reduced pitch elevation, a combination that plausibly corresponds to a more confident, less strained production mode. This is notable because it supports the idea that shadowing can trigger not only mechanical imitation but also a rapid recalibration of speaking "setting" (pace + pitch range).

Summary of Acoustic Measures

Table 1. Summary of Utterance Duration and Mean F0 (Hz) for Model and Learner Recordings, With Changes from First to Second Shadowing Attempt

Measure	Model	Student A	Student B	Student C
Utterance duration (s)				
Attempt 1		4.77	5.94	2.03
Attempt 2	2.65	5.03	4.38	1.74
Δ duration (Attempt 2 – Attempt 1)		+0.26	−1.56	−0.29
Mean F0 (Hz)				
Attempt 1		244.6	232.0	328.7
Attempt 2	257.7	230.5	/	236.7
Δ mean F0 (Attempt 2 – Attempt 1)		−14.1	/	−92.0

Table 1 consolidates the key acoustic indicators. Two cross-learner patterns stand out. First, tempo generally moved toward the model in Students B and C, while Student A's tempo did not—yet Student A still showed improvement in contour stability. Second, mean F0 declined from the first to the second attempt in Students A and C, consistent with a reduction in excessive pitch height and possibly reduced articulatory tension.

For Student B, mean F0 for the second attempt is not reported. This decision reflects a methodological constraint rather than a data omission: the pitch track for this sample contained unstable segments that prevented a reliable mean F0 estimate under the same analysis settings applied to the other recordings. To maintain comparability across learners, no alternative parameterization was introduced for

this case. Consequently, acoustic interpretation for Student B's second attempt was based primarily on contour shape and utterance duration rather than on a single summary statistic.

Results of Native-Speaker Auditory Evaluation

Two native Japanese speakers evaluated learner recordings along four dimensions: accent, intonation, word rhythm (special morae such as sokuon, hatsuon, yōon, chōon), and overall improvement (second attempt compared with first).

Figure 10. Comparison of Accent Evaluation Scores Between the First and Second Shadowing Attempts (by Student)

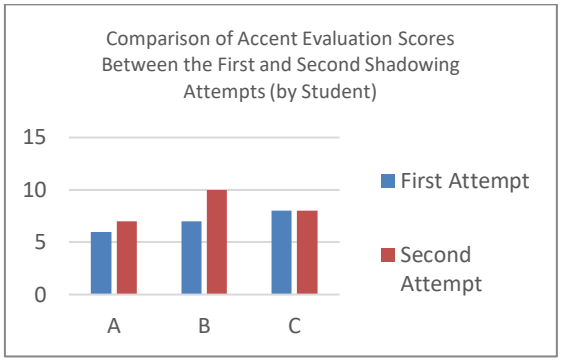


Figure 11. Comparison of Intonation Evaluation Scores Between the First and Second Shadowing Attempts (by Student)

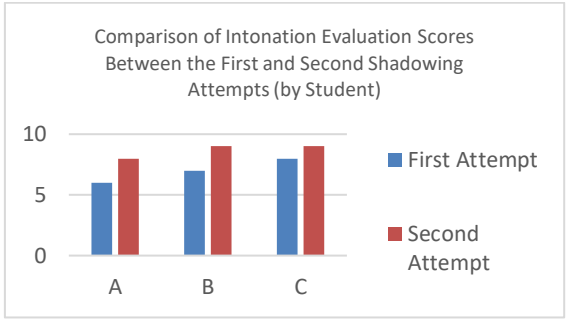


Figure 12. Comparison of Word Rhythm Evaluation Scores Between the First and Second Shadowing Attempts (by Student)

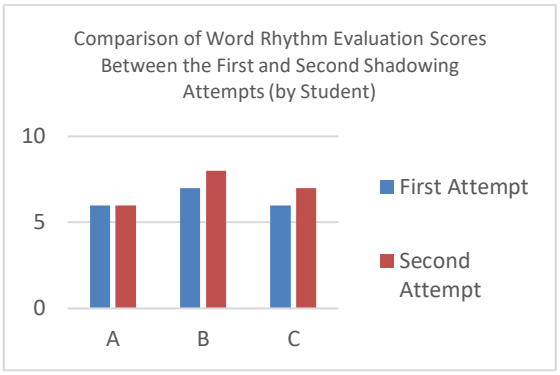
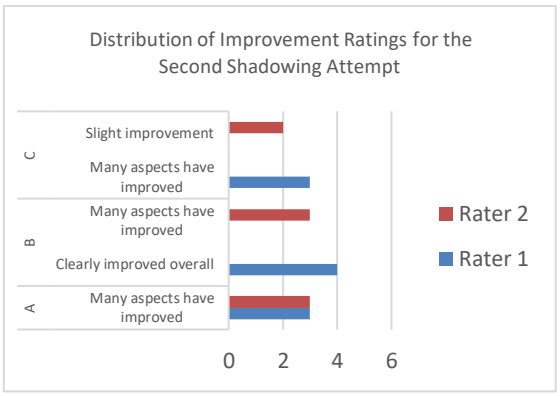


Figure 13. Distribution of Improvement Ratings for the Second Shadowing Attempt (Compared to the First Attempt)



Across Figures 10–12, second-attempt ratings were consistently higher for all three learners, indicating perceived improvement in accent, intonation naturalness, and rhythm. Student B tended to receive the highest evaluations, with ratings frequently approaching the "almost no errors" end of the scale. Students A and C also showed noticeable improvement, though the gains were more category-dependent, consistent with the acoustic observation that learners do not improve evenly across sub-features.

Figure 13 further supports this interpretation. Student B's second attempt was most often judged as "clearly improved overall," whereas Students A and C were more often placed in intermediate categories such as "many areas improved" or "slightly improved." This distribution matters because it indicates that native listeners recognized improvement even when the acoustic record showed mixed signals (e.g., Student A's longer duration). In other

words, the perceptual results suggest that what listeners reward is not speed alone but coherence—a more integrated rhythm and pitch flow.

Discussion

The present study asked whether repeated shadowing practice yields observable pronunciation development and what kind of improvement is captured when acoustic inspection is paired with native-speaker perception. The answer is not simply "it works," but more specific: a second attempt tends to reorganize speech toward greater prosodic stability, often reducing disruptive segmentation and supporting a more native-like phrase-level flow.

This finding aligns with research that reports prosodic gains through shadowing with both auditory judgement and acoustic visualization, yet the present data add a practical nuance: improvement can appear within a very small time window, not only across long training cycles. In Student A, the key change was smoother contour behavior and reduced mean F0 despite no gain in speed; in Student B, gains in tempo and phrase prosody co-occurred with a risk of mora-length shortening; in Student C, both tempo alignment and reduced pitch elevation emerged together. That variability underscores a point sometimes obscured in outcome-focused reports: shadowing does not push all learners along the same "single axis" of improvement.

Compared with classroom-based interventions that primarily report aggregate test gains, this study highlights the micro-structure of change. It also differs from long-term advanced-learner work that foregrounds reflective awareness over time; here, the emphasis is on how production adjusts between adjacent attempts, and how that adjustment is perceived by native listeners. Finally, the perceptual ratings show that listeners may evaluate progress through a holistic lens—valuing smoother intonation movement and rhythm integration—so acoustic change should be interpreted not as an end in itself, but as a route to perceived naturalness.

At the same time, the study's boundaries should be stated clearly. The analysis is based on three learners, and the resulting patterns should therefore be understood as learner-specific rather than broadly generalizable across all learners of Japanese. In addition, one acoustic summary value (Student B's second mean F0) was unavailable, and some improvements involved trade-offs (e.g., speed vs. long-vowel accuracy). These limitations do not weaken the central observation, but they do narrow the scope of interpretation. What the data support is a realistic claim: regular shadowing practice can generate perceptible prosodic improvement, but the form and extent of that improvement differ by learner and by feature, which is precisely why dual-method triangulation remains useful. Future research should extend this line of inquiry by tracking a larger number of learners over a longer period and by examining whether the prosodic gains observed between repeated attempts are sustained in broader speaking performance over time.

CONCLUSION

This study examined whether autonomous oral practice—implemented through repeated shadowing—can yield observable pronunciation development among intermediate learners of Japanese. Using a dual-perspective design, the analysis combined (a) PRAAT-based inspection of fundamental frequency (F0) contours and temporal measures, and (b) auditory evaluations by native Japanese speakers. Together, these methods allowed the study to address not only whether change occurred, but also what kind of change becomes visible acoustically and what kind becomes salient perceptually.

Three conclusions follow from the findings. First, across learners, the second attempt of the same shadowing material tended to show greater prosodic coherence: pitch movement became less fragmented, phrase-level flow appeared more stable, and native-speaker judges more often perceived the output as improved. This suggests that the value of shadowing in autonomous practice is not limited to accumulating long-term gains; it can also

support within-session adjustment, where learners rapidly recalibrate rhythm and intonation organization after an initial production. Second, improvement was not uniform across learners or features. Some learners moved closer to the model mainly through contour smoothing, while others showed a clearer shift in tempo; at the same time, speed gains could coincide with local losses in timing accuracy (e.g., long-vowel shortening). This pattern indicates that "better" shadowing performance is best treated as a balance between fluency-oriented adaptation and accuracy-oriented control, rather than a single-direction progression. Third, the alignment between acoustic patterns and auditory judgments was informative: native listeners appeared to reward overall organization and naturalness even when individual acoustic indicators did not move in the same direction for every learner. For evaluating autonomous training, this supports the practical relevance of combining acoustic evidence with perceptual evidence, because neither alone fully captures what counts as improvement in communicative terms.

These conclusions connect naturally with earlier work that has argued for the effectiveness of shadowing while also emphasizing methodological triangulation. For example, studies employing both auditory and acoustic approaches have noted that perceptually "better" speech does not always equal full attainment of micro-level targets, and that improvement may involve changes in global delivery rather than isolated segmental fixes. The present results are compatible with that view: perceptual gains were robust, while acoustic changes highlighted trade-offs and learner-specific pathways. At the same time, the scope of the present study should be defined carefully. Because the analysis is based on three intermediate learners, the findings should be understood as learner-specific patterns observed within this dataset rather than as broadly generalizable tendencies across all learners of Japanese. In this sense, the study offers a close description of how prosodic adjustment may emerge under learner-managed shadowing, rather than a claim of universal effectiveness.

In terms of contribution, the study advances Japanese language education in a modest but concrete way. It offers evidence that even under limited and learner-managed conditions, shadowing can support prosodic development that is meaningful to native listeners, while also clarifying why simplistic claims ("shadowing makes pronunciation better") are insufficient: improvement emerges through different profiles and may involve trade-offs. This perspective is useful for educators who adopt shadowing as a routine supplement to pronunciation training, and for learners who practice independently and need realistic expectations about what improves first. Future research should extend this line of inquiry by following a larger number of learners over a longer period, examining whether the gains observed between repeated attempts are retained over time, and exploring what kinds of feedback or instructional support help learners maintain both fluency and accuracy during autonomous shadowing practice.

REFERENCES

Takahashi, E., Fukuda, N., Iwashita, M., & Sakoda, K. (2010). *Jokyu reberu gakushusha ni taisuru shadoingu no kenkyu: Gakushusha no kidzuki to kyoshi no shien* [A Study of Shadowing for Advanced Language Learners : Learners' noticing and instructors' support]. *Bulletin of the Graduate School of Education, Hiroshima University, Part II (Arts and Science Education)*, *59*, 299–308. https://doi.org/10.15027/31048

Takahashi, E., Fukuda, N., Iwashita, M., & Sakoda, K. (2010). *Jokyu reberu gakushusha ni taisuru shadoingu no kenkyu: Gakushusha no kidzuki to kyoshi no shien* [A Study of Shadowing for Advanced Language Learners : Learners' noticing and instructors' support]. *Bulletin of the Graduate School of Education, Hiroshima University, Part II (Arts and Science Education)*, *59*, 299–308. https://doi.org/10.15027/31048

Yin, S. (2015). *Cong jiaoshi tingjue pingjia he yuyin fenxi ruanjian kan gendu fa xunlian dui gaishan Riyu yunlu cuowu de xiaoguo* [Effects of shadowing

training on correcting Japanese prosodic errors from the perspectives of teachers' auditory evaluation and acoustic analysis software]. *Foreign Language Teaching Theory and Practice*, *4*, 70–74.

Okumura, K. (2015). *Betonamu-jin Nihongo gakushusha no shadoingu onse no henka: Akusento inneshon no bunseki kara* [Changes in the "Shadowing Voice" of Vietnamese Learners of Japanese: Through an analysis of accent and intonation]. *Kanagawa University Language Studies*, 38, 23–41.

Wang, P., & Hu, X. (2019). *Shadoingu ni yoru Nihongo gengo noryoku no kojo no jisshoteki kenkyu: Shokyu reberu daigaku Nichigo no gakushusha o taisho ni* [The empirical study on improving Japanese language skills by shadowing: For learners of beginner level in China]. *Bulletin of the Institute of Social Systems Studies, Chuo Gakuin University*, 20(1), 1–9.

Widyaningtyas, I. H. (2017). Tanggapan Mahasiswa terhadap Kegiatan Shadowing dalam Mata Kuliah Choudokkai. *CHI'E Jurnal Pendidikan Bahasa Jepang (Journal of Japanese Learning and Teaching)*, 5(1), 6-8. https://doi.org/10.15294/chie.v5i1.10411

Sumiyoshi, H., & Svetanant, C. (2017). Motivation and Attitude Towards Shadowing: Learners Perspectives in Japanese as a Foreign Language. *Asian-Pacific Journal of Second and Foreign Language Education*, *2*(1), 16. https://doi.org/10.1186/s40862-017-0039-6

Sumiyoshi, H. (2019). The Effect of Shadowing: Exploring the Speed Variety of Model Audio and Sound Recognition Ability in the Japanese as a Foreign Language Context. *Electronic Journal of Foreign Language Teaching*, *16*(1), 5-21. https://doi.org/10.56040/hdsm1611

Utsugi, A. (2022). *Praat nyumon* [Introduction to Praat]. https://sites.google.com/site/utsakr/Home/praat

Matsumi, N., Han, X., Yu, Y., Sato, T., Fei, X., Tome, M., & Matsubara, A. (2013). *Chugoku kokunai no chukyu Nihongo gakushusha ni okeru shadoingu no yukosei: Shadoingu kunren to ondoku kunren o donyu shita jikken jugyo o toshite* [The Effectiveness of the Shadowing for Intermediate Chinese Learners of Japanese in China : By using the shadowing and oral repeated reading training in Japanese classes]. *Hiroshima Journal of School Education*, *19*, 113–122. https://doi.org/10.15027/34708

Hamada, Y. (2019). Shadowing: What is it? How to use it. Where will it go?. *RELC Journal*, *50*(3), 386-393. https://doi.org/10.1177/0033688218771380

Yanagimoto, D. (2024). *Nihongo gakushusha no shadoingu suikoji no joho shori katei: Chokakuteki go ninchi sokudo no kanten kara* [Cognitive Processing in Shadowing Performance of Japanese Language Learners: Insights from Auditory Word Recognition Speed]. *The Annual Bulletin of the International Centre, Gakushuin University*, 10, 5–14.

Vasilevna, K. A. (2025). Methodology of applying shadowing in Japanese language teaching to enhance listening proficiency. *Modern American Journal of Linguistics, Education, and Pedagogy*, *1*(3), 189-200.

Appendix. Auditory Evaluation Questionnaire (English Translation) — Compact Table

Item	Rating scale (options)
1. Accent (accuracy of pitch accent within words)	☐ Many errors overall ☐ Several clear errors and many minor errors ☐ Few clear errors but many minor errors ☐ Few clear errors and few minor errors ☐ Almost no errors
2. Intonation (naturalness of overall pitch movement in the sentence)	☐ Many errors overall ☐ Several clear errors and many minor errors ☐ Few clear errors but many minor errors ☐ Few clear errors and few minor errors ☐ Almost no errors
3. Word Rhythm (accuracy and fluency of special morae such as geminates, nasals, contracted sounds, and long vowels)	☐ Not adequately produced overall ☐ Some parts below standard, though some words are relatively fluent

	☐ Most words are relatively fluent, though some are below standard ☐ Fluent overall
4. Overall Improvement (comparison between first and second shadowing attempts)	☐ No improvement observed ☐ Slight improvement observed ☐ Many areas improved ☐ Clearly improved overall