A Comparison of Sound Change Production Between a Native English Speaker and Japanese University Students:

An Acoustic Analysis of Elision

英語母語話者と日本人大学生の音変化産出頻度の比較

—脱落現象に焦点を当てた音響分析—

山　上　英　絵*

Hanae YAMAGAMI

1. Introduction

Unlike written language, spoken language is often produced as a continuous stream of sounds, making it challenging to understand a speaker if one is not accustomed to listening to natural speed of speech. This difficulty arises because, in natural speech, sounds often influence one another, making it difficult to distinguish each sound, even when one knows the pronunciation of each word (Nakamura, 1997, p. 39). These connected speech phenomena, known as sound changes, occur when neighboring sounds influence each other, resulting in the connection and deletion of sounds.

Although there are exceptions, those who can produce certain sounds are more likely to be able to perceive those sounds (Baghrahi et al., 2014, p. 208; Nakamura, 1997, p. 40). This theory is based on the "perception-production loop", which explains that the way individuals produce a certain sound results from the mental registration of the sounds they have heard over time (Beddor, 2015, para 1). This means that if someone can produce a particular sound, they are also able to perceive that sound registered in their mind.

Despite its importance in communication, Japanese English textbooks do not explicitly cover sound change, and its instruction highly depends on the teacher (Ueda & Otsuka, 2014, p. 13). With limited time and resources, teachers are required to effectively teach sound change to students; thus, it is paramount to identify the types of sound change that are particularly challenging for Japanese students.

There are many interesting types of sound change across various languages, attracting the attention of researchers. However, due to this large variety of sound changes, it has been challenging to compare and analyze data on a consistent scale. To address this challenge, Yamagami (2023) systematically categorized common phonetic sound changes into three hierarchical tree diagrams: assimilation (Figure 1), elision (Figure 2), and linking (Figure 3). These diagrams are organized from broader to narrower categories, allowing us to understand different types of sound change systematically and succinctly. Despite this advantage,

* やまがみ　はなえ　文学研究科英文学専攻博士前期課程
指導教員：中西　弘

some of these sound changes may remain theoretical, and whether they actually occur in daily conversation remains unclear. Thus, it is also crucial to collect speech audio from native speakers to investigate the occurrences of these sound changes in natural speech.

This paper focuses specifically on elision among the three types of sound change discussed in Yamagami (2023). Compared to assimilation and linking, elision is more likely to hinder listeners' ability to catch and process information. While assimilation and linking involve changes that alter or connect sounds, they do not result in the disappearance of sounds. In contrast, elision involves the omission of sounds, which Crystal (2003, p. 158) defines as the "omission of sounds in connected speech" (cited in Hasan, 2012, p. 454). This characteristic of elision can impair learners' ability to perceive and produce it. Additionally, Young Jr. (2015, p. 37) explains that although elision exists in Japanese as well, it occurs for reasons different from those in English. According to him, while elision in English serves to maintain the natural rhythm of speech, Japanese elision arises from a preference for light syllables, where a single consonant is followed by a vowel (C+V). This difference may also affect Japanese learners' ability to perceive and produce English elisions.

Building on the discussions above, this paper aims to: 1) investigate whether native speakers consistently produce the sound changes documented in the literature, and 2) identify specific types of elision that pose challenges for Japanese English learners in production. The research questions are presented below.

1) To what extent do native speakers consistently produce elisions as documented in the literature?
2) Which types of elision pose challenges for Japanese learners of English in production?

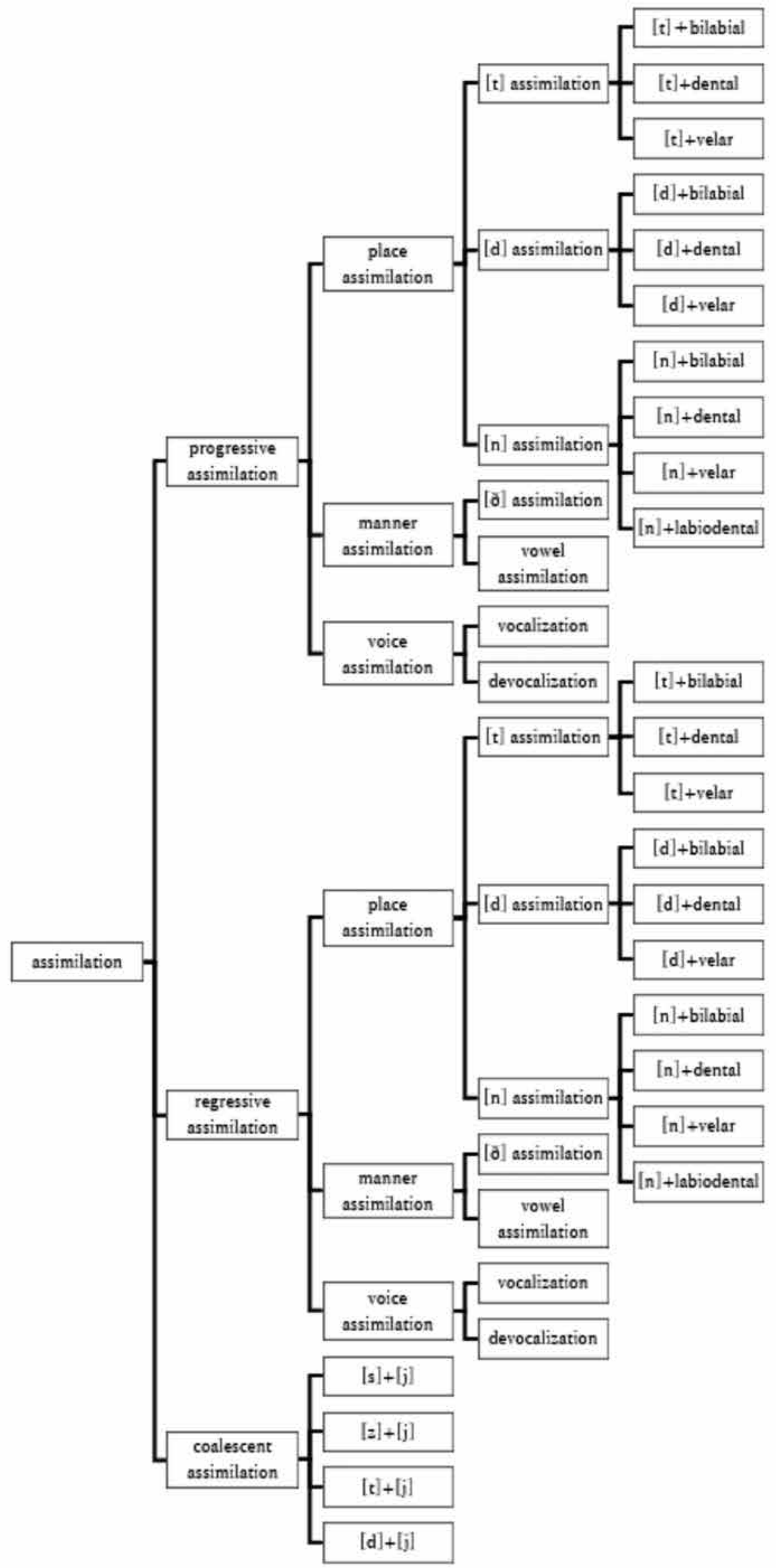

Figure1. Tree diagram of assimilation classification

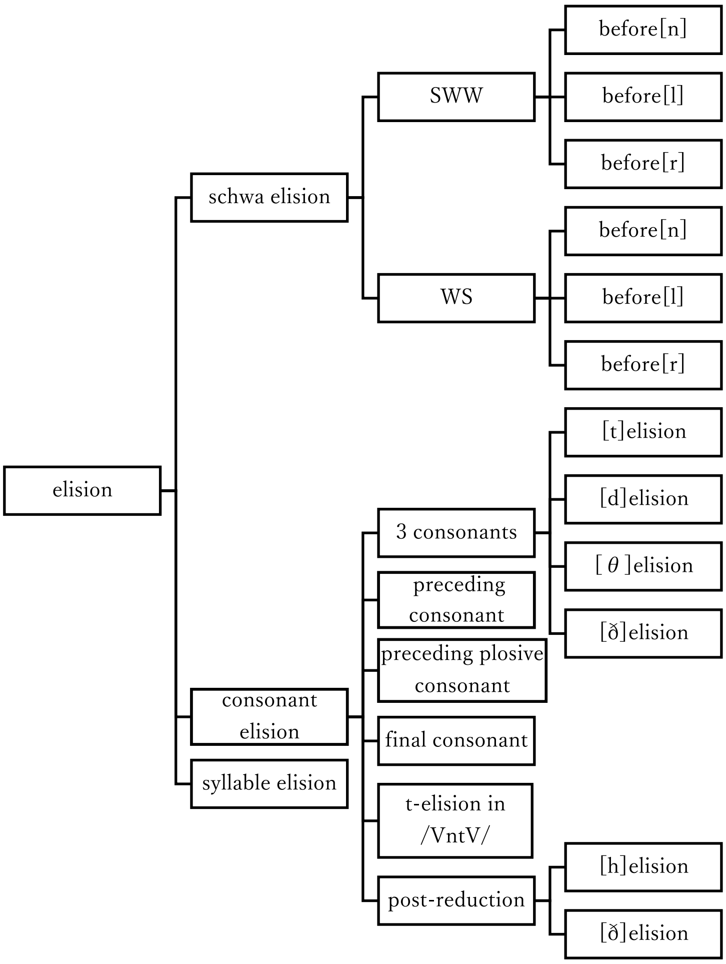

Figure2. Tree diagram of elision classification

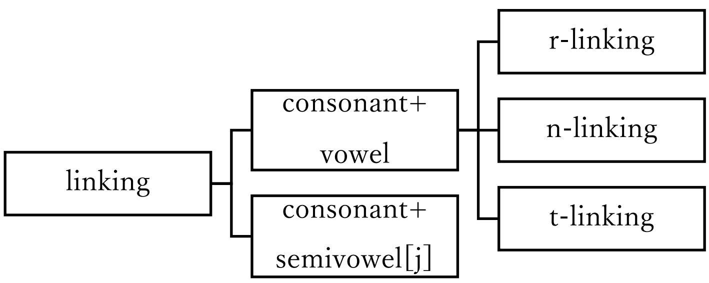

Figure3. Tree diagram of linking classification

2. Types of elision

This section will explore the types of elision which will be examined in this paper. In Yamagami (2023), elision is classified into as many detailed branches as possible, down to the specific sounds involved in each sound change. However, this paper focuses on five broader categories due to limitations in the available stimuli: a) schwa elision, b) preceding consonant elision, c) preceding plosive consonant elision, d) post-reduction elision, and e) t-elision in /VntV/ (where V refers to a vowel).

First, *schwa elision* refers to the omission of the schwa sound, identified as [ə] in the IPA, making it inaudible. For instance, the schwa [ə] in the word *fashion* [fæʃən] is often reduced and dropped, resulting in the pronunciation [fæʃn], especially in fast and casual speech. Figure 4 demonstrates the spectrogram of the word *fashion* [fæʃn] produced by a native speaker, showing the reduction and elision of [ə].

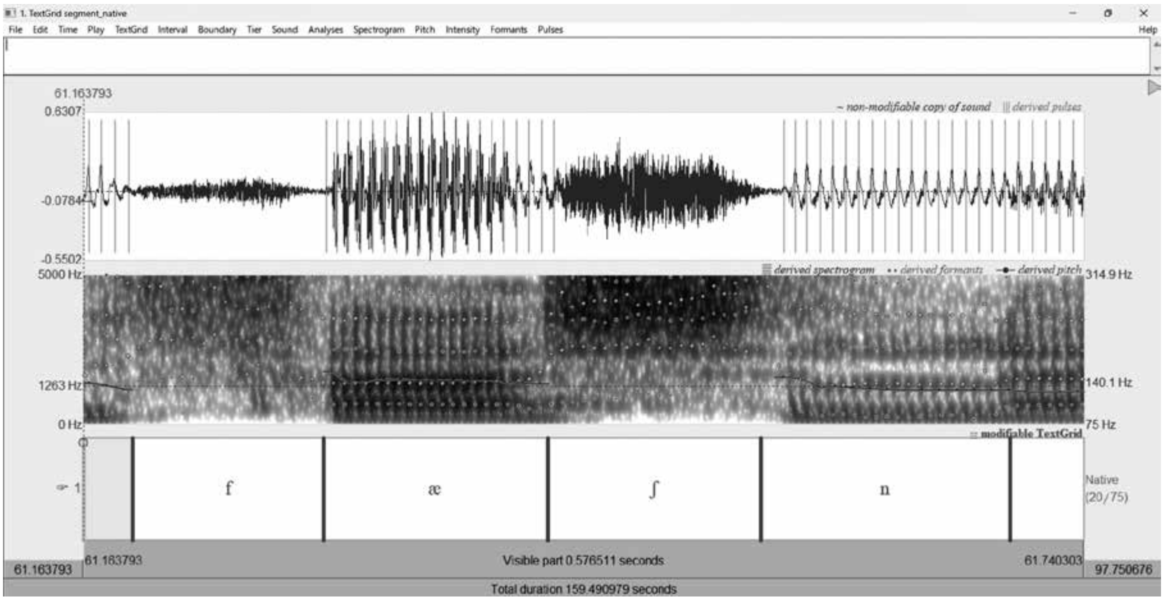

Figure 4. Spectrogram of *fashion* [fæʃn] by a native speaker

Second, *preceding consonant elision* occurs when two identical consonants appear consecutively between two words, as in *hot tea* [hɑːt tiː], *top page* [tɑp peɪdʒ], and so on. In this environment, the preceding consonant is likely to be dropped, leaving only the following consonant. This leads to the pronunciations [hɑ tiː] and [tɑ peɪdʒ], respectively. Figure 5 shows the spectrogram of the word *hot tea* [hɑ tiː] produced by a native speaker. Here, the first [t] is dropped.

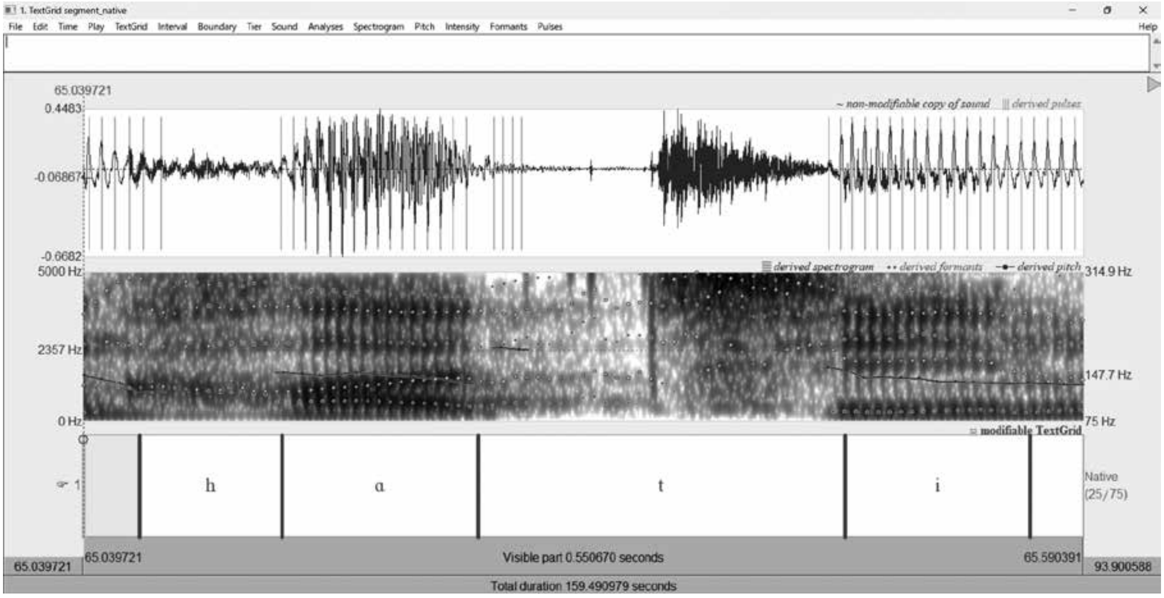

Figure 5. Spectrogram of *hot tea* [hɑ tiː] by a native speaker

Third, *preceding plosive consonant elision* is observed when a plosive consonant is followed by another consonant between two words. Examples of this include [k] and [f] in *ask for* [æsk fɔːr], or [g] and [t] in *jog to* [dʒɑːg tuː]. In such cases, the preceding plosive consonants [k] and [g] are likely to be dropped, leaving only the following consonants [f] and [t]. These changes result in the pronunciations [æs fɔːr] and [dʒɑ tuː], respectively. Figure 6 presents the spectrogram of the word *ask for* [æs fɔːr] produced by a native speaker, with the [k] sound omitted.

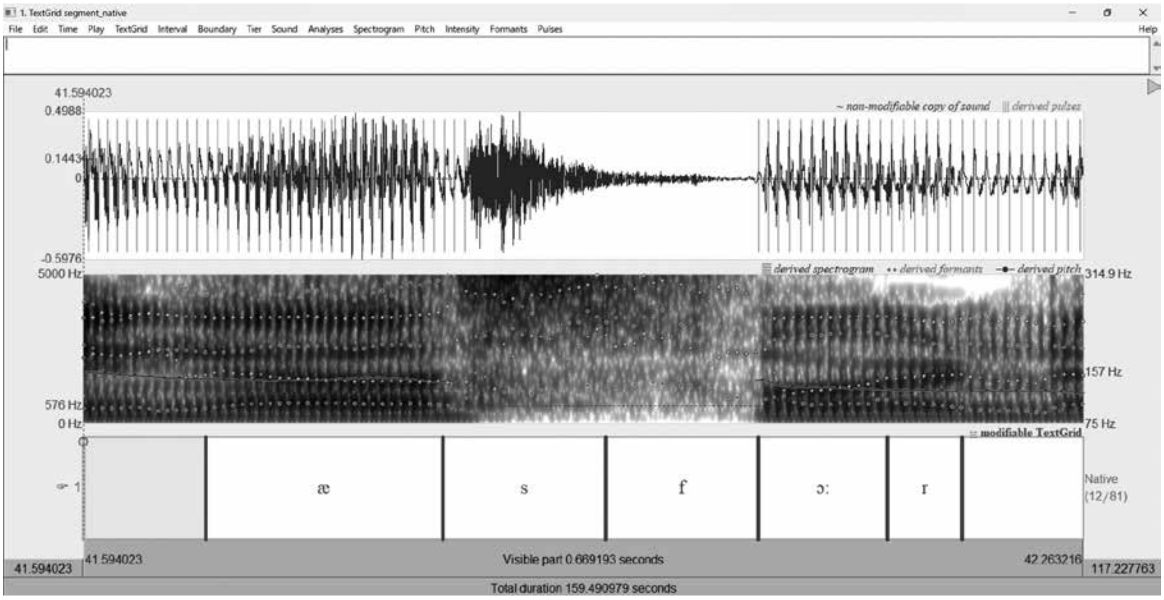

Figure 6. Spectrogram of *ask for* [æs fɔːr] by a native speaker

Fourth, *post-reduction elision* happens when sounds such as [h] and [ð] are dropped following their reduction. These sounds are typically found in pronouns such as *him* or *them*. When these pronouns appear in speech, as in *call him* [kɑːl hɪm] or *tell them* [tel ðem], they tend to be weakened, eventually leading to the elision of [h] or [ð], resulting in pronunciations like [kɑːl ɪm] or [tel əm]. Figure 7 shows the spectrogram of *call him* [kɑːl ɪm] produced by a native speaker, where [h] is reduced and dropped.

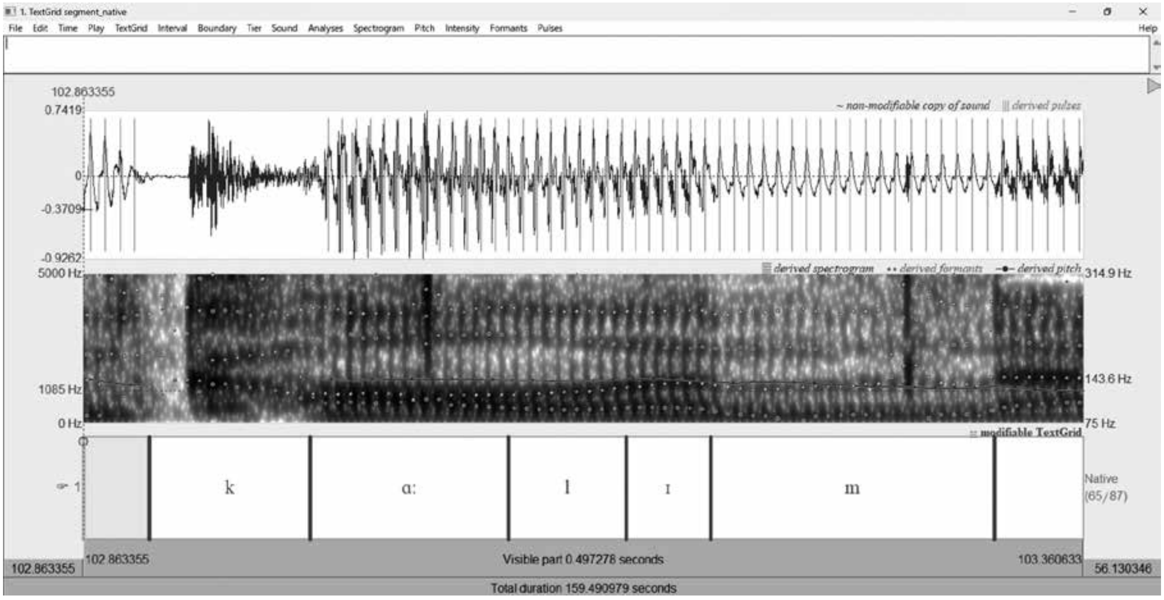

Figure 7. Spectrogram of *call him* [kɑːlɪm] by a native speaker

Finally, in General American (GA), *t-elision in /VntV/* occurs when an [nt] cluster is positioned between two vowels, as in *center* [sɛnt̬ɚ]. In this context, the [nt] cluster is influenced by the surrounding vowels, leading to a change in pronunciation to [ɾ̃ɾ]. Progressive assimilation then takes place between these two sounds, with [ɾ] assimilating to the preceding nasalized [ɾ̃], resulting in [ɾ̃ɾ̃]. Consequently, the [t] becomes inaudible, and *center* [sɛnt̬ɚ] is pronounced as [sɛnɚ]. Figure 8 illustrates the spectrogram of the word *center* [sɛnɚ] produced by a native speaker, where the [t] is assimilated and elided.

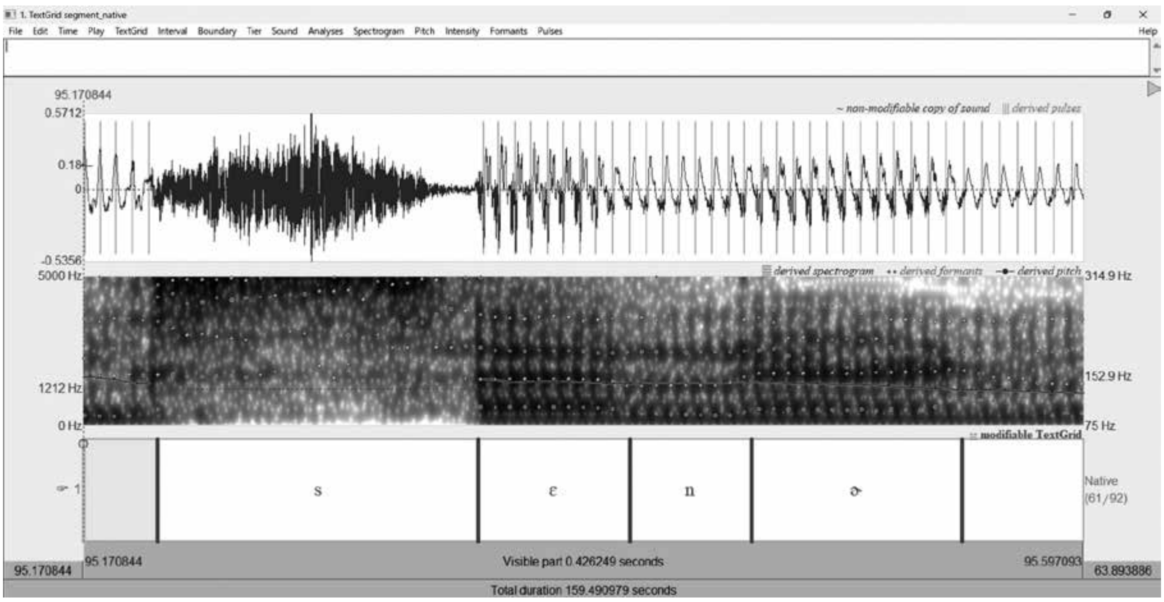

Figure 8. Spectrogram of *center* [sɛnɚ] by a native speaker

3. Methodology

3.1. Participants

A native English speaker from Canada participated in this research to observe the occurrences of elision. For comparison, ten Japanese university students, who are not majoring in English, were also selected. Five students have a B1 level according to the Common European Framework of Reference for Languages (CEFR), while the other five have an A1 level. Students from different CEFR levels were included in this research to examine the differences among Japanese students based on their English proficiency.

3.2. Selection of stimuli

The present study focused on the five categories of potential elision environments in connected speech. For each category, simple lexical items were carefully selected as experimental stimuli (e.g., *certainly* for schwa elision, *top pick* for preceding consonant elision, *look to* for preceding plosive consonant elision, *call him* for post-reduction elision, and *center* for t-elision in /VntV/). All target items were embedded within the carrier sentence, "I'm reading _____ this time." Speech samples were collected from the native English speaker and Japanese learners of English, who read the carrier sentences aloud. To regulate the speech speed, the native speaker participant was first recorded reading dummy and real samples, with instructions to speak faster than their normal pace. The Japanese participants were then instructed to listen to the recorded audio of dummy samples produced by the native speaker before performing the reading task at a similar speed. These productions were digitally recorded and analyzed acoustically using Praat software.

a. Schwa elision

In the following sections, the conditions and rationale for choosing the words presented in Tables 9 through 13 will be explained.

For schwa elision, three conditions were established: 1) [n] follows [ə]; 2) each word contains 2 or 3 syllables; 3) the consonant preceding [ə] is equally represented by voiced and voiceless sounds. The first condition ensures consistency across all words. The second condition examines the influence of the number of syllables on the results. The third condition investigates the impact of voicing on the outcomes.

t+ə+n	3 syllables 2 syllables	cert(ai)nly cert(ai)n	d+ə+n	3 syllables 2 syllables	sudd(e)nly sudd(e)n
k+ə+n	3 syllables 2 syllables	balc(o)ny bac(o)n	g+ə+n	3 syllables 2 syllables	drag(o)ness drag(o)n
m+ə+n	3 syllables 2 syllables	comm(o)nly comm(o)n	l+ə+n	3 syllables 2 syllables	col(o)ny poll(e)n
p+ə+n	3 syllables 2 syllables	op(e)ning op(e)n	ʃ+ə+n	3 syllables 2 syllables	nat(io)nal fash(io)n

Table 9. Word items for schwa elision

b. Preceding consonant elision

Three conditions were set for preceding consonant elision: 1) each word contains one syllable; 2) the relevant consonants are either bilabial or alveolar; and 3) each group (bilabial and alveolar) includes consonants with the same place of articulation but different voicing (voiced/voiceless). The first condition ensures uniformity in the number of syllables across all words. The second condition investigates differences in production between consonants with the same manner of articulation but different places of articulation. Finally, the third condition examines patterns of consonant elision between voiced and voiceless consonants.

Bilabial	/p/+/p/	to(p) pick	to(p) page
	/b/+/b/	we(b) belt	we(b) book
Alveolar	/t/+/t/	ho(t) take	ho(t) tea
	/d/+/d/	foo(d) day	foo(d) drive

Table 10. Word items for preceding consonant elision

c. Preceding plosive consonant elision

For preceding plosive consonant elision, three conditions were considered: 1) each word contains one syllable; 2) the preceding plosive consonants are either voiced or voiceless; and 3) the following consonants are either plosives or fricatives. The first condition ensures consistency across all stimuli, while the second and third conditions aim to clarify whether voicing and the manner of articulation influence the realizations of elision.

/k/+/t/	loo(k) to	as(k) to
/g/+/t/	hu(g) to	jo(g) to
/k/+/f/	loo(k) for	as(k) for
/g/+/f/	hu(g) for	jo(g) for

Table 11. Word items for preceding plosive elision

d. Post-reduction elision

The stimuli for post-elision are based on three conditions: 1) each word consists of one syllable, 2) the first word is either *call* or *tell*, and 3) the following word is either *him* or *them*. The first condition ensures consistency across all words, while the second and third conditions focus on words where this type of elision commonly occurs (Katayama et al., 1996, p. 127).

/l/+/h/	call (h)im	tell (h)im
/l/+/ð/	call (th)em	tell (th)em

Table 12. Word items for post-reduction elision

e. T-elision in /VntV/

Three conditions were set for t-elision in /VntV/: 1) all words contain two syllables; 2) [t] appears in either /nter/ or /nty/; and 3) the vowels preceding [n] are either [ɛ] or [aʊ]. The first condition ensures consistency across all stimuli, while the second and third conditions examine the influence of surrounding environments and neighboring sounds on the occurrence of [t] elision. /nter/ and /nty/ are common environments for this type of elision, and [ɛ] and [aʊ] were selected because they meet the strict conditions outlined above, despite the limited number of available words.

nter	cen(t)er	coun(t)er
nty	plen(t)y	coun(t)y

Table 13. Word items for t-elision in /VntV/

3.3. Methods

The study employed non-parametric statistical methods to analyze between-group differences in elision production. The test was chosen because it is suitable for comparing three or more independent groups, particularly when the sample size is not large (Liu, 2015, p. 7). The traditional ANOVA assumes a normal distribution of data and equal variances across groups, which can lead to inaccurate results (Hecke, 2012, p. 2). Thus, the Kruskal-Wallis test is appropriate for the present study because the groups are independent, and the data are assumed not to show a normal distribution. In this study, a p-value of less than 0.05 was regarded as indicating significant differences.

Post-hoc analyses were conducted using Dunn's test to identify specific between-group differences (Dinno, 2015, p. 294). The p-value was adjusted using the Bonferroni method to minimize the risk of Type I error (false positives) when performing multiple statistical tests (Armstrong, 2014, p. 502).

4. Results

This section will provide the results on: 1) how the three groups differed in the occurrence of elisions, and 2) which types of elision were easier or more challenging for the Japanese participants.

a. Schwa elision

Table 14 shows the number of elision occurrences of the three groups. Out of 16 instances, the native speaker produced five schwa elisions. The B1 group showed a wider range, from 0 to 6 occurrences, while individuals in the A1 group exhibited similar patterns within their group.

Level	Participant	Elision count
	Native	5
B1	Japanese A	4
	Japanese I	6
	Japanese R	3
	Japanese S	0
	Japanese B	0
A1	Japanese O	2
	Japanese H	2
	Japanese K	2
	Japanese D	2
	Japanese N	2

Table 14. Number of schwa elision occurrences

Figure 15 shows the comparison of successful schwa elisions among the native speaker, A1 learners, and B1 learners, while Table 16 presents their basic statistics. The Kruskal-Wallis test revealed a statistically significant difference between the groups ($H(2) = 8.12$, $p = .017$) with a large effect size ($r = .742$). Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed a hierarchical pattern of schwa elision ability. The native speaker demonstrated the highest number of schwa elisions (five elisions), which was significantly different from A1 learners ($p = .007$) with a large effect size ($r = .531$). The difference between the native speaker and B1 learners ($p = .047$) showed a medium-to-large effect size ($r = .371$), though it was marginally significant after Bonferroni correction. B1 learners also performed significantly better than A1 learners ($p = .035$), with a medium-to-large effect size ($r = .396$).

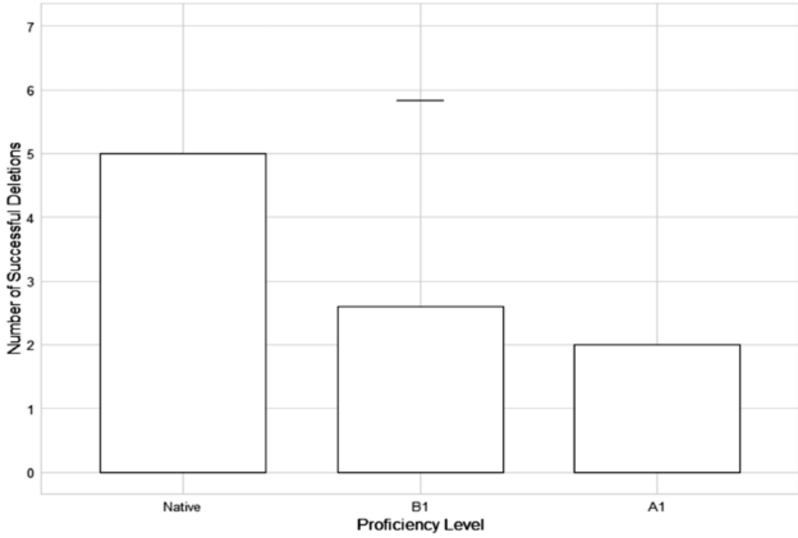

Figure 15. Comparison of successful schwa elisions among the native speaker, B1 learners, and A1 learners

	Native	B1	A1
Mean	5	2.6	2
Median	5	3	2
Standard Deviation	-	2.608	0
Minimum	5	0	2
Maximum	5	6	2
Count	1	5	5

Table 16. Basic statistics of schwa elision occurrences

Differences in schwa occurrences between two-syllable and three-syllable words were examined using the Kruskal-Wallis test across the three groups. Figure 17 illustrates the schwa elision rates in two-syllable words. The results indicated no statistically significant differences among the groups, $H(2) = 4.897$, $p = .086$. Despite the lack of statistical significance, the effect sizes suggested practically meaningful differences between the native speaker and learners. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed no statistically significant pairwise differences between the native speaker and B1 learners ($p = .139$), the native speaker and A1 learners ($p = .102$), or B1 and A1 learners ($p = 1.000$). However, effect size calculations provided additional insights into the practical significance of these differences. The comparison between the native speaker and A1 learners showed a medium effect size ($r = .262$), as did the comparison between the native speaker and B1 learners ($r = .247$). In contrast, the difference between B1 and A1 learners showed a negligible effect size ($r = .042$).

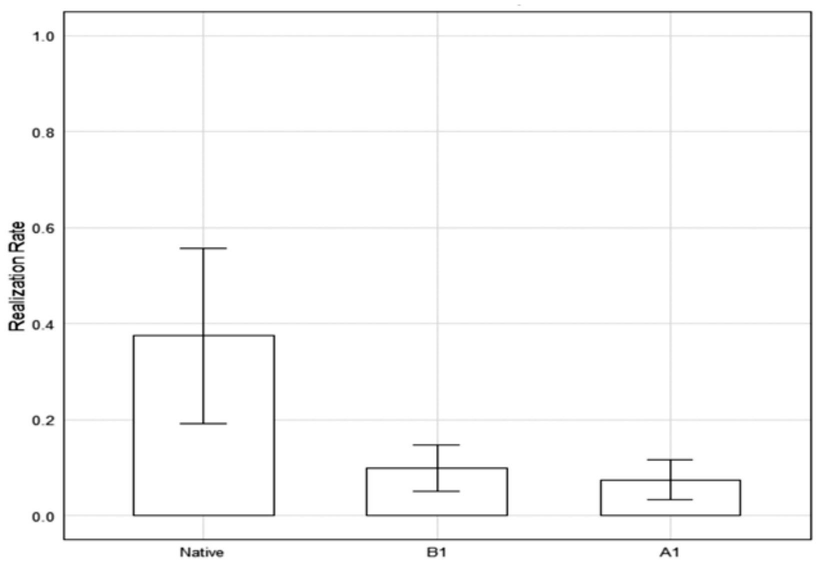

Figure 17. Schwa elision rates in two-syllable words

Figure 18 shows the schwa elision rates in three-syllable words. The Kruskal-Wallis test showed no statistically significant differences among the groups, H(2) = 0.368, p = .832. The effect sizes for all comparisons were notably small, suggesting minimal practical differences between the groups in their elision patterns for three-syllable words. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons confirmed the absence of significant differences between any pair of groups (all p-values = 1.000). Effect size calculations further supported these findings, with small effect sizes observed between the native speaker and B1 learners (r = .082), the native speaker and A1 learners (r = .108), and a negligible effect size between B1 and A1 learners (r = .033).

While medium differences were observed between the native speaker and A1 learners, as well as between the native speaker and B1 learners in three-syllable words, these differences were small in two-syllable words. Although not statistically significant, these findings suggest that Japanese students are more capable of producing schwa elision similar to the native speaker in three-syllable words than in two-syllable words.

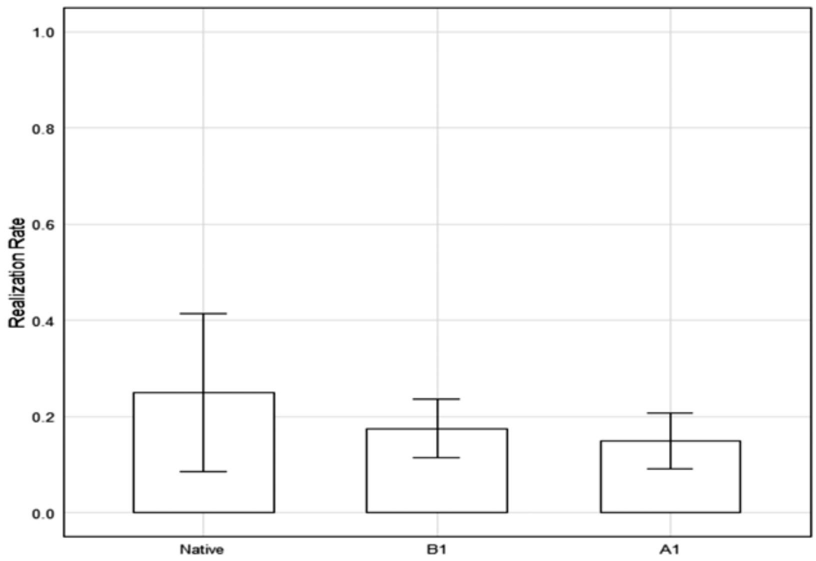

Figure 18. Schwa elision rates in three-syllable words

The differences in schwa elision rates in voiceless consonant + schwa contexts and voiced consonant + schwa contexts among the three groups were also examined using the Kruskal-Wallis test. Figure 19 presents the schwa elision rates in voiceless consonant + schwa contexts. The Kruskal-Wallis test revealed a significant difference between groups (H(2) = 6.82, p = .033). Follow-up Dunn's tests with Bonferroni correction showed that the native speaker's performance was significantly different from that of A1 learners (p = .013, r = .386) and B1 learners (p = .024, r = .354), with medium effect sizes in both comparisons. No significant difference was found between B1 and A1 learners (p = 1.000, r = .035), with a negligible effect size.

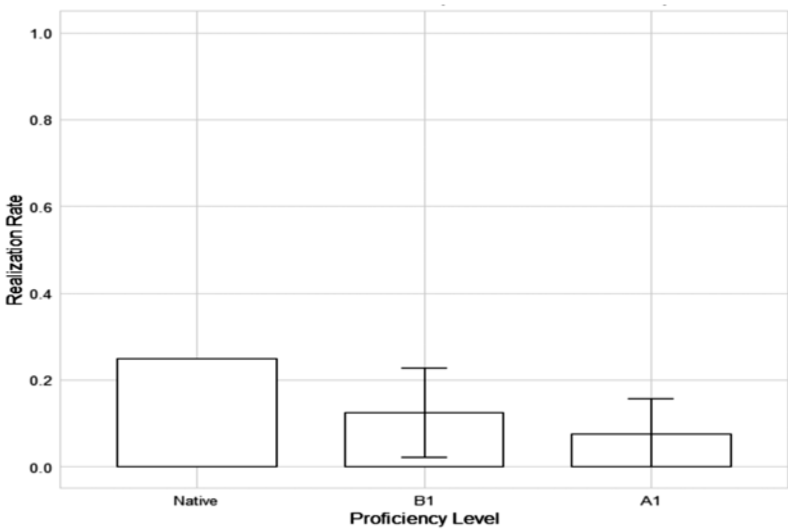

Figure 19. Schwa elision rates in voiceless consonant + schwa contexts

Figure 20 illustrates the schwa elision rates in voiced consonant + schwa contexts. The Kruskal-Wallis test revealed significant differences between the groups (H(2) = 7.94, p = .019). Dunn's tests with Bonferroni correction showed that the native speaker's performance was significantly higher than that of both A1 learners (p = .010, r = .398) and B1 learners (p = .018, r = .370), with medium effect sizes in both comparisons. No significant difference was found between B1 and A1 learners (p = 1.000), with a negligible effect size (r = .031).

Regarding the differences based on the voicing of the preceding consonant, Japanese students exhibited similar rates of elision in both voiceless consonant + schwa and voiced consonant + schwa contexts.

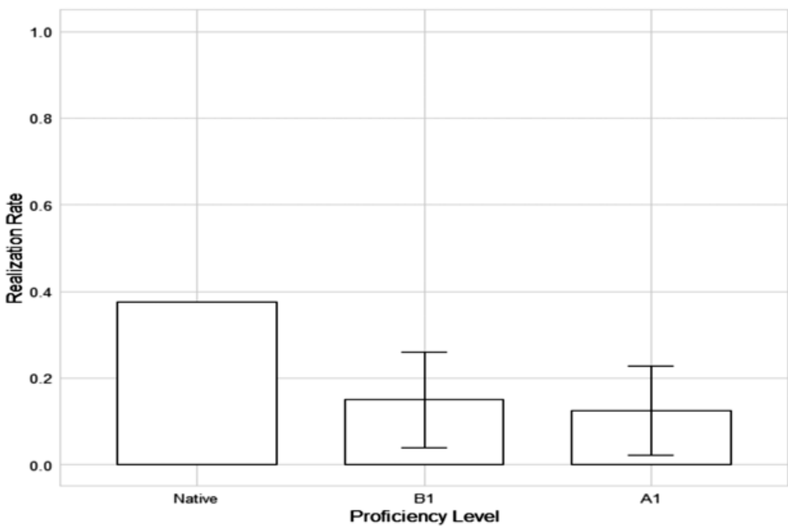

Figure 20. Schwa elision rates in voiced consonant + schwa contexts

b. Preceding consonant elision

Table 21 shows the number of preceding consonant elision occurrences of the three groups. Out of eight instances, the native speaker produced consonant elisions in all cases. The B1 and A1 groups also exhibited a reasonable number of elisions, with some participants producing as many elisions as the native speaker.

Level	Participant	Elision count
	Native	8
B1	Japanese A	7
	Japanese I	8
	Japanese R	3
	Japanese S	5
	Japanese B	8
A1	Japanese O	8
	Japanese H	6
	Japanese K	5
	Japanese D	7
	Japanese N	5

Table 21. Number of occurrences of preceding consonant elision

Figure 22 shows the comparison of successful preceding consonant elisions among the native speaker, A1 learners, and B1 learners, while Table 23 presents their basic statistics. The Kruskal-Wallis test revealed no statistically significant differences between the groups ($H(2) = 1.83$, $p = .401$), despite a medium effect size ($r = .384$). Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed no significant differences between any pair of groups: native speaker vs. B1 learners ($p = .342$, $r = .227$), native speaker vs. A1 learners ($p = .342$, $r = .227$), and B1 vs. A1 learners ($p = 1.000$, $r = .000$). The effect sizes for the comparisons involving the native speaker and the learner groups were small to medium ($r = .227$), while the comparison between the two learner groups showed no effect ($r = .000$), reflecting their identical mean performance (both $M = 6.2$).

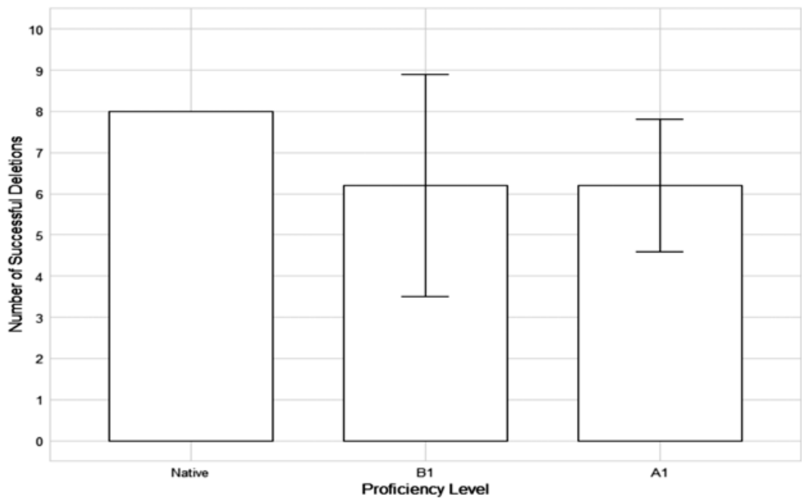

Figure 22. Comparison of successful preceding consonant elisions among the native speaker, B1 learners, and A1 learners

	Native	B1	A1
Mean	8	6.2	6.2
Median	8	7	6
Standard Deviation	-	2.168	1.304
Minimum	8	3	5
Maximum	8	8	8
Count	1	5	5

Table 23. Basic statistics of occurrences of preceding consonant elision

Next, the Kruskal-Wallis test was conducted to observe the different realizations of preceding consonant elision in bilabial and alveolar consonant sequences across the three groups. Figure 24 illustrates the elision rates in bilabial + bilabial and alveolar + alveolar contexts.

For bilabial + bilabial contexts, the results indicated no statistically significant differences, $H(2) = 1.750$, $p = .417$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons confirmed the absence of significant differences between any pair of groups (all p-values > .05). Effect size calculations revealed small to medium effects between the native speaker and B1 learners ($r = .277$), the native speaker and A1 learners ($r = .236$), and a negligible effect between B1 and A1 learners ($r = .054$).

For alveolar + alveolar contexts, the Kruskal-Wallis test also indicated no statistically significant differences, $H(2) = 1.233$, $p = .540$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons confirmed the absence of significant differences between any pair of groups (all p-values = 1.000). Effect size calculations revealed small to medium effects between the native speaker and B1 learners ($r = .236$), the native speaker and A1 learners ($r = .189$), and a negligible effect between B1 and A1 learners ($r = .058$).

The practical differences among the three groups in both bilabial and alveolar consonant sequences followed similar patterns. This suggests that the Japanese learner groups produce preceding consonant elision with similar frequencies in these two environments.

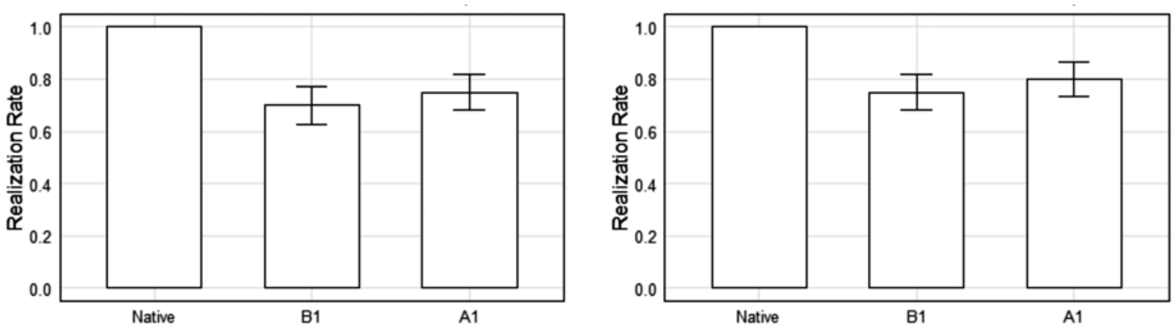

Figure 24. Elision rates in bilabial consonant sequences (left) and alveolar consonant sequences (right)

The differences between voiceless and voiced consonant sequences were also examined using the Kruskal-Wallis test (Figure 25). For voiceless + voiceless contexts, the results indicated no statistically significant differences, $H(2) = 0.857$, $p = .651$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons confirmed the absence of significant differences between any pair of groups (all p-values = 1.000). Effect size calculations revealed small effects between the native speaker and B1 learners (r = .183) and between the native speaker and A1 learners (r = .183), and no effect between B1 and A1 learners (r = .000).

In contrast, for voiced + voiced contexts, the Kruskal-Wallis test indicated statistically significant differences, H(2) = 6.124, p = .047. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners (p = .041), while differences between the native speaker and A1 learners (p = .156) and the differences between B1 and A1 learners (p = .223) were not statistically significant. Effect size calculations showed large effects between the native speaker and B1 learners (r = .456), a medium effect between the native speaker and A1 learners (r = .289), and a small to medium effect between B1 and A1 learners (r = .200).

This comparison suggests that the two Japanese groups produce preceding consonant elision more frequently in voiceless + voiceless contexts than in voiced + voiced contexts, with the B1 group showing a statistical difference and the A1 group showing a practical difference.

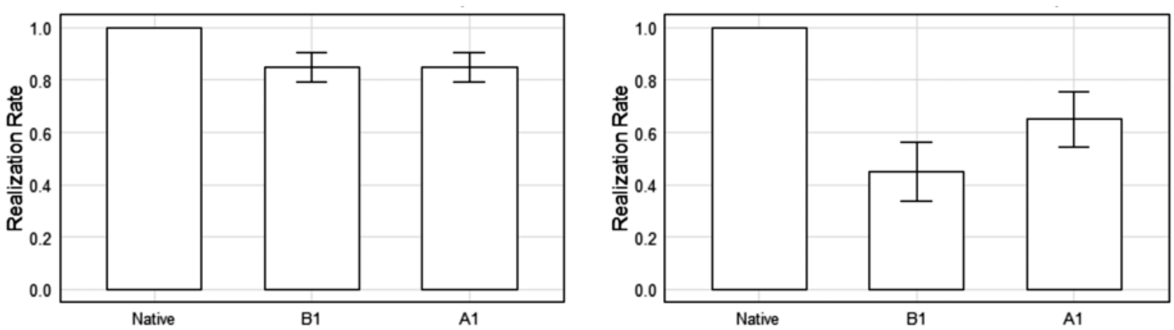

Figure 25. Elision rates in voiceless + voiceless (left) and voiced + voiced (right)

c. Preceding plosive consonant elision

Table 26 shows the number of occurrences of preceding plosive consonant elision of the three groups. Out of eight instances, the native speaker exhibited five consonant elisions. Participants from the B1 and A1 groups demonstrated varying rates of elision.

Level	Participant	Elision count
	Native	5
B1	Japanese A	0
	Japanese I	4
	Japanese R	1
	Japanese S	0
	Japanese B	0
A1	Japanese O	2
	Japanese H	5
	Japanese K	1
	Japanese D	0
	Japanese N	1

Table 26. Number of occurrences of preceding plosive consonant elision

Figure 27 shows the comparison of successful preceding plosive consonant elisions among the native speaker, A1 learners, and B1 learners, while Table 28 presents their basic statistics. The Kruskal-Wallis test revealed a statistically significant difference between the groups (H(2) = 6.94, p = .031), with a large effect size (r = .502). Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed that the native speaker's performance (5 elisions) was significantly different from that of the B1 learners (p = .022), with a medium-to-large effect size (r = .459). The difference between the native speaker and A1 learners was not statistically significant but showed a trend toward significance (p = .051), with a medium effect size (r= .397). No significant difference was found between B1 and A1 learners (p = 1.000), with a negligible effect size (r = .062).

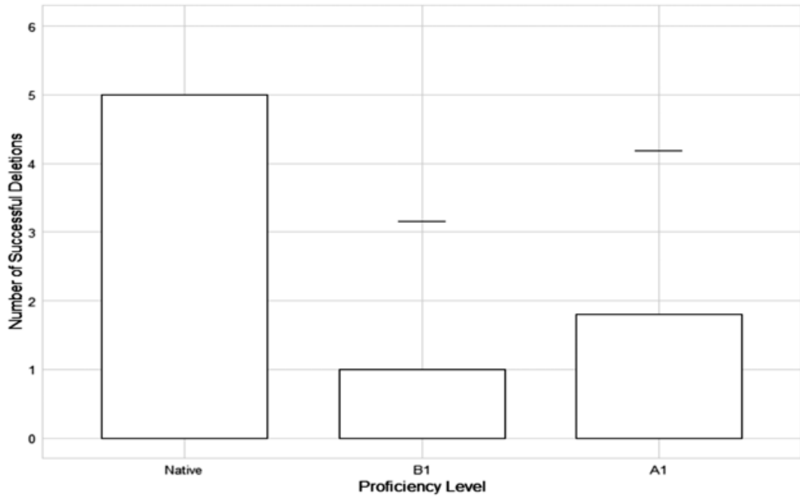

Figure 27. Comparison of successful preceding plosive consonant elisions among the native speaker, B1 learners, and A1 learners

	Native	B1	A1
Mean	5	1	1.8
Median	5	0	1
Standard Deviation	-	1.732	1.924
Minimum	5	0	0
Maximum	5	4	5
Count	1	5	5

Table 28. Basic statistics of occurrences of preceding plosive consonant elision

The differences between voiceless + voiceless and voiced + voiceless contexts were analyzed using the Kruskal-Wallis test to observe the realization rates among the three groups (Figure 29). For the voiceless + voiceless context, the results indicated statistically significant differences, H(2) = 8.952, p = .011. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners (p = .009), and between the native speaker and A1 learners (p = .023), while no statistically significant difference was found between B1 and A1 learners (p = 1.000). Effect size calculations revealed large effects between the native speaker and B1 learners (r = .548), the native speaker and A1 learners (r = .471), and a small effect between B1 and A1 learners (r = .071).

On the other hand, for voiced + voiceless contexts, the results indicated no statistically significant differences, $H(2) = 2.458$, $p = .293$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed no significant differences between any pair of groups (all p-values > .05). Effect size calculations revealed medium effects between the native speaker and B1 learners ($r = .316$), a small to medium effect between the native speaker and A1 learners ($r = .224$), and a small effect between B1 and A1 learners ($r = .122$).

These results suggest that Japanese students produce preceding consonant elision more similarly to the native speaker in voiced + voiceless contexts than in voiceless + voiceless contexts.

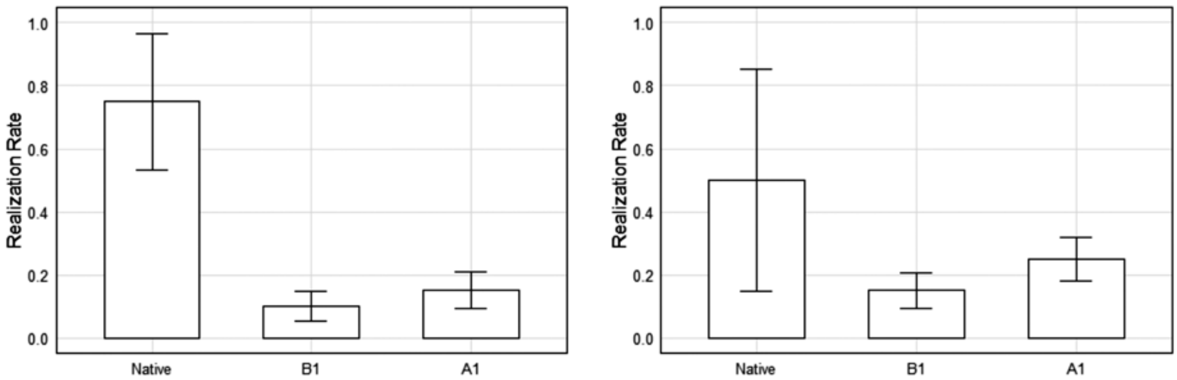

Figure 29. Elision rates of voiceless + voiceless (left) and voiced + voiceless (right)

The differences in elision occurrences between plosive + plosive and plosive + fricative contexts were also examined (Figure 30). For the plosive + plosive contexts, the Kruskal-Wallis test indicated statistically significant differences, $H(2) = 15.327$, $p < .001$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners ($p < .001$), and between the native speaker and A1 learners ($p = .002$). The difference between B1 and A1 learners was not statistically significant ($p = .682$). Effect size calculations revealed very large effects between the native speaker and B1 learners ($r = .745$), the native speaker and A1 learners ($r = .632$), and a small effect between B1 and A1 learners ($r = .183$).

On the other hand, for the plosive + fricative contexts, the Kruskal-Wallis test showed no statistically significant differences, $H(2) = 0.687$, $p = .709$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons confirmed the absence of significant differences between any pair of groups (all p-values > .05). Effect size calculations revealed small effects between the native speaker and B1 learners ($r = .122$), no effect between the native speaker and A1 learners ($r = .000$), and a small effect between B1 and A1 learners ($r = .122$).

The results suggest that Japanese students produce preceding consonant elision more frequently in plosive + fricative contexts than in plosive + plosive contexts.

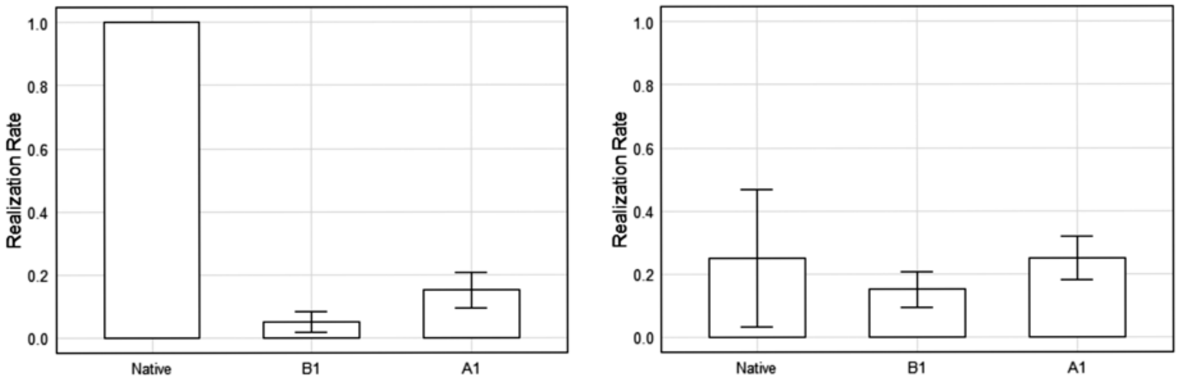

Figure 30. Elision rates of plosive + plosive (left) and plosive + fricative (right)

d. Post-reduction elision

Table 31 shows the number of post-reduction elision occurrences of the three groups. While the native speaker exhibited three elisions out of four instances, no participants from the B1 and A1 groups produced elisions in this environment.

Level	Participant	Elision count
	Native	3
B1	Japanese A	0
	Japanese I	0
	Japanese R	0
	Japanese S	0
	Japanese B	0
A1	Japanese O	0
	Japanese H	0
	Japanese K	0
	Japanese D	0
	Japanese N	0

Table 31. Number of occurrences of post-reduction elision

Figure 32 shows the comparison of successful post-reduction elisions among the native speaker, A1 learners, and B1 learners, while Table 33 presents their basic statistics. The Kruskal-Wallis test revealed a highly significant difference between the groups ($H(2) = 20.00$, $p < .001$) with a large effect size ($r = .853$). Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed that the native speaker's performance (3 deletions) was significantly different from both B1 learners ($p < .001$, $r = .592$) and A1 learners ($p < .001$, $r = .592$), with large effect sizes in both comparisons. There was no difference between B1 and A1 learners ($p = 1.000$, $r = .000$) as neither group achieved any successful deletions.

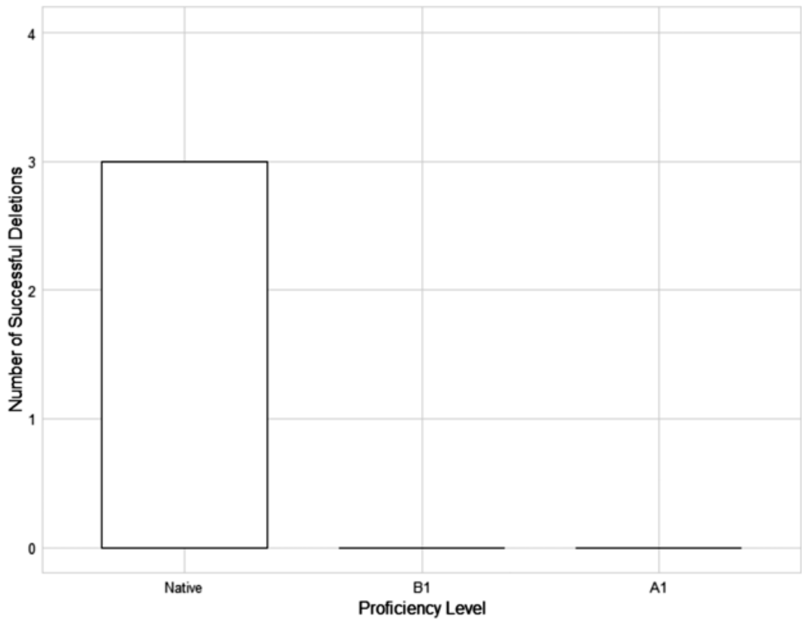

Figure 32. Comparison of successful post-reduction elisions among the native speaker, B1 learners, and A1 learners

	Native	B1	A1
Mean	3	0	0
Median	3	0	0
Standard Deviation	-	-	-
Minimum	3	0	0
Maximum	3	0	0
Count	1	5	5

Table 33. Basic statistics of occurrences of post-reduction elision

The differences in h-elision and ð-elision occurrences among the three groups were examined (Figure 34). For h-elision, the Kruskal-Wallis test indicated statistically significant differences, $H(2) = 20.000$, $p < .001$. Effect sizes were calculated to assess the practical significance of the group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners ($p < .001$), as well as between the native speaker and A1 learners ($p < .001$). No significant difference was found between B1 and A1 learners ($p = 1.000$), as both groups showed identical (zero) performance. Effect size calculations revealed perfect effects between the native speaker and both learner groups ($r = 1.000$ for both comparisons), while no effect was observed between B1 and A1 learners ($r = .000$).

Statistically significant differences were also found for ð-elision among the three groups, $H(2) = 10.000$, $p = .007$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners ($p = .004$), as well as between the native speaker and A1 learners ($p = .004$). No significant difference was found between B1 and A1 learners ($p = 1.000$), as both groups showed identical (zero) performance. Effect size calculations revealed large effects between the native speaker and both learner groups ($r = .707$ for both comparisons), while no effect was observed between B1 and A1 learners ($r = .000$).

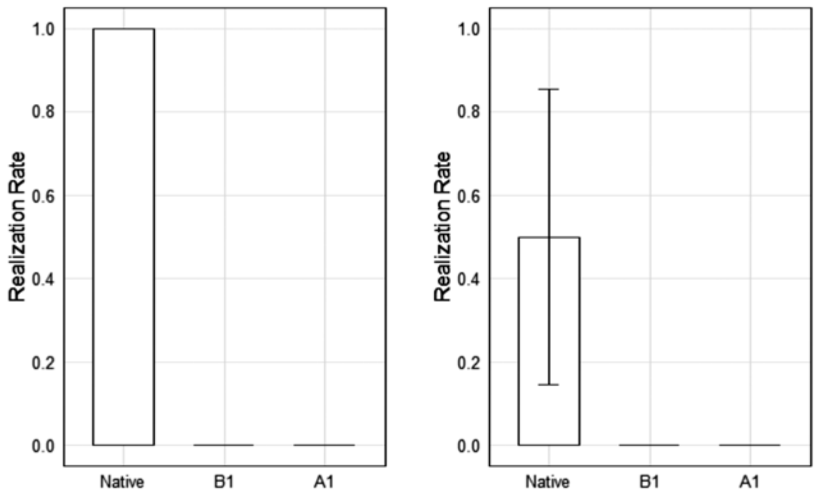

Figure 34. Elision rates of h-elision (left) and ð-elision (right)

e. T-Elision in /VntV/

Table 35 shows the number of t-elision occurrences of the three groups in /VntV/ environments. Similar to post-reduction elision, the native speaker exhibited three elisions out of four instances, while no participants from the B1 or A1 groups produced elisions in this environment.

Level	Participant	Elision count
	Native	3
B1	Japanese A	0
	Japanese I	0
	Japanese R	0
	Japanese S	0
	Japanese B	0
A1	Japanese O	0
	Japanese H	0
	Japanese K	0
	Japanese D	0
	Japanese N	0

Table 35. Number of occurrences of t-elision in /VntV/

Figure 36 shows the comparison of successful t-elisions in /VntV/ contexts, while Table 37 presents their basic statistics. The Kruskal-Wallis test revealed a highly significant difference between the groups, $H(2) = 20.00$, $p < .001$, with a large effect size ($r = .853$). Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed that the native speaker's performance (3 elisions) was significantly different from both B1 learners ($p < .001$, $r = .592$) and A1 learners ($p < .001$, $r = .592$), with large effect sizes in both comparisons. No significant difference was found between B1 and A1 learners ($p = 1.000$, $r = .000$) as neither group produced any elisions.

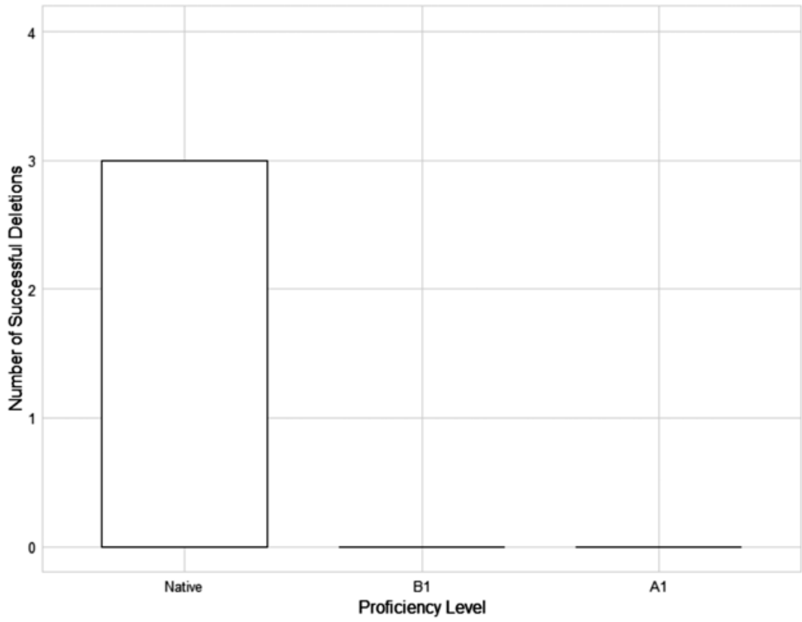

Figure 36. Comparison of successful t-elisions in /VntV/ contexts among the native speaker, B1 learners, and A1 learners

	Native	B1	A1
Mean	3	0	0
Median	3	0	0
Standard Deviation	-	-	-
Minimum	3	0	0
Maximum	3	0	0
Count	1	5	5

Table 37. Basic statistics of occurrences of t-elision in /VntV/

The occurrences of elision in contexts of /nter/ and /nty/ were examined using the Kruskal-Wallis test (Figure 38). For the /nter/ contexts, the results indicated statistically significant differences, $H(2) = 20.000$, $p < .001$. Effect sizes were calculated to assess the practical significance of group differences.Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners ($p < .001$), as well as between the native speaker and A1 learners ($p < .001$). No significant difference was found between the B1 and A1 learners ($p = 1.000$), as both groups showed identical (zero) performance. Effect size calculations revealed perfect effects between the native speaker and both learner groups ($r = 1.000$ for both comparisons), while no effect was observed between B1 and A1 learners ($r = .000$).

For the /nty/ context, the Kruskal-Wallis test also revealed statistically significant differences, $H(2) = 10.000$, $p = .007$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed significant differences between the native speaker and B1 learners ($p = .004$), as well as between the native speaker and A1 learners ($p = .004$). No significant difference was found between the B1 and A1 learners ($p = 1.000$), as both groups exhibited identical (zero) performance. Effect size calculations indicated large effects between the native speaker and both learner groups ($r = .707$ for both comparisons), while no effect was observed between B1 and A1 learners ($r = .000$).

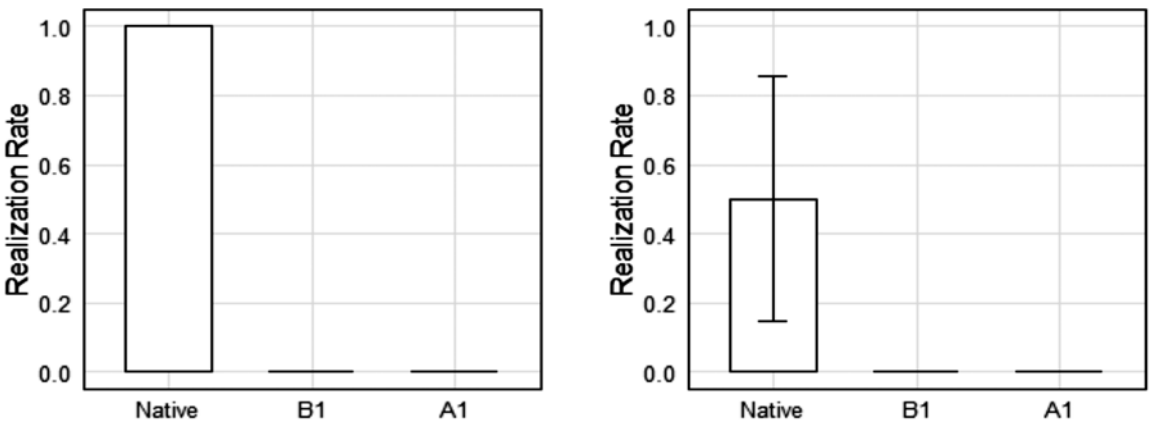

Figure 38. Elision rates in /nter/ (left) and /nty/ (right) contexts

Another examination was conducted to investigate differences in elision occurrences in [ɛnt] and [aʊnt] sequences (Figure 39). For the [ɛnt] sequence, the Kruskal-Wallis test revealed statistically significant differences, $H(2) = 10.000$, $p = .007$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons showed significant differences between the native speaker and B1 learners ($p = .004$), as well as between the native speaker and A1 learners ($p = .004$). No significant difference was found between B1 and A1 learners ($p = 1.000$), as both groups showed identical (zero) performance. Effect size calculations revealed large effects between the native speaker and both learner groups ($r = .707$ for both comparisons), while no effect between B1 and A1 learners ($r = .000$).

For the [aʊnt] sequence, the Kruskal-Wallis test also showed statistically significant differences, $H(2) = 20.000$, $p < .001$. Effect sizes were calculated to assess the practical significance of group differences. Follow-up Dunn's tests with Bonferroni correction for multiple comparisons revealed significant differences between the native speaker and B1 learners ($p < .001$), as well as between the native speaker and A1 learners ($p < .001$). No significant difference was found between B1 and A1 learners ($p = 1.000$), as both groups exhibited identical (zero) performance. Effect size calculations revealed perfect effects between the native speaker and both learner groups ($r = 1.000$ for both comparisons), while no effect was observed between B1 and A1 learners ($r = .000$).

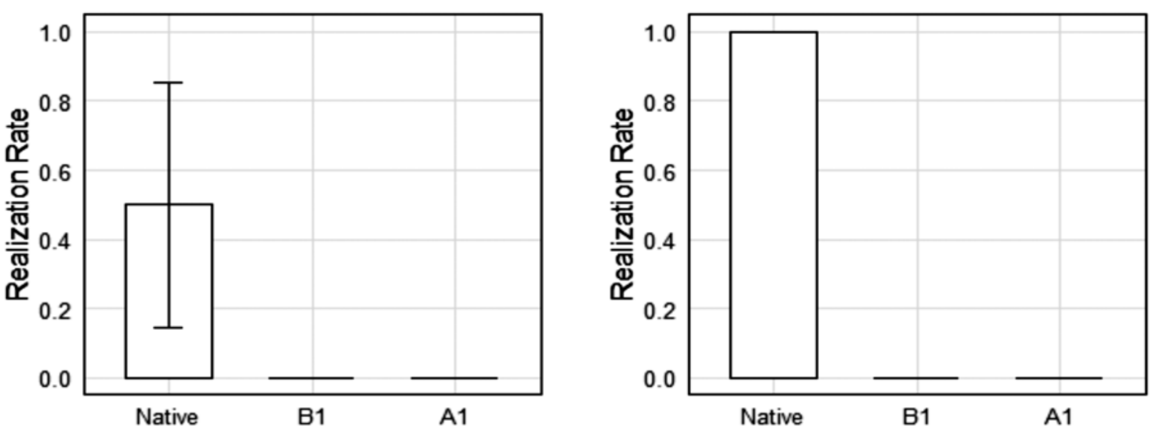

Figure 39. Elision rates in /nter/ (left) and /nty/ (right) contexts

The following is the summary of the results.

1. Overall, native speakers do not necessarily elide schwas, and the frequency of different types of elisions varies, with some types of elision occurring more frequently than others.

a. Schwa elision

2. There was a statistically significant difference in the overall realization of schwa elision between the native speaker, B1 learners, and A1 learners, with the B1 group performing significantly better than the A1 group.
3. Japanese students were more likely to elide schwa in three-syllable words than in two-syllable words at a practical level.
4. Japanese students showed similar performance in producing schwa elision in both voiceless consonant + schwa and voiced consonant + schwa contexts.

b. Preceding consonant elision

5. There were no statistically significant differences in the overall production of preceding consonant elision among the three groups, with the two Japanese student groups exhibiting identical performance.
6. Japanese students showed a similar pattern of preceding consonant elision in bilabial + bilabial and alveolar + alveolar contexts.
7. Japanese students produce preceding consonant elision more frequently in voiceless + voiceless contexts than in voiced + voiced contexts, with the B1 group showing a statistically significant difference and the A1 group showing a practical difference.

c. Preceding plosive consonant elision

8. There was a statistically significant difference in the overall realization of preceding plosive consonant elision across the three groups.
9. Japanese students performed plosive consonant elision better in voiced + voiceless contexts than in voiceless + voiceless contexts.
10. Japanese students produced elision more frequently in plosive + fricative contexts than in plosive + plosive contexts.

d. Post-reduction elision

11. There was a significant difference in the overall occurrences of post-reduction elision between the native speaker and the two Japanese groups.
12. Japanese students from both proficiency levels exhibited no post-reduction elision, regardless of the consonants expected to elide.

e. T-elision in /VntV/

13. There was a significant difference in the overall production of t-elision in /VntV/environment between the native speaker and the two Japanese groups.
14. Japanese students from both proficiency levels did not exhibit t-elision in /VntV/environment, regardless of the context in which t-elision occurs or the vowels preceding [nt].

5. Discussion

The findings of this study have several important implications for understanding elision patterns in English and their pedagogical applications. Result 1 provides valuable insights into sound change research. While various types of elisions are documented in the existing literature, native speakers are likely to produce different types of sound changes with varying frequencies. This suggests that English learners could benefit from focusing on the elisions most frequently produced by native speakers, leading to more effective and efficient learning.

Results 2, 5, 8, 11, and 13 indicate that Japanese students have difficulty producing certain types of elisions, while they can produce other types of elision as frequently as the native speaker. For example, Japanese students produce schwa elision, preceding plosive consonant elision, post-reduction elision, and t-elision in /VntV/ much less frequently than the native speaker. On the other hand, they can perform preceding consonant elision at a similar level to that of the native speaker. Students should particularly focus on practicing post-reduction elision and t-elision in /VntV/ context, as indicated by Results 12 and 14, which highlight a gap in the production of elision in these contexts. Although t-elision in /VntV/ is a feature commonly found in American and Canadian English, most Japanese English learners are taught American English in public education, so it is one of the types of sound change they can acquire. The significant gaps in the realization of post-reduction elision and t-elision in /VntV/ between the native speaker and the Japanese participants suggest that the students may not be fully aware of the sound changes expected to happen during speech production, whereas native speakers exhibit a high frequency of these elisions.

This paper explored not only the differences in elision types but also those based on more detailed contexts. In schwa elision, Japanese students find it easier to elide schwa in three-syllable words compared to two-syllable words (Results 3). For preceding consonant elision, they are more likely to realize elision in voiceless + voiceless contexts than in voiced + voiced contexts (Result 7). Similarly, for preceding plosive consonant elision, Japanese students elide the preceding consonants more frequently in voiced + voiceless contexts than in voiceless + voiceless contexts (Result 9), and more frequently in plosive + fricative contexts than in plosive + plosive contexts (Result 10). These findings suggest that teachers should focus on the more challenging contexts where elision occurs when teaching their students the production and perception of elision.

To address the production gap between native speakers and Japanese students, focusing on specific elisions and practicing them could enhance students' natural speech flow and their perception of spoken English. One effective approach is to have students explicitly notice these elisions. By marking elisions on listening scripts and carefully analyzing the audio, students can consciously process elisions they may have previously overlooked. They can then practice reproducing these elisions by imitating the audio. This training helps students internalize the sound changes, leading to improvements in both their speech production and listening comprehension skills.

One limitation of this study is that it involves only one native English speaker. Including data from more native speakers would lead to more accurate and convincing findings. Additionally, the present study focused solely on elision within limited contexts. Future research should explore assimilation and linking, as well as other contexts of elision.

6. Conclusion

The present study compared elision production between a native English speaker and Japanese English learners with differing proficiency levels through a comparative analysis. Two key findings emerged from this investigation. First, it was found that native speakers do not consistently produce the sound changes documented in the literature. Additionally, it was observed that schwa elision, preceding plosive consonant elision, and, particularly, post-reduction elision and t-elision in /VntV/ were challenging for Japanese students to produce.

This study further identified specific contexts in which Japanese students experience difficulty with elisions. For instance, they are more likely to elide schwa in three-syllable words than in two-syllable words. They also elide preceding consonants more often in voiceless + voiceless contexts than in voiced + voiced contexts. Similarly, preceding plosive consonants are elided more often in voiced + voiceless contexts than in voiceless + voiceless contexts, and more often in plosive + fricative contexts than in plosive + plosive contexts.

Japanese students did not exhibit post-reduction elision or t-elision in /VntV/ contexts, indicating a lack of awareness of the elisions expected in these situations. These findings suggest that varying degrees of difficulty exist in producing elisions among Japanese English learners, emphasizing the need for targeted training. Teachers should prioritize explicit instruction and practice of problematic elision patterns, particularly those that show the largest gaps between native and learner production. Such focused instruction could enhance both the naturalness of learners' spoken English and their ability to process connected speech during real-time listening comprehension.

References

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Baghrahi, A. K., Shariati, M., & Tajadini, M. (2014). The effect of assimilation and elision teaching on listening comprehension of EFL junior high school students. *International Journal of Language Learning and Applied Linguistics World*, *5*(1), 264-273.

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. *Proceedings of the International Congress of Phonetic Sciences*. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1041.pdf

Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn's Test. *The Stata Journal*, *15*(1), 292-300. https://doi.org/10.1177/1536867X1501500117

Hasan, A. (2012). A phonological study of elision in standard English and standard Arabic. *Journal of University of Babylon*, *20*(2), 454-473.

Hecke, T. V. (2012). Power study of ANOVA versus Kruskal-Wallis test. *Journal of Statistics and Management Systems*, *15*(2-3), 241-247. https://doi.org/10.1080/09720510.2012.10701623

Katayama, Y., Nagase, Y., & Joto, A. (1996). *Foundations of English phonetics: Focusing on sound changes and prosody*. Kenkyusha.

Liu, H. (2015). *Comparing Welch's ANOVA, a Kruskal-Wallis test and traditional ANOVA in case of heterogeneity of variance* [Master's thesis, Virginia Commonwealth University]. *VCU Scholars Compass*. https://scholarscompass.vcu.edu/etd/3985

Nakamura, H. (1997). A study on the effect of learning sound changes in movie English on EFL learners' listening abilities. *The Association for Teaching English through Multimedia*, *3*, 39-49. https://doi.org/10.24499/atem.3.0_39

Ueda, H., & Otsuka, T. (2014). An analysis of pronunciation instruction items in Japanese junior high school English textbooks: The changes under the new course of study. *Journal of Osaka Jogakuin University*, *10*, 1-15.

Yamagami, H. (2023). Reconsidering the classification of sound change. *Seinan Gakuin University Graduate School Bulletin*, *16*, 125-140.

Young Jr., J. A. (2015). *Exploring Japanese learners' perception, production, and beliefs concerning spoken English contractions* [Doctoral Dissertations, University of Illinois at Urbana-Champaign]. ProQuest Dissertations and Theses. https://www.proquest.com/docview/1816974360