Identification of Mora in Sokuon, Tan'on, and Chouon by Japanese language learners in Indonesia

Rike Febriyanti*, Camelia Vonita Febri, Ary Sudiargo
Universitas Brawijaya, Indonesia
*Corresponding Author; Email: febriyanti_rike@ub.ac.id

ABSTRACT

Japanese is regarded as *mora*-timed, while Indonesian is syllable-timed because it provides equal weight and time for all syllables. Indonesian learners are struggling in differentiating one *mora* and two *mora*s, and single consonant and double consonants. This study investigated Indonesian respondents' ability to identify double consonants (*sokuon*), short vowels (*tan'on*) dan long vowels (*chouon*). 62 Indonesian respondents were involved as they were asked to listen carefully to a set of sounds containing the three distinctive sounds once and to write the sentence they heard. The result showed that 18% of 62 respondents failed to identify the short vowel (*tan'on*) because they omitted certain sounds, changed the word form inaccurately, and put words in an inaccurate position when they wrote the sentence. In identifying double consonants (*sokuon*), 11% of 62 respondents failed because they inaccurately changed the word form when writing the sentence. In identifying long vowels (*chouon*), only 6% of 62 respondents failed because they added unnecessary sounds and changed the word form inaccurately. In conclusion, although Japanese and Indonesian languages are different in their rhythm classes, Indonesian learners' ability to identify the distinctive sounds of the Japanese language is relatively high although listening to the material more than once.

Keywords: Short vowel, *Mora*, double consonant, long vowel

INTRODUCTION

Most assumptions about non-native speech perception and production view the L1 (first language) phonetic inventory as a 'filter' in L2 (second language) phonetic learning. At their core are the 'Native Language Magnet Model', the 'Perceptual Assimilation Model' (Rosen et al., 2015) and the 'Speech Learning Model' (Flege & Bohn, 2021). The idea is that the perceptual space and production of L2 segments can be altered by the L1 phonetic categories of the learner, so learning outcomes are not always as expected. The Speech Learning Model (SLM) and Perceptual Assimilation Model-L2 (PAM-L2) (Tyler, 2019), link non-native segment learning to L1 phonetic inventory by incorporating L2 experience, L2 production and perception, L2 phonetic categories, which are important for learning outcomes.

SLM theory seeks to explain how L2 can form new L2 phonetic categories by considering the perceived distance between L2 and L1. New phonetic categories for L2 sounds when the learner discovers (or sees) phonetic dissimilarities between the L2 sound and the closest L1 sound in phonetic space, and the learner will be more successful in learning 'new' L2 segments compared to 'similar' L2 segments. This research is based on the idea that the greater the perceived dissimilarity of the L2 segment from the closest L1, the more likely a new category will be formed for the L2 sound.

It has been widely reported that native speakers have a harder time understanding non-native speech than native speakers (Liu & Takeda 2021). According to a large body of evidence, proficiency of speakers' second language (L2) affects intelligibility as stated by (Mitterer et al. 2020). On the other hand, foreign-accented speech does not make all listeners less comprehensible. The shared language background of the speaker and listener, as well as individual language experience, influence the

intelligibility of foreign-accented speech. (Yazawa et al., 2023) created the term "interlanguage speech intelligibility benefit" (ISIB) to characterize a situation in which L2 learners can distinguish foreign-accented speech as accurately as native speech. Their use of the term "interlanguage" refers to the possibility that L2 speakers and listeners can benefit from each other's language.

PAM-L2 seeks to analyze perceived L1-L2 dissimilarity at the phonetic and phonological levels and predict success in L2 perceptual acquisition across different scenarios: (1) when only one L2 phonological category is perceptually assimilated to a given L1 phonological category, (2) when both L2 phonological categories are perceptually assimilated to the same L1, with equally good or unequally good examples of those categories, (3) when there is no L1-L2 phonological assimilation.

SLM predicts that small L1-L2 dissimilarities perceived by L2 learners may result in failure of L2 category formation: Specifically, L2 learners may be able to produce target segments with strong characteristics similar to nearby L1 counterparts. Given the foregoing, this study aims to investigate how learners with different L1 backgrounds can produce their learning outcomes for L2 segments.

The Indonesian language (Bahasa Indonesia), is a unity language formed from hundreds of languages spoken in the Indonesian archipelago as specified by (Rahmi, 2016). The official language of the Republic of Indonesia is Indonesian, which is spoken by almost 250 million people in 37 provinces. Based on the 2018 survey on Japanese language education abroad conducted by the Japan Foundation (Djafri & Wahidati 2020), the total number of Japanese learners in the world currently has reached 3,851,774 people. China ranks first with 1,004,625 learners, followed by Indonesia with 709,479 learners. The number of Japanese learners in Indonesia is massive although Indonesian and Japanese languages are different in their rhythm as classified by (Liu & Takeda, 2021b).

While Indonesian is a syllable-timed language because it gives equal weight and time for all syllables (Lasut 2015), Japanese is categorized into *mora*-timed language, which differentiates bi*mora*ic syllables and mono*mora*ic syllables as mentioned by (Liu & Takeda 2021). (Garellek 2019) proposes *mora*-timed languages, which are mainly exemplified by Japanese and suggests that "[a] *mora* is a unit of timing, in the sense that each *mora* … ha[s] approximately the same duration." Meanwhile, (Tateishi, 2017) claims that the *mora* is "the unit of phonological distance" in Japanese and can only be defined as "something of which a long syllable consists of two and a short syllable consists of one."

In Japanese, there are several unique sounds, such as short vowels (*tan'on*), double consonants (*sokuon*), and long vowels (*chouon*), which distinguish meanings in (Shiro, n.d.). Syllables in Japanese are called onsetsu 「音節」 (Ota et al., 2018), which are identical to syllables in Indonesian or can be said to be the smallest unit of a word according to (Prastiti et al. 2015). In contrast to Indonesian, in Japanese, there is a *mora* / Haku which is the smallest unit of a word, for mention in Indonesian it is called a beat as stated by (Fujimoto et al. 2021). Furthermore, (Febriyanti & Indrowati 2015) stated that Indonesian learners are struggling in differentiating one *mora* and two *mora*s, and single consonant and double consonants. Then the previous notion is supported by (Pimentel et al. 2020) that the L2 learner tends to transfer the habits of his native language structure to the foreign language.

One of the reasons for the difficulties identifying sounds or pronouncing unique Japanese sounds is that if they hear the wrong words or sentences, the production produced by the learner's hearing will also be wrong. So that it will affect the understanding of meaning because in Japanese, it is known as a language rich in letters but lacks sound variety. (Baydar 2020) and (Kubozono 2017) states that the acquisition of rhythm in Japanese is about the types of sounds, such as short vowel (*tan'on*), double consonants (*sokuon*), and long vowels (*chouon*). Studies such by (Ofuka, 1997; Pendidikan et al., 2021; Wahyuni & Sutedi, 2020) on the perception as well as the production of suprasegmental sounds in Japanese has been conducted. The results of the research indicate that learning Japanese sounds is challenging for most non-mora languages. Besides, those studies characterize learners' capacity to recognize patterns in Japanese accents, this study concentrated on how learners perceive Japanese accents but none of them touched the unique Japanese sounds like this study. This study emphasizes to find how well Japanese language learners in Indonesia who speak their mother tongue as a syllable-timed language can distinguish between short and long *mora* in Japanese as represented by short vowel (*tan'on*), double consonants (*sokuon*), and long vowels (*chouon*).

Stress, intonation, and rhythm are prosodic elements that have been discovered to contribute more to understandability and to a lesser extent to foreign accentedness than segmental features, but there hasn't been much research on them (Dahmen et al., 2023). Furthermore, (Rasier & Hiligsmann, 2007) stated that transfer at the suprasegmental level is more crucial since it is "cumulative," whereas transfer at the segmental level is "self-limiting." Transfer can sometimes be beneficial ('positive transfer'). For instance, if two languages share some characteristics, being able to draw from one's prior knowledge of the first language helps speed up learning the second. The widespread use of knowledge from the other language in different contexts may be harmful (referred to as "negative transfer"). For instance, if two languages differ in a certain way, then a transfer from one would probably result in an unlawful structure in the other. This might be especially typical if two languages are only subtly different in the finer features and superficially identical as stated by by (Romero & Manjarres, 2017).

METHOD

The research employs a descriptive qualitative method as suggested by (Rose et al., 2019; Yazawa et al., 2023) to investigate the respondents' ability to identify unique sounds in Japanese language. The present research employed a listening method as the basic technique and descriptive approach for data collection. The participants in this study were 62 respondents from Brawijaya University's Japanese Language Education program who had completed their third semester using *Minna No Nihongo Sokyu* chapter 14 page 116 and chapter 16 page 136, which discussesed double consonants (*sokuon*) in Japanese and contained about 47 vocabularies. Respondents was also completed questions via online surveys and inquiries due to the COVID-19 pandemic.

Respondents were first invited to listen to the audio, and then they were asked to identify responses from the audio they had listened to. They were asked to respond in accordance with the question, and so on until the last question is asked. Second, if any individuals made mistakes throughout the identification process, the research participants would be questioned to determine the reasons that lead to identification errors. Participants would listen to audio recordings of native Japanese speakers reciting various sentences in Japanese in response to the specified sentences, with the voice structured as naturally as possible as follows (Saji et al., 2019):

1. 来てください。　　（kite kudasai）
2. 着てください。　　（kite kudasai）
3. 切ってください。　（kitte kudasai）
4. 切手ください。　　（kitte kudasai）
5. 聞いてください。　（kiite kudasai）

The research participants were then requested to respond by continuing the current sentences in the offered column by writing the sentences they had heard. Hiragana letters must be used to fill in the blanks. The following were the sentences that the participants need to study:

1. 明日戻って来てください。（ashita modotte kite kudasai）
2. この着物を着てください。(kono kimono wo kite kudasai)
3. この野菜を切ってください。(kono yasai wo kitte kudasai)
4. この手紙の切手をください。(kono tegami no kitte wo kudasai)
5. この録音を聞いてください。(kono rokuon wo kiite kudasai)

The possible errors made by the participants will be classified using the categories proposed by (Febriyanti & Husna, 2020), are:

"(1) deduction (omission), important elements which are left unwritten or unspoken, (2) addition, unimportant parts which are added without significance effect, and (3) morphological mistakes (misformation), errors that modify the shape/format of words, (4) errors in placement (misordering), an error in the arrangement of an item in a phrase."

To find out the research subject's understanding of the sound recording of native speakers, the author presented a set of data in the form of a resource identification table with 5 sections which were divided into 3 types of distinctive sounds in Japanese pronounced by native speakers, regarding short sounds (*tan'on*) in 「来てください」 and 「着てください」, double consonant (*sokuon*) in 「切って ください」 and 「切手をください」, and long sound (*chouon*) in 「聞いてください」.

In the self-paced condition, participants were given control over an embedded media player after viewing the items in their original version (single listening). As a result, each participant was free to decide if they wanted to listen to the full sound file or only certain parts. The participants in this study could listen directly from the beginning to the end, stop whenever they wanted, or even replay the recording from the start. What they couldn't do was listen to the next recording or read the next question until they had finished answering the current one in front of them. The question was visible during the hearing in the typical single-listening setting. This was done so that, in the absence of any potential effects from longer reading times for the items, we could examine any additional benefits that self-paced listening might have. In the self-paced listening condition, participants were free to take as much time as necessary. Every question received a score of 0 if the participant answered incorrectly, and 1 point was given for accurate responses.

As proposed by Petermann & Daseking (2019) in (Kochva et al., 2021), two tasks—a word reading task and a phrase reading task—were utilized to assess the reading accuracy and speed of written text decoding. There were five sets in the word reading task. As previously noted, there were five minimal pairs in phrases in the phrases reading task. Participants in both tasks were instructed to read the sentences or words aloud precisely and clearly, and researchers kept track of the total amount of time needed for each set as well as the number of mistakes made. We invited participants to complete a series of questions that assessed their ability to distinguish between the minimum pairings we suggested in the five forms of sentences we previously described, following their listening to a series of sound files. After collecting recordings from all participants, several follow-up interviews were conducted in this study, particularly with participants who exhibited unique phonological phenomena. This was done to gain a deeper understanding of the respondents' conditions when pronouncing and why they did so. In addition to recording right and wrong responses and conducting in-depth interviews to learn more about the root causes of errors, we also documented unusual events that might occur when participants heard our questions and responded.

RESULTS AND DISCUSSION
Results

For the first question regarding the type of sound (*tan'on*) 「来てください」, the percentage findings acquired roughly 82 percent of accurate responses and 18 percent of inaccurate responses from respondents who responded the question. For the second question regarding (*tan'on*) 「着てください 」, respondents responded the question with the percentage results obtained about 95% of correct responses and 5% of inaccurate responses. For the third question regarding (*sokuon*) 「切ってくださ い」, respondents responded the question with a percentage of 89% correct responses and 11% inaccurate responses. For the fourth response regarding (*sokuon*) 「切手をください」, respondents

responded the question with a percentage of 68% correct responses and 32% inaccurate responses. To response the fifth question regarding (*chouon*) 「聞いてください」, respondents responded questions with a percentage of 94% correct responses and 6% inaccurate responses. The above overall results can be illustrated in Figure 1.
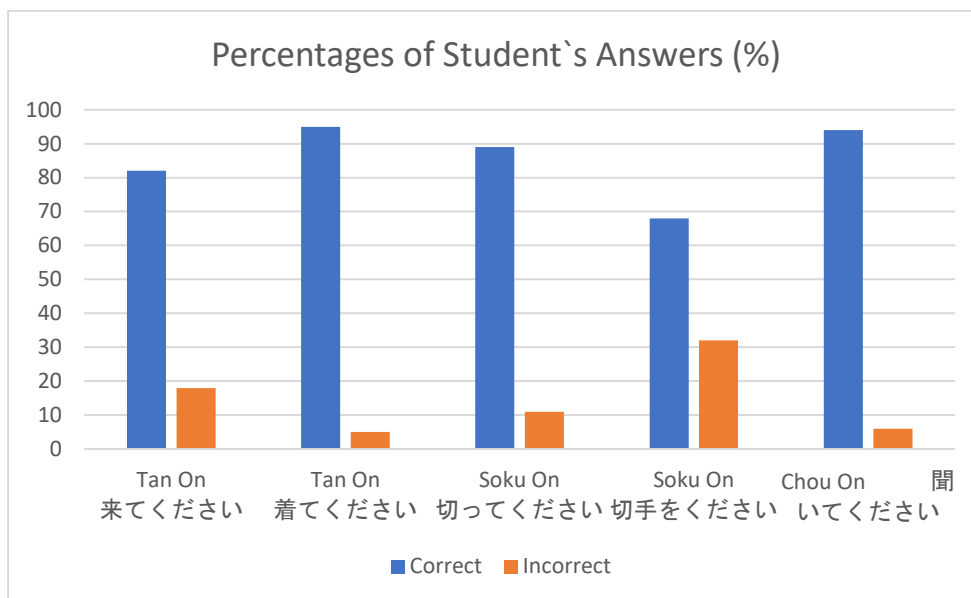

Figure 1. Student`s Responses

From the Figure 1 above we could understand five primary initial data. First, the majority of responders can recognize short sounds (*tan'on*) 「来てください」 in the sentence number 1 「明日戻って来てください」, with 82 percent of accurate responses and 18 percent of inaccurate responses - 4 respondents replied inaccurately due to an omission by writing only 「ください」 (kudasai), 3 respondents inaccurately answered due to a change in word form (misformation) by writing 「きってください」 (kitte kudasai), and 4 respondents inaccurately answered due to a word placement error (misordering) by writing 「いってください」 (itte kudasai)（3 respondents）and 「どってください」 (dotte kudasai) (1 student). The illustration of respondents' ability to identify *tan'on* 「来てください」 will be illustrated in Figure 2.
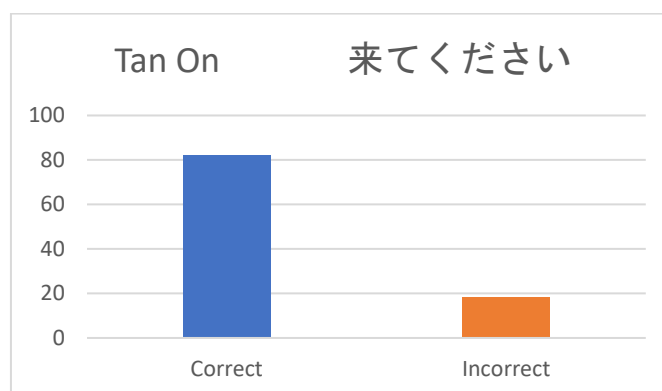

Figure 2. Respondents' Identification in *Tan'on* in Text 1

Second, the majority of responders could recognize short sounds (*tan'on*)「着てください」in the sentence number 2 「この着物を着てください」, with 95 percent of accurate responses and 5 percent of inaccurate responses. The ability of the respondents to identify short sounds (*tan'on*) in sentence number 2 pronounced by native speakers is high. 59 respondents responded correctly, and 3 respondents responded inaccurately. 3 respondents that responded inaccurately due to an error in modifying the word's form (misformation). 1 respondent wrote「きいてください」（kiite kudasai）and 2 respondents wrote「きってください」(kitte kudasai). The illustration of this situation is shown in Figure 3.
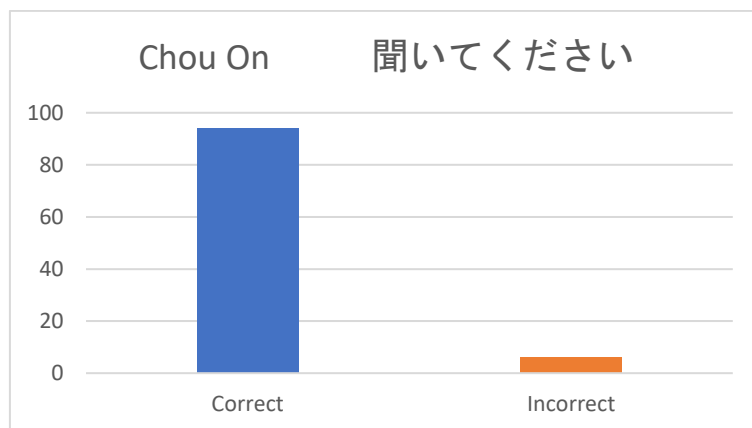

Figure 3. Respondents' Identification in *Tan'on* in text 2

Third, almost all respondents can recognize the sound of double consonants (*sokuon*) in「切ってください」in the sentence number 3 「この野菜を切ってください。」, resulting in an overall proportion of 89 percent accurate responses and 11% inaccurate responses. In detail, all participants made mistake in changing the form word (misformation). The ability of respondents to identify double consonants (*sokuon*) in sentence number 3 (*kono yasai wo kitte kudasai*) pronounced by native speakers was relatively in the middle stage because respondents could response correctly (55 respondents), and for those who could not identify double consonants (*sokuon*), 5 respondents showed inaccurate responses by writing 「きいてください」(kiite kudasai) meanwhile 2 respondents showed by writing

「きてください」(kite kudasai). The following condition is described in Figure 4.
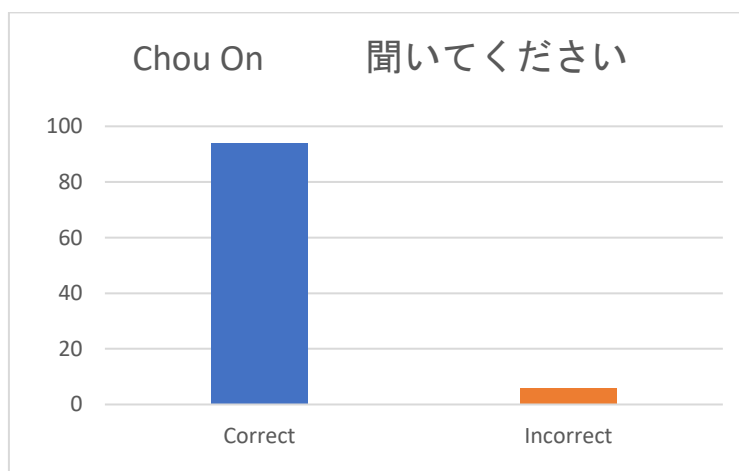

Figure 4. Respondents' Identification in *Sokuon* in Text 3

Fourth, more than half of the respondents correctly identified the sound of double consonants (*sokuon*) in 「切手をください」, with 68 percent of the responses was accurate and 32 percent (20 respondents) was inaccurate. 9 respondents who responded inaccurately for doing subtraction (omission) such as 「きってください」 (kitte kudasai). Then there were 9 respondents who responded wrong by changing word form (misformation) such as 「きいてください」 (kiie kudasai) and there were 2 respondents who responded wrong due to misordering by writing 「かいてください」 (kaite kudasai). The following Figure 5 shows the state.
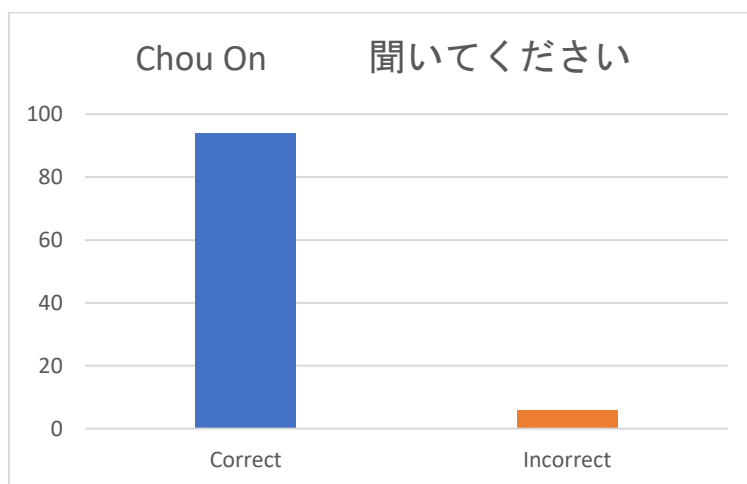

Figure 5. Respondents' Identification in *Sokuon* in text 2

Fifth, almost all participants could recognize long vowel (*chouon*) in word 「聞いてください」 in sentence 「この録音を聞いてください。」, resulting in a 94 percent accurate response rate and a 6 percent inaccurate response rate. In this segment, we could find that 1 respondent who responded inaccurately because of unnecessary addition such as 「きいてくあださい」 (kiite kuadasai) and then there were 3 respondents who responded inaccurately because of misformation such as 「きってください」 (kitte kudasai). This respondents' achievement will be portrayed in the following Figure 6.
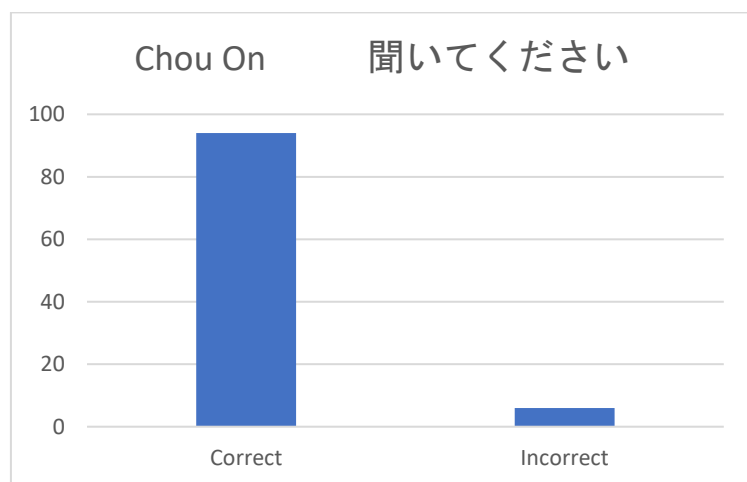

Figure 6. Respondents' Identification in *Chouon* in Text 2

Discussion

The first issue I must bring up in this discussion is how many times the respondents listened to the audio file containing the words they were required to answer. Despite our instructions for the respondents to just listen to the recording file once, they listened it multiple times. The respondents said they listened to the recorded file an average of twice before responding to the questions for the first and second sentences. That implies that they did so twice—once while following directions and once on their own. The respondents said they heard an average of three times for the third and fifth questions, thrice as directed and three times on their own initiative. The final one, notably sentence number 4, was repeated five times. Once as directed and four times on their own initiative. This demonstrates the study's multiple dimensions. Firstly, even responders who had taken multiple semesters of Japanese classes lacked the confidence to listen to the recording just once. Due to their limited exposure to native speakers' stories, this is quite likely. They are also highly likely to be insecure since, up until now, insufficient emphasis has been placed on the distinctive sounds of Japanese in a straightforward and organized way, as (Fujita, 2022) explains.

Because of the COVID-19 epidemic during this research and in order to determine the error factor, the author conducted an interview stage by asking respondents through a private room on WhatsApp when the respondent was unable to identify the type of short sound (*tan'on*), the sound of double consonants (*sokuon*), and long sounds (*chouon*) that have been pronounced by native speakers. The majority of the responses given in response to the first question by respondents were accurate. Based on the findings above, it can be said that listening to sentences that have been uttered by native speakers tends to make respondents' responses to interview question number 1 more accurate. They felt a little bit confident in their response to interview question 1 because they believed the responses they have heard were accurate. This fact shows that during the pandemic, the online Japanese language teaching and learning process was still slightly promote a student-centered learning environment and two-way interaction via computer networks or the Internet without going through a server in a classroom setting for language learning as stated by (Winarni et al., 2022).

Secondly, the writer conducts the interview stage by instructing the respondent to re-listen to the recording of the inaccurate question by responding the question again, done by asking respondents through private rooms on WhatsApp, when the respondent is unable to recognize the type of short sound (*tan'on*), the sound of double consonants (*sokuon*), and long sounds (*chouon*) that have been pronounced by native speakers. This similar condition has been experienced by (Wahyuni & Sutedi, 2020) who pointed out that condition was because of the respondents' lack of ability in identifying hiragana, katana or kanji. When listening to sentences pronounced by native speakers, respondents who provided an average response to the second question tended to repeat the words 1 to 3 times. According to this analysis, it can be inferred that because the respondents' responses were insufficient, the respondents took their own initiative that they listen to the recording repeatedly until they could respond to the questions correctly. This was why the responses to interview question number 2 from the respondents tended to repeat one to three times throughout the interview listen to how natural speakers pronounce sentences.

The third point, when a respondent is unable to distinguish between the types of short sounds (*tan'on*), double consonant sounds (*sokuon*), and long sounds (*chouon*) as pronounced by native speakers, the writer conducts an interview stage by asking respondents in private Whatsapp rooms in order to determine the error factor. The majority of respondents who gave an average response to the third question were aware of the type of double consonant sound (*sokuon*), but they were less knowledgeable about the specifics of the double consonant sound (*sokuon*). Based on this data, it can be deduced that respondents generally tend to know the general type of double consonant sound (*sokuon*), but not the specifics, because they believe that native speakers pronounce words and sentences correctly. The similarity of the average causes respondents to pause before responding to questions. Respondents are therefore inaccurate in their responses and unable to recognize the different sound kinds, particularly the double consonant sounds (*sokuon*s) present in phrases produced by native speakers. The similar problem was faced by (Ren, 2022) while she was investigating Chinese students who learnt to pronounce Japanese geminate construents. She considered that while Chinese respondents learning Japanese geminate construents such as *sokuon* and tan'on having much

lower perception levels than JNS notwithstanding group differences (such as L1 background and Japanese language competency), they nevertheless made use of their L1 to improve their perception performance.

The fourth point, although respondents who gave an average response to the fourth question tended to understand the sound of double consonants (*sokuon*), they did not truly comprehend it in depth because during lecture learning, the teacher only provided explanations based on the material in the textbook and did not go into detail about the understanding of multiple sounds (*sokuon*). This issue has been also pointed by (Seguel & Sonesson, 2022) who investigated the similar problematic issue in teaching Japanese in Sweden where the tutors did not provide enough time to master unique sounds in Japanese such as geminate construents so the Swedish only learning them at a glance. This investigation leads to the conclusion that while respondents generally grasp the sound of double consonants (*sokuon*), they do not necessarily comprehend it thoroughly because they believe that there is no distinct explanation of the sound (*sokuon*) and that the explanation of double consonants (*sokuon*s) in learning at lectures is exclusively limited to textbooks. Respondents are therefore misinformed in their responses and fail to recognize the different sound kinds, particularly the double consonant sounds (*sokuon*s) present in sentences produced by native speakers.

The fifth point, when given information in the form of text, where it can be read repeatedly and understood easily, respondents were more likely to be able to identify sounds (*sokuon*), whereas when given information in the form of pronunciation, where there might not be enough time to complete the listening process (choukai), respondents were less likely to be able to understand and to identify duplicate sounds (*sokuon*). so that there was uncertainty when the respondent provides a response. Based on the results of this investigation, it could be said that because written solutions could be easily read repeatedly, they were typically simpler to comprehend and identify (*sokuon*) than pronunciation. When there was a (*sokuon*) in the pronunciation, the respondents found it difficult to respond the question. As a result, the respondent misidentified some sound categories, particularly some double consonant sound categories (*sokuon*) used by native speakers.

The sixth question was answered by respondents with an average response that tended to be less difficult if they understood the double sound (*sokuon*) in text form, but the respondent felt difficult if it was in the form of pronunciation because there were respondents who had only recently learned Japanese in college and were less familiar with the dual sound (*sokuon*). Then, for those who have studied Japanese for a considerable amount of time, namely since junior high or high school, they still find it challenging because the teacher does not provide a grasp of the different sorts of Japanese pronunciation sounds. In general, respondents' responses are simpler to comprehend and recognize (*sokuon*) in text form since the pronunciation is simple to read. Because some respondents have only recently acquired Japanese in lectures, they are less familiar with the different sorts of sounds, particularly the double sound type (*sokuon*). Then, for those who have been exposed to Japanese language for a long time, particularly since junior high or high school, they still find it challenging since the teacher does not provide clear instruction on the different forms of Japanese pronunciation sounds. As a result, respondents misidentified some sound categories, particularly those including double consonants (*sokuon*s) in phrases spoken by native speakers.

The last point, the respondents who provided an average response to the seventh question frequently lacked independent listening practice to native speakers pronouncing sentences or words, and when they did, they primarily relied on choukai learning during lectures. As a result, the respondents' hearing was unaccustomed to understanding and identifying different sound types, particularly Japanese double consonant sounds (*sokuon*). The respondents become confused and uncertain when identifying duplicate sounds as a result of their lack of knowledge of the vocabulary that they have learned and understood since they believe that words produced by native speakers are all essentially the same and similar. This investigation has led to the conclusion that for respondents to understand the meaning of the sentence, they must listen to it several times because, on average, they haven't had much practice listening to sentences or words said by native speakers on their own. When asked to identify the different sorts of sounds, particularly the different double consonant sounds (*sokuon*s) in phrases produced by

native speakers, respondents become hesitant and make mistakes due to a lack of terminology that is memorized and understood.

CONCLUSION

The first conclusion was that 83.6% of the 310 responses provided by the respondents were accurate, while 13.8% were incorrect. All of these responses were given after the respondents had heard between two and five times, greatly exceeding the researchers' instructions. We assume that the percentage of accurate responses would not have been as high and the percentage of incorrect answers would have been significantly greater if the respondents had listened only once because having control over the hearing input may enhance the measurement precision of listening assessments in the first attempt. Self-pacing also helps test-takers or, more generally, with their second language (L2) comprehension skills. in particular, of students who struggle with reading and may process information more slowly, with swiftness and little working memory.

The second inference is that, among the five statements we requested responders to hear, *Sokuon* 「切手ください」 in sentence number four 「この手紙の切手をください。」 (kono tegami no kitte wo kudasai) is the most challenging. This inference is made because the respondents dared to defy the researchers' instructions by repeating the hearing five times in sentence number four. This was further supported by the study's lowest percentage of accurate responses (69.2%) and largest number of incorrect responses (30.8%). Maybe if the participants only heard the recorded file once, the already low percentage of accurate responses would be considerably lower and the number of participants who gave incorrect answers much higher. In this instance, we have not been able to demonstrate why the respondents find the *sokuon* in phrase number four to be the most challenging distinctive sound form. Of course, this calls for more thorough investigation utilizing a wider range of examples of sentences.

The second less difficult item was the *tan'on* 「来てください」 in the first sentence 「明日戻って来てください。」 （ashita modotte kite kudasai) which were responded accurately by 78% of respondents and the rest were responded inaccurately. Again, this number was the subject of doubt since the respondents were confessing that they were listening twice.

The research shows that the respondent's ability to detect double consonant sounds (*sokuon*) is first assessed at the hearing and comprehending stages. Because respondents did not independently practice listening to phrases or words uttered by native speakers, they relied only on listening courses (*choukai)* learned during lectures when practicing listening. As a result, respondents are unfamiliar with recognizing and distinguishing different sorts of sounds, particularly double consonants (*sokuon*) in Japanese. Next, respondents' ability to identify double consonant sounds (*sokuon*) is influenced by their lack of memorized and understood vocabulary, which causes respondents to be confused and doubtful when identifying duplicate sounds because they assume that native speakers' words are almost all the same and similar. Furthermore, respondents find it simpler to recognize double consonant sounds (*sokuon)* of respondents studying Japanese at Brawijaya University by reading texts rather than by pronouncing them. Because they believe that if they go over the reading material with the responder, the respondent will be able to read it again. However, if when the respondents perform identification by pronunciation, it is different since there is a limited amount of time available, and responders cannot do it often enough to learn native speakers' pronunciation.

Finally, this research demonstrates that Japanese respondents in Indonesia are quite adaptable while transitioning from identifying vowels and single consonants to long double vowels and double consonants. We believe that the respondents to this study still lack the confidence necessary to follow our directions and listen to unique Japanese noises just once. The data we offer still need to be taken into consideration because it turns out that we are powerless to force our respondents to follow our instructions. We believe we will have time to conduct research where we can control our respondents after the COVID-19 epidemic has passed, resulting in more legitimate and accurate research findings. We also aim to be able to improve the environment by adopting sound system technology that is better equipped to deliver respondents with equal-quality sound. The previous notions can contribute to the field of phonology and Japanese language education in Indonesia in a

few ways, despite limitations acknowledged by the researchers. This study highlights the challenges Indonesian learners face with Sokuon (double consonants) compared to Tan'on (shortened vowels). This provides data on specific sound features that learners might struggle with. The study suggests that repeated listening improves accuracy in comprehension tasks. This emphasizes the importance of providing learners with multiple exposures to unfamiliar sounds during listening exercises. Next, the findings suggest a link between vocabulary knowledge and sound identification. Then, this highlights the need for integrated learning approaches that combine vocabulary development with listening practice. Besides, the study indicates that participants found recognizing Sokuon easier in written text compared to spoken pronunciation. Furthermore, this suggests a gap between learners' reading and listening comprehension skills, and this points towards incorporating pronunciation practice alongside reading activities. Finally, the research acknowledges a lack of confidence among participants in following instructions to listen only once. This might be a factor in future studies and highlights the importance of building learner confidence and clear communication of expectations.

REFERENCES

Baydar, S. (2020). *EVALUATION OF A VISUAL FEEDBACK TOOL FOR SPELLING*. *August*.

Dahmen, S., Grice, M., & Roessig, S. (2023). Prosodic and Segmental Aspects of Pronunciation Training and Their Effects on L2. *Languages*, *8*(1), 74. https://doi.org/10.3390/languages8010074

Djafri, F., & Wahidati, L. (2020). Study in Japan and the Motivation of Japanese Language Learners in Higher Educational Institutions in Indonesia. *Izumi*, *9*(2), 112–120. https://doi.org/10.14710/izumi.9.2.112-120

Febriyanti, R., & Husna, L. (2020). Analisis Kesalahan Berbahasa Jepang. *PELANGI SASTRA*.

Febriyanti, R., & Indrowati, S. A. (2015). THE PRONUNCIATION OF LONG VOWEL SOUNDS (CHOUON) BY STUDENTS, TEACHERS AND NATIVE SPEAKERS OF DEPARTMENT OF JAPANESE LANGUAGE TEACHING OF BRAWIJAYA UNIVERSITY. *EDUCATE*, *3*.

Flege, J. E., & Bohn, O. (2021). *(c) 2021 Cambridge University Press 1he Revised Speech Learning Model (SLM-r)* (Issue c).

Fujimoto, M., Shinohara, S., & Mochihashi, D. (2021). Articulation of geminate obstruents in the Ikema dialect of Miyako Ryukyuan: A real-time MRI analysis. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100321000013

Fujita, R. (2022). the Role of Speech-in-Noise in Japanese Efl Learners' Listening Comprehension Process and Their Use of Contextual Information. *International Journal of Listening*, *36*(2), 118–137. https://doi.org/10.1080/10904018.2021.1963252

Garellek, M. (2019). The phonetics of voice. *The Routledge Handbook of Phonetics*, *January*, 75–106. https://doi.org/10.4324/9780429056253-5

Kochva, I. B., Vágvölgyi, R., Dresler, T., Nagengast, B., Schröter, H., & Schrader, J. (2021). Basic reading and reading - related language skills in adults with deficient reading comprehension who read a transparent orthography. *Reading and Writing*, *34*(9), 2357–2379. https://doi.org/10.1007/s11145-021-10147-4

Kubozono, H. (2017). *Moras and Syllables in Japanese Dialects **. *25580098*, 1–13.

Lasut, P. A. (2015). Word Stress Contribution in Second Language Acquisition. *Journal of Language and Literature*, *15*(2), 163–167.

Liu, S., & Takeda, K. (2021a). Mora-timed, stress-timed, and syllable-timed rhythm classes: Clues in English speech production by bilingual speakers. *Acta Linguistica Academica*, *68*, 350–369. https://doi.org/10.1556/2062.2021.00469

Liu, S., & Takeda, K. (2021b). Mora-timed, stress-timed, and syllable-timed rhythm classes: Clues in English speech production by bilingual speakers. *Acta Linguistica Academica*, *68*(3), 350–369. https://doi.org/10.1556/2062.2021.00469

Mitterer, H., Eger, N. A., & Reinisch, E. (2020). My English sounds better than yours: Second-language learners perceive their own accent as better than that of their peers. *PLoS ONE*, *15*(2), 1–12.

https://doi.org/10.1371/journal.pone.0227643

Ofuka, E. (1997). *Long and.*

Ota, M., Yamane, N., & Mazuka, R. (2018). The effects of lexical pitch accent on infant word recognition in Japanese. *Frontiers in Psychology*, *8*(JAN), 1–13. https://doi.org/10.3389/fpsyg.2017.02354

Pendidikan, J., Jepang, B., & Najoan, F. R. (2021). *JAPANEDU : The Perceptual Ability to Identify Japanese Language Accent Patterns by Indonesian Learners*. *06*(02), 144–152.

Pimentel, T., Roark, B., & Cotterell, R. (2020). Phonotactic complexity and its trade-offs. *Transactions of the Association for Computational Linguistics*, 8, 1–18. https://doi.org/10.1162/tacl_a_00296

Prastiti, N., Samopa, F., Informasi, J. S., Teknik, F., Madura, U. T., Informasi, J. S., & Informasi, F. T. (2015). *Analisis Pelafalan Bahasa Jepang Sebagai IDENTIFIKASI PENULISAN AKSARA JEPANG BERDASARKAN AKSEN*. *1*(3), 173–181.

Rahmi, R. (2016). the Development of Language Policy in Indonesia. *Englisia Journal*, *3*(1), 9. https://doi.org/10.22373/ej.v3i1.622

Rasier, L., & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux Cahiers de Linguistique Française*, *28*(July 2015), 41–66. http://books.google.co.kr/books?hl=en&lr=&id=JPJP64OMqlUC&oi=fnd&pg=PA41&dq=L2+phonology+transfer&ots=u3hey_XYZ5&sig=Uyrs-B3gvfM6SawM0n3Ca8SdaOo%5Cnhttp://books.google.co.kr/books?hl=en&lr=&id=JPJP64OMqlUC&oi=fnd&pg=PA41&dq=prosodic+transfer+hypothesis&ot

Ren, H. (2022). Perception of Japanese singleton and geminate contrasts: A case of Chinese learners with different dialectal backgrounds. *Frontiers in Psychology*, *13*(December), 1–18. https://doi.org/10.3389/fpsyg.2022.1070107

Romero, Y., & Manjarres, M. P. (2017). How Does the First Language Have an Influence on Language Learning? A Case study in an English ESL Classroom. *English Language Teaching*, *10*(7), 123. https://doi.org/10.5539/elt.v10n7p123

Rose, H., McKinley, J., & Baffoe-Djan, J. B. (2019). Data Collection Research Methods in Applied Linguistics. *Data Collection Research Methods in Applied Linguistics*, *November*. https://doi.org/10.5040/9781350025875

Rosen, A., Trauer, T., Hadzi-Pavlovic, D., Parker, G., Patton, J. R., Cronin, M. E., Bassett, D. S., Koppel, A. E., Zimpher, N. L., Thurlings, M., Evers, A. T., Vermeulen, M., Obanya, P., Avsec, S., Nurzarina Amran, Liu, S. H., Petko, D., Aesaert, K., Van Braak, J., … Brown, N. (2015). No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析 Title. *Teaching and Teacher Education*, *12*(1), 1–17. http://dx.doi.org/10.1080/01443410.2015.1044943%0Ahttp://dx.doi.org/10.1016/j.sbspro.2010.03.581%0Ahttps://publications.europa.eu/en/publication-detail/-/publication/2547ebf4-bd21-46e8-88e9-f53c1b3b927f/language-en%0Ahttp://europa.eu/.%0Ahttp://www.leg.st

Saji, N., Akita, K., Kantartzis, K., Kita, S., & Imai, M. (2019). Cross-linguistically shared and language-specific sound symbolism in novel words elicited by locomotion videos in Japanese and English. *PLoS ONE*, *14*(7), 1–24. https://doi.org/10.1371/journal.pone.0218707

Seguel, S., & Sonesson, K. R. (2022). *Teaching Kanji at Swedish Universities On Teaching Kanji , Learning Strategies , and How These Are Perceived by Teachers and Students of L2 Japanese in Sweden*.

Shiro, K. (n.d.). *A perceptual study on Japanese intonation*.

Tateishi, K. (2017). More Arguments against Japanese as a Mora Language. *Proceedings of the 34th West Coast Conference on Formal Linguistics*, 529–535.

Tyler, M. D. (2019). *PAM-L2 and phonological category acquisition in the foreign language classroom PAM-L2 and Phonological Category Acquisition in the Foreign Language Classroom. May*.

Wahyuni, M. I., & Sutedi, D. (2020). *Error Analysis of Sokuon and Choon Used by Japanese Learners*. *424*(Icollite 2019), 11–14. https://doi.org/10.2991/assehr.k.200325.044

Winarni, R. S., Rasiban, L. M., & Juangsih, J. (2022). The Effect of YouTube Video Learning Media

on the Students' Basic Japanese Listening Ability. *Proceedings of the Sixth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2022)*, 455–459. https://doi.org/10.2991/978-2-494069-91-6_71

Yazawa, K., Konishi, T., Whang, J., Escudero, P., & Kondo, M. (2023). Spectral and temporal implementation of Japanese speakers' English vowel categories: A corpus-based study. *Laboratory Phonology*, *24*(1), 1–33. https://doi.org/10.16995/labphon.6427