

Prioritized English Sounds in Language Learning

ISONO Toru

要旨

学習時間が有限である以上、優先的に習得すべき事項を明確にする取り組みは学習の効率化という面において非常に重要なものである。第二言語音声習得においては、ネイティブ・スピーカーと遜色ない発話能力の習得が究極のゴールとして存在する一方で、発話の明瞭性を少ない労力で上げるというアプローチも一つの現実的なゴールとして認識されている。しかし、具体的に何を優先して学習していけばいいかという問題に関しての議論はまだ深まっていない。本研究では、各英語音がもつ習得上の重要性とその音の習得に必要な労力を考え併せながら、どの英語音を優先的に習得していくべきかを提言していく。「重要性」の度合いに関しては機能負荷量 (functional load) で示された数値を本研究では参照していく。「習得上の労力」、つまり日本人英語学習者による知覚・調音両面における英語音の習得傾向に関しては、筆者が過去に行ってきたいくつかの実証研究結果を総合的に再検証することで明らかにする。これらの研究は異なった年に、異なる被験者を対象に、異なる目的・分析方法を用いて行われたものではあるが、被験者は皆、調査時には英語を専門に学ぶ大学1・2年生であったという共通点がある。よって、それらの結果を総合的に再分析することにより、当該グループが示す一般的な傾向を知ることができるであろう。この議論をもとに、日本における英語学習上比較的大きな注目を集めている英語音 ([p], [b], [t], [d], [k], [g], [f], [v], [θ], [ð], [s], [z], [l], [r], [æ], [ʌ]) を対象に、どのような優先順序で習得していくべきかを提案することで本論の結論とする。

Keywords: English sounds, Japanese learners, Perception, Articulation, Functional load

1. Introduction

As far as the goal of perception/pronunciation teaching to Second Language (L2) learners is concerned, there are two kinds of approaches. One approach is to fix the goal at a near-native level. This approach considers any misperception/foreign accent as a problematic feature and aims to eliminate it altogether. The audio-lingual method, which was strongly influenced by behaviorism, had a similar idea and claimed that no errors produced by an L2 learner should be overlooked. In this sense, as far as the acquisition of L2 sounds is concerned, an L2 learner's task is, theoretically, to master the correct perceptions/articulations of all L2 sounds, and a language teacher's task is to correct all deviations in order to make it possible for an L2 learner to perceive/pronounce the norm of the target language.

The other approach sets the goal to improve L2 learners' perceptions and pronunciations so that they can understand the L2 content and be understood by native speakers without too much difficulty. For the past several decades, this approach has been supported by many experts and scholars (e.g., Pennington & Richards, 1986; Crawford, 1987; Morley, 1991). This assumption is partly affected by the concept of interlanguage that not all features deviating from the norm are necessarily considered as errors that should be corrected. Mastering the correct perceptions and articulations of all L2 sounds is, naturally enough, a supremely difficult task. Demanding that L2 learners attempt such a task from the beginning stage might cause them to accelerate the use of processes which make up for the insufficiency of their L2 knowledge.

In addition, correcting all deviations does not necessarily serve the aim of nurturing competence in communication. It has been said that a language teacher does better by leaving minor errors which do not hinder communication without correction, so that the flow of communication can continue comfortably without interruption. With regard to this point, an important but difficult question emerges: what kind of deviations should a language teacher correct with priority? The purpose of this paper is to propose the ideal order of teaching English sounds, by mainly taking into consideration the following two factors, *Importance* and *Acquisition Rate*.

2. Importance and Acquisition Rate

Importance is determined by the extent to which an error hinders communication. At the phoneme level, when a particular phoneme is replaced by another existing phoneme, this deviation will theoretically lead native listeners to misunderstand or fail to recognize what is said. However, in practice, since phonemes occur within words in real conversation, other sounds and phonological patterns within the word can often make it possible for listeners to understand a word which contains deviations in phonemes. Therefore, in phonology, the importance of distinguishing phonemes is represented by the concept known as functional load (Catford, 1987, cited in Derwing & Munro, 2014). A phoneme is considered to have a high functional load when it contributes to the differentiation of a large number of words. In general, phonemes that form numerous minimal pairs tend to have a high functional load and should be prioritized in L2 sound instruction. Although this concept was originally devised with articulation in mind, it also offers significant pedagogical implications for perception, which is closely related to articulation.

Acquisition Rate is deeply connected to one of the central issues in Second Language Acquisition (SLA) research, which is to clarify why some L2 sounds are acquired earlier and more easily than others. This question has been argued in terms of the phonological distance between First Language (L1) and L2 and the amount of attention L2 learners pay to L2 sounds. For example, Major and Kim (1996) claim that even L2 sounds which are generally considered difficult can be acquired more quickly and earlier than others as a result of the considerable amount of attention being paid by learners. Items that attract attention are usually things which are greatly different from L1 systems because they are perceived as “new” by L2 learners. Concerning the data on Acquisition Rate, we refer to the experimental results¹ reported by the present author (Isono, 2019; 2020; 2025a; 2025b; 2026). All of these studies investigated Japanese learners’ acquisition of English sounds, but each study focused on different target sounds (e.g., the 2019 study examined plosives, but the 2020 study did fricatives). In addition, the first four studies focused solely on perception, while the 2026 study investigated the relationship between perception and articulation of English sounds. Thus, while it is true that these surveys lack uniformity in terms of adopted criteria, because of having been conducted in different years with different subjects and for different research purposes, the subjects were consistently first-

and second-year students enrolled in the faculty of International Communication, Aichi University, at the time of each survey. Therefore, by comprehensively reanalyzing the findings of these five studies, we assume that we can present results that more accurately reflect the acquisition process of English sounds by Japanese university students.

In addition to the above two, this paper also refers to two more factors, *Difficulty* and *Acquisition Speed*. Difficulty refers to the degree of effort required for L2 learners to accurately perceive and produce sounds. From the 1940s to the 1960s, SLA researchers believed that they could establish more effective teaching methods by conducting Contrastive Analysis (CA) which is a technique comparing points of similarity and difference between L1 and L2. The basis of CA lay in the assumption that when L1 and L2 are similar, L2 elements are easy to learn; however, when they are different, L2 elements become difficult for learners to acquire. Unlike other fields in SLA, we can say that CA predictions are insightful in the acquisition of L2 sounds, and using CA to examine differences between the English and Japanese sound systems has been successful in its own right by explaining why Japanese learners have difficulties with some particular aspects of English sounds. Furthermore, methods for comparing L1 and L2 sound systems become increasingly refined. The Perceptual Assimilation Model (PAM) (Best, 1995; Best and Tyler, 2007) proposes several patterns in the relationship between L1 and L2 sound systems, and assigns a level of difficulty to each pattern.

Acquisition Speed shares many aspects with Acquisition Rate, but this paper distinguishes between the two according to the following definitions. Acquisition Rate, which was outlined before, is an indicator that, from a long-term perspective, clarifies which sounds can be acquired at an early stage. On the other hand, Acquisition Speed emphasizes short-term effects and demonstrates how quickly learners can progress after receiving explicit instruction.

In this paper, the arguments will be developed mainly using the first two assumptions (Importance and Acquisition Rate), and the other two (Difficulty and Acquisition Speed) will be employed as supplementary explanations for the ideal teaching order of English sounds, which will be proposed in this paper.

3. Five Levels of Acquisition Rate

The most important part of this paper, to which the greatest number of pages is

devoted, is the establishment of Acquisition Rate. For this purpose, we comprehensively examine the previous data concerning Japanese learners' perceptual and articulatory abilities with respect to English sounds, and propose acquisition rates of Japanese learners for each English sound using the following five levels:

Level 1 – Learners attain a high level of ability regardless of their proficiency levels

Level 2 – Beginners may experience slight difficulties

Level 3 – Learners attain a fairly level of ability regardless of their proficiency levels

Level 4 – Only advanced learners attain a high level of ability

Level 5 – Even advanced learners cannot attain a high level of ability

As for Importance, we refer to the previous discussions of functional load as mentioned earlier, so these indicators will be presented together in tables showing the results of the above discussion of Acquisition Rate.

3–1. Plosives

Isono (2019) examined the perceptual abilities of 53 subjects with regard to the six English plosives ([p], [b], [t], [d], [k], [g]) in word-initial and word-final positions.

Table 1. Perceptual Accuracy for Plosives in Word-initial and Word-final Position

	Word-initial	Word-final	Difference
[p]	98.52% (2.54)	93.92% (6.82)	4.60%
[b]	96.71% (3.94)	88.39% (19.7)	8.32%
[t]	98.86% (2.27)	94.49% (5.70)	4.37%
[d]	95.15% (9.63)	97.52% (2.74)	–2.37%
[k]	95.41% (5.03)	96.88% (4.72)	–1.47%
[g]	90.03% (12.1)	92.07% (12.1)	–2.04%

(Isono, 2019, p. 5)

Based on the concept of markedness, features which are simple and common in natural languages are said to be unmarked, but those which are complex and less common are said to be marked. In language acquisition, marked features are, naturally enough, more difficult to acquire than unmarked ones. Concerning the positions where

voicing contrasts occur, the word-final position is considered to be a more marked position than word-initial. This hypothesis was supported by the results of [p], [b], and [t], but not by those for [d], [k], and [g]. In addition, concerning voicing contrast, voiced plosives are said to be more marked than voiceless ones. This hypothesis was generally supported, but the accuracy rates for the voiced plosives were also high enough. These results suggest that Japanese learners do not have difficulty in distinguishing English plosives, no matter whether they are voiced or placed in word-final position, within these six English plosives.

However, as the following findings indicate, when the target sounds of the investigation included not only plosives but also other consonants, the accuracy rates of voiced sounds became lower.

Table 2. Perceptual Accuracy for Plosives

	Group 1 (Beginners)	Group 2 (Intermediate)	Group 3 (Advanced)
[p]	89.6% (14.3)	94.0% (9.95)	96.4% (7.85)
[b]	62.7% (28.0)	62.0% (26.8)	74.1% (22.0)
[t]	89.3% (17.8)	94.0% (9.08)	95.2% (11.2)
[d]	78.6% (21.9)	86.1% (16.1)	93.8% (14.0)
[k]	94.6% (9.72)	95.6% (12.1)	98.5% (9.85)
[g]	73.2% (25.4)	78.6% (23.7)	87.6% (12.1)

(Isono, 2025b, p. 7)

Isono (2025b) divided the Japanese subjects into the above three groups based on their TOEIC listening scores, and investigated how accurately they could perceive English sounds, including fricatives and other consonants. The research methods remained the same as in Isono (2019), except that the number of items increased. As we can see from the table, the accuracy rates on the voiced plosives ([b], [d], [g]) were significantly lower than those on the voiceless plosives ([p], [t], [k]). Considering the above reported results that Japanese learners do not have difficulty in distinguishing English plosives within the six English plosives, this discrepancy can be attributed to the misperception with other fricatives and affricates, such as [v], [ð], [dʒ], which will be examined later.

Tendencies in the articulatory abilities of Japanese learners are discussed with

reference to Isono (2026), who investigated the relationship between perceptual and articulatory abilities with regard to English sounds after dividing the subjects into two groups based on their perceptual abilities.

Table 3. Articulatory Accuracy for English Plosives

	[p]	[b]	[t]	[d]	[k]	[g]
Accuracy	89.7%	86.2%	86.7%	96.4%	90.0%	72.0%

(Isono, 2026, p. 10, revised by the author)

Concerning articulatory abilities for English plosives, the Japanese learners seemed to attain high accuracy, except for [g]. As for the example of misarticulation of [g], there were, as expected, instances of the corresponding voiceless plosives [k], but also instances of [j] and [dʒ]. Another point worth mentioning concerns the articulatory abilities for [t]. Overall, it showed a high accuracy rate of 86.7%, but in the less-advanced group, many instances of misarticulation as [d] were observed, which contributed to a drop in accuracy to 75.0%. Conversely, no such instances were not found in the advanced group (Advanced – 100%, Beginner – 75.0%). The importance of distinguishing voiced and voiceless plosives is frequently emphasized in English learning. As far as the present results are concerned, Japanese learners of English can distinguish them without difficulty in terms of auditory perception, but experience slight difficulties in terms of articulation.

The summary of the discussion thus far is presented in the below table, based on the criteria outlined earlier.

Table 4. Levels of Acquisition Rate for English Plosives

	Perception	Articulation	Importance (functional load)
[p]	Lv. 1	Lv. 1	p/b (98), p/k (92), p/t (87), p/h (85), p/f (77)
[b]	Lv. 4	Lv. 1	p/b (98), b/d (82), b/w (76), b/g (71), b/v (29)
[t]	Lv. 1	Lv. 2	t/k (81), t/s (81), t/d (73), t/tʃ (39), t/θ (18)
[d]	Lv. 2	Lv. 1	b/d (82), d/l (79), d/r (75), d/g (56), d/n (53), d/dʒ (39), d/ð (19), d/z (7)
[k]	Lv. 1	Lv. 1	k/h (100), p/k (92), k/g (50)
[g]	Lv. 2	Lv. 3	b/g (71), d/g (56), k/g (50), g/w (49), g/dʒ (31)

In the column of Importance (functional load), the numbers in the parentheses indicate the percentage of functional load. For example, many words are distinguished by the contrast between [p] and [b] (e.g., pat/bat, pan/ban), therefore, this sound pair is said to have a high functional load of 98%, and its importance for acquisition is very high. In contrast, sound pairs with low functional load pair are less important for acquisition, because they distinguish fewer words.

From the table, it becomes clear to which sounds class time should be devoted. Although [p] and [k] are included in many sound pairs with high functional loads, class time does not need to be devoted to these sounds, because Japanese learners already demonstrate high accuracy in both perception and articulation. Japanese learners show a Level 4 rate in the perception of [b], but the distinction between [b] and [v], which contributes to this Level 4 rating, has a relatively low percentage of 29%. Therefore, when considering its level, its priority for learning is not particularly high. Similarly, Japanese learners, especially beginners, may have difficulty in the perception of [d] (Level 2), but the distinction between [d] and [ð], which is often cited as a typical error among Japanese learners, has a low percentage of 19%, so its necessity for learning is also not particularly high.

On the other hand, Japanese learners may have articulatory difficulty in distinguishing [t] and [d] at the early stages of learning, and this sound pair has a relatively high functional load (73%), so the priority of acquiring accurate articulation ability is relatively high. Among English plosives, when considering both Acquisition Rate and Importance, the acquisition of [g] should be given the highest priority in both of perceptual and articulatory abilities.

3-2. Fricatives

As for the reanalysis of fricatives, we follow the same procedure as that used for plosives. Isono (2020) investigated the perceptual abilities of 52 Japanese learners with regard to English fricatives and plosives in word-initial and word-final positions. The target sounds were six English fricatives ([f] – [v], [θ] – [ð], and [s] – [z]) and six English plosives in word-initial position, and all of those sounds except [ð] in word-final position².

The experimental results revealed high accuracy rates for English plosives (95.1%), whereas the perception of English fricatives was considerably lower (69.2%), as shown in the table below.

Table 5. Perceptual Accuracy for Fricatives in Word-initial and Word-final Position

	Word-initial	Word-final
[f]	88.00% (10.88)	82.09% (27.64)
[v]	49.06% (30.70)	50.00% (36.79)
[θ]	64.41% (40.11)	73.53% (28.08)
[ð]	63.86% (26.08)	—
[s]	82.23% (14.20)	80.57% (19.69)
[z]	59.28% (30.05)	63.31% (24.82)

(Isono, 2020, p. 44)

Among the English fricatives listed in the above table, those not presented in Japanese are [f], [v], [θ], and [ð]. While the Japanese fricative [ɸ] corresponds to the English fricative [f], confusions between [v] and [b], [θ] and [s], and [ð] and [z] are often observed. The sound relationships like [v] – [b], [θ] – [s], and [ð] – [z] are classified as “Uncategorized-Categorized Assimilation” in PAM, and those sounds are considered relatively difficult, especially for beginners who tend to confuse them. The above results generally supported these assumptions.

Another point that becomes clear was that there was no substantial difference in perception between English fricatives in word-initial position and those in word-final position. Since similar tendencies were observed in Japanese learners’ perception of English plosives, it is evident that whether sounds occur at the beginning or end of a word is not particularly significant in Japanese learners’ perception of English plosives and fricatives.

Next, we examine how the perceptual accuracy of the above fricatives improves across learning stages, based on Isono (2025b) who investigated how accurately Japanese learners at different proficiency levels could perceive English sounds, including plosives, fricatives, and other consonants.

Table 6. Perceptual Accuracy for Fricatives

	Group 1 (Beginners)	Group 2 (Intermediate)	Group 3 (Advanced)
[f]	74.8% (22.4)	75.2% (20.6)	76.1% (28.2)
[v]	63.9% (23.5)	75.4% (22.9)	78.2% (20.6)
[θ]	40.8% (24.8)	34.7% (29.1)	55.8% (24.3)
[ð]	32.9% (30.7)	51.3% (33.9)	63.8% (31.5)
[s]	69.1% (30.8)	88.4% (14.2)	93.2% (14.0)
[z]	56.0% (32.3)	66.5% (29.1)	45.8% (31.0)

(Isono, 2025b, p. 9, revised by the author)

The results in the above table can be categorized into the following three tendencies. First, there was no significant difference in the perceptual accuracy of the English fricative [f] among the groups, although the rates were slightly lower than expected (Level 3). Second, the results for the English fricatives [v] and [s] were relatively low among beginners but improved as learners advance to higher levels (Level 2). Third, the results for the English fricatives [θ], [ð], and [z] showed low overall accuracies across all the groups. Concerning the results for [θ] and [ð], although significant improvements were observed in Group 3, the accuracy rates remained below 70% (Level 5).

As for Japanese learners' articulatory abilities for these English fricatives, we refer to Isono (2026), as in the analysis of English plosives. However, since the English fricative [f] was not included in the test items in Isono (2026), it is not shown in the below table.

Table 7. Articulatory Accuracy for English Fricatives

	[v]	[θ]	[ð]	[s]	[z]
Accuracy	44.1%	50.8%	67.6%	65.5%	62.8%

(Isono, 2026, p. 12, revised by the author³)

As the table shows, while the articulatory accuracy for the first two ([v] and [θ]) was relatively low, it was comparatively high for the latter three ([ð], [s], and [z]). When the first two results were examined separately for advanced and beginner groups, both fricatives showed similar tendencies ([v]: Advanced – 61.5%, Beginner – 26.7%; [θ]:

Advanced – 60.7%, Beginner – 42.3%). Since the beginners’ results (26.7% and 42.3%) should be regarded as indicating “significant” difficulty rather than “slight” difficulty, the articulation of these sounds is classified as Level 4 in this paper.

Concerning the latter three fricatives ([ð], [s], and [z]), the articulatory accuracy was around 65%, however, when examined separately for beginners and advanced learners in the same way as above, distinct trends emerged for each group. Similarly to the first two fricatives ([v] and [θ]), the fricative that showed a significant difference between the groups was [z] (Advanced – 85.7%, Beginner – 40.0%), and thus this sound is also classified as Level 4. On the other hand, the beginner group attained fairly accuracy for the other two fricatives. As for the articulation of the fricative [s], it is classified as Level 3 in this paper, since the percentages were not particularly high and significant differences were not observed between the groups (Advanced – 67.8%, Beginner – 76.3%). Regarding the articulation of the fricative [ð] (Advanced – 76.0%, Beginner – 59.9%), it is classified as Level 2 in this paper due to the relatively low percentage for beginners.

The discussion of Japanese learners’ acquisition of English fricatives can be summarized in the below table, together with the functional load values.

Table 8. Levels of Acquisition Rate for English Fricatives

	Perception	Articulation	Importance (functional load)
[f]	Lv. 3	—	p/f (77), f/h (69), f/s (64), f/v (23), f/hw (13)
[v]	Lv. 2	Lv. 4	b/v (29), f/v (23), v/w (22), v/ð (11), v/z (2)
[θ]	Lv. 5	Lv. 4	s/θ (21), t/θ (18), f/θ (15), θ/ð (1)
[ð]	Lv. 5	Lv. 2	d/ð (19), v/ð (11), θ/ð (1), z/ð (1)
[s]	Lv. 2	Lv. 3	s/h (85), t/s (81), f/s (64), s/ʃ (53), s/tʃ (37), s/θ (21), s/z (6)
[z]	Lv. 5	Lv. 4	d/z (7), s/z (6), v/z (2), z/ð (1)

English fricatives have attracted considerable attention in English teaching, as they include sounds not present in Japanese, which often leads learners to substitute Japanese sounds for them. This tendency might explain the above results showing the low acquisition rates of English fricatives. Specifically, the perceptual abilities for the English [θ], [ð], and [z] are not sufficiently developed even among advanced learners, and the articulation abilities of [v], [θ], and [z] are unlikely to be attained by many learners except at advanced levels. However, as indicated by the low functional load values for these

sounds, even if learners make considerable efforts to acquire the ability to distinguish them, this does not significantly contribute to distinguishing English words in actual communication. Mastering these sounds is important for the purpose of achieving native-like English pronunciation, however, when the goal is to improve intelligibility with minimum effort, their priority becomes lower.

From the above table, it is clear that mastering the accurate perception and proper articulation of the English [s] has a much higher priority. Although both Japanese and English have this sound, learners cannot perceive or produce it with the same degree of accuracy as English plosives. In addition, as the high functional load values in many pairs of sound contrasts show, mastering this sounds may greatly contribute to distinguishing English words in actual communication, unlike other English fricatives.

3-3. Approximants

Among the English approximants ([l], [r], [j], and [w]), we focus on the acquisition of [l] and [r], which have always been central in English teaching in Japan and to which considerable effort has been devoted. In the early stages of learning, the English [l] and [r] are assumed to be assimilated to the Japanese [r] sound without distinction. This relationship is considered to correspond to “Single-Category Assimilation” in PAM. As learning progresses, learners are thought to come to distinguish the two English sounds, with the assumption that the English [l] is similar to the Japanese [r] sound but the English [r] is not. This pattern corresponds to “Category-Goodness Assimilation” in PAM-2, which applies to learners with a certain experience of language learning. “Single-Category Assimilation” and “Category-Goodness Assimilation” are considered to be among the most difficult patterns.

The below table shows the results of Isono (2025b), which investigated the perceptual abilities of Japanese learners for the English [l] and [r].

Table 9. Perceptual Accuracy for the English [l] and [r]

	Group 1 (Beginners)	Group 2 (Intermediate)	Group 3 (Advanced)
[l]	47.9% (24.5)	52.5% (27.3)	80.2% (20.0)
[r]	49.3% (29.8)	55.6% (31.9)	80.0% (24.4)

(Isono, 2025b, p. 11, revised by the author)

As shown in the table, similar tendencies were observed for the two sounds. In both cases, advanced learners attained significantly higher abilities to distinguish the two sounds. Therefore, the perception of these sounds corresponds to Level 4 according to the criteria outlined before.

On the other hand, the developmental process of the articulatory abilities for these sounds shows different tendencies, as reported by Isono (2026).

Table 10. Articulatory Accuracy for English [l] and [r]

	[l]	[r]
Accuracy	82.2%	62.6%

(Isono, 2026, p. 15, revised by the author)

The results showed that the Japanese learners attained high accuracy in articulating the English [l], and large differences were not observed across proficiency levels (Advanced – 89.3%, Beginner – 77.6%). On the other hand, this pattern was not observed for the English [r], and the beginners showed considerably low articulatory accuracy (Advanced – 82.1%; Beginner – 44.1%). Based on these results, the acquisition process of articulating the English [l] corresponds to Level 1, whereas that of articulating the English [r] corresponds to Level 4.

Table 11. Levels of Acquisition Rate for the English [l] and [r]

	Perception	Articulation	Importance (functional load)
[l]	Lv. 4	Lv. 1	l/r (83), d/l (79), n/l (61)
[r]	Lv. 4	Lv. 4	l/r (83), d/r (75), n/r (41)

Considering the functional load values for those sounds, it is obvious that they should be prioritized in learning, regardless of their difficulty. One encouraging finding is that Japanese learners can improve the intelligibility of their English [l] without much effort. In addition, the perception of English [l] and [r], as well as the articulation of English [r], can be achieved by advanced learners, although they require a certain amount of time to master.

3–4. Vowels

As with English consonants, there are many pairs of English vowels forming minimal pairs. Among these, this paper focuses on the pair of [æ] and [ʌ], which receive considerable attention in teaching because both are often substituted by the Japanese [a]. The table below shows the Japanese learners' perceptual and articulatory abilities for those English vowels, as reported by Isono (2026).

Table 12. Perceptual and Articulatory Accuracies for English [æ] and [ʌ]

	[æ]	[ʌ]
Perceptual Accuracy	67.6%	71.8%
Articulatory Accuracy	80.0%	81.4%

(Isono, 2026, p. 16, revised by the author)

The perceptual accuracy rates for the English [æ] and the English [ʌ] were 67.6% and 71.8%, respectively. Considering that the task required a forced choice between the English [æ] and [ʌ], we cannot say that the two vowels can be distinguished with absolute certainty by Japanese learners. The difficulty in distinguishing these vowels was also reported by Isono (2025a), and no significant improvement was observed when comparing the perceptual results before and after the English phonetic course.

Turning to the articulation accuracy, however, we see that the articulation accuracy rates were higher than the perceptual accuracy rates in both cases. The high articulation accuracy for the English [ʌ] was easily expected, given its similarity to the Japanese [a], and no significant difference was observed across their proficiency levels (Advanced – 80.6%, Beginner – 79.0%). On the other hand, for the English [æ], which is considerably different from the Japanese [a], relatively high accuracy rates were also observed overall (Advanced – 81.8%, Beginner – 77.0%). These results can be summarized in the table below.

Table 13. Levels of Acquisition Rate for the English [æ] and [ʌ]

	Perception	Articulation	Importance (functional load)
[æ]	Lv. 3	Lv. 1	cat/cut (68)
[ʌ]	Lv. 3	Lv. 1	cat/cut (68)

Taking into account the high functional load values, there is considerable justification for the approach taken in English education in Japan, which emphasizes distinguishing between these vowels. However, it should be noted that the articulation tends to develop earlier than the perception. This tendency is often observed when learners pay a great amount of attention to a target sound as a result of perceiving it as “new”, and when the sound does not require complex articulatory control.

In addition to the above vowels, the distinctions between [i:] and [ɪ] (beet/bit) and between [ɔ:] and [ou] (bought/boat) are also associated with high functional load values. The Japanese sound system is characterized by the fact that long vowels and the corresponding short vowels have the same sound quality and are distinguished only by the sound duration, and that long vowels and diphthongs are not clearly distinguished. Taking these characteristics into account, both are important to master; however, the latter is much more important in terms of intelligibility, as Isono (2003) demonstrated that Japanese learners’ prolongation of long vowels was intelligible enough for native speakers to distinguish long vowels from short ones.

4. Conclusion

Based on the discussion in this paper, we can classify English sounds into tiers to effectively improve Japanese learners’ perception and the intelligibility of their pronunciation, as follows.

Table 14. Prioritized English Sounds

	(P = Perception, A = Articulation)
Tier 1	[g] (P – Lv. 2, A – Lv. 3), [s] (P – Lv. 2, A – Lv. 3), [l] (P – Lv. 4, A – Lv. 1), [r] (P – Lv. 4, A – Lv. 4) [æ] (P – Lv. 3, A – Lv. 1), [ʌ] (P – Lv. 3, A – Lv. 1)
Tier 2	[b] (P – Lv. 4, A – Lv. 1), [v] (P – Lv. 2, A – Lv. 4) [θ] (P – Lv. 5, A – Lv. 4), [ð] (P – Lv. 5, A – Lv. 2)
Tier 3	[z] (P – Lv. 5, A – Lv. 4)
Others	[p], [t], [d], [k]

The English sounds included in Tier 1 are those which are not extremely difficult and have high functional load values, and are therefore very important to master.

Although Japanese learners, especially beginners, might have difficulty with those sounds, it is reasonable to assume that articulating the English [l], [æ], and [ʌ] does not require much effort even for them. The English sounds included in Tier 2 are important for mastering native-like pronunciation, but they do not contribute to distinguishing between words to the same extent as those classified as Tier 1. Considering their lower importance and the difficulty of perceiving and articulating them together, the priority of acquiring these sounds is much lower than that of the sounds classified as Tier 1. The English [z] is classified as Tier 3 because it is extremely difficult for Japanese learners to master in both perception and articulation, and it rarely contributes to distinguishing between words, even when it is mastered. In terms of efficiency in learning, it would be more efficient to focus on learning other sounds rather than spending time on this sound. The four English sounds classified as “Others” are sufficiently intelligible when the corresponding Japanese sounds are substituted, so there is no need to pay particular attention to them, as long as the purpose of learning is to improve intelligibility.

Affricates and diphthongs are not included in the above table, as they were not examined in the series of studies by Isono (2019, 2020, 2025a, 2025b, 2026). By incorporating such information, it should become clearer which sounds we should prioritize in order to effectively improve the intelligibility of our English pronunciation.

Notes

1. Due to lack of space, the survey methods and subjects' information for each research are omitted. For details, please refer to the published papers.
2. The English fricative [ð] in word-final position was not examined in Isono (2020), due to the low number of the existing English words having the sound at the end.
3. The data of vat, zeal, and zinc, which were included for the analysis in the original research, are excluded in this analysis, because it was found that they were not suitable words to be tested.

References

- Best, C. T. 1995. A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research*. Baltimore: York Press, 171–204.
- Best, C. T. and Tyler, M. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In O. S. Bohn and M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning*. Amsterdam: John Benjamins Publishing Company, 13–34.

- Crawford, W. 1987. The Pronunciation Monitor: L2 acquisition considerations and pedagogical priorities. In Morley, J. (eds.), *Current Perspectives on Pronunciation: practices anchored in theory*. Washington, DC: TESOL.
- Derwing, T. & Munro, M. 2014. Once you have been speaking a second language for years, it's too late to change your pronunciation. In Grant, L. (Eds.), *Pronunciation Myths: Applying Second Language Research to Classroom Teaching* (pp. 34–55), University of Michigan Press.
- Isono, T. 2003. *Japanese Learners' Interlanguage Phonology: with special reference to English vowels and plosives*. Unpublished Ph.D. thesis. University of Essex.
- Isono, T. 2019. Perception of English Plosives by Japanese Learners. 『言語と文化』 第40号, 1–9.
- Isono, T. 2020. A Study of Japanese Learners' Perceptual Identification of English Fricatives. 『言語と文化』 第42号, 39–49.
- 磯野 徹. 2025a. 音声知覚能力の向上に及ぼす英語音声学の授業効果, 『言語と文化』 第51号, 61–78.
- 磯野 徹. 2025b. 子音知覚と聴解能力の関係に関する研究, 『言語と文化』 第52号, 1–15.
- 磯野 徹. 2026. 日本人英語学習者における知覚能力と調音能力の関係性, 『言語と文化』 第53号, 3–19.
- Major, R. and Kim, E. 1996. The similarity differential rate hypothesis. *Language Learning* 36.4, 505–522.
- Morley, J. 1991. The Pronunciation Component of Teaching English to Speakers of Other Languages. *TESOL Quarterly*, 25: 481–520.
- Pennington, M. and Richards, J. 1986. Pronunciation Revisited. *TESOL Quarterly*, 20: 207–225.