Voice Onset Time in World Englishes: Focus on Japanese EFL Learners

Hiroshi Nakanishi

1. Introduction

1.1 Universality and Diversity in VOT across languages

Voice Onset Time (VOT) is a crucial parameter for distinguishing voicing in stop consonants. Defined as the time interval between the release of a stop closure and the onset of vocal fold vibration (Shimizu, 1999), VOT serves as a key criterion for understanding both universal phonetic tendencies and language-specific adaptations. The foundational work of Lisker and Abramson (1964) provided a cross-linguistic perspective on VOT, while Cho and Ladefoged (1999) expanded upon this by analyzing a wider range of languages and refining the understanding of physiological factors affecting VOT.

Lisker and Abramson (1964) conducted a cross-linguistic analysis of VOT across 11 languages, identifying three primary categories of VOT: negative (pre-voicing), near-zero (short-lag), and positive (long-lag). Their results demonstrated significant cross-linguistic variations, as well as systematic differences in VOT by place of articulation. For example, in English, aspirated stops (/p/, /t/, /k/) exhibited VOT values of 58ms, 70ms, and 80ms, respectively, while Spanish unaspirated stops had much shorter VOT values (4ms, 9ms, and 29ms). While the characteristics of VOT vary across languages, a universal trend was observed: velar stops (/k/) consistently showed longer VOT than bilabial (/p/) or alveolar (/t/) stops, a finding attributed to physiological constraints such as the larger supraglottal cavity behind velar closures.

Table 1

Comparative VOT Data for Voiceless Stops (Lisker & Abramson,1964)

	Bilabial (/p/)	Alveolar (/t/)	Velar (/k/)
English (voiceless aspirated)	58ms	70ms	80ms
Spanish (voiceless unaspirated)	4ms	9ms	29ms
Korean (voiceless aspirated)	91ms	94ms	126ms
Korean (voiceless unaspirated)	7ms	11ms	19ms
Thai (voiceless aspirated)	64ms	65ms	100ms

Building on Lisker and Abramson's framework, Cho and Ladefoged (1999) investigated VOT across 18 languages, emphasizing the interaction between universal constraints and language-specific phonological systems. Their findings confirmed the universality of longer VOT values for velar stops but introduced a more fine-grained classification of voiceless aspirated stops, distinguishing between short, typical, and extremely long aspiration. For instance, Navajo exhibited remarkably long VOT values for aspirated stops, compared to languages like Gaelic, where the values were more moderate. They discuss that in Navajo and Tlingit, the reason for the exceptionally long VOT of voiceless aspirated stops may not only be attributed to the movement speed of articulators and aerodynamic factors but also to language-specific phonological rules. Furthermore, it explains that velar sounds (/k/) tend to have longer VOTs compared to bilabial (/p/) and alveolar sounds (/t/) due to factors such as aerodynamic considerations, the movement speed of articulators, and the contact area between the articulators.

Table 2

Comparative VOT Data for Voiceless Stops (Cho & Ladefoged, 1999)

	Bilabial (/p/)	Alveolar (/t/)	Velar (/k/)
Navajo (voiceless aspirated)		130ms	154ms
Navajo (voiceless unaspirated)	12ms	6ms	45ms
Gaelic (voiceless aspirated)	64ms		73ms
Tlingit (voiceless aspirated)			128ms

From here, the focus shifts to research on Japanese, which is the native language of the Japanese learners of English who serve as the participants in

this study. Shimizu (1996) acoustically investigated the voicing contrasts of stop consonants in Japanese, focusing particularly on the measurement of VOT. In his study, VOT measurements were conducted to examine (1) voiced-voiceless distinctions across three places of articulation (bilabial, alveolar, velar) and (2) the influence of following vowels, using data from six native Japanese speakers.

Table 3

VOT of Japanese Word-Initial Consonants by Following Vowel (Shimizu, 1996)

	/p/	/t/	/k/	/b/	/d/	/g/
/a/	37ms	29ms	53ms	-93ms	-73ms	-63ms
/i/	48ms		87ms	-92ms		-81ms
/u/			73ms			-70ms
/e/		29ms	55ms		-95ms	-83ms
/o/	40ms	31ms	63ms	-83ms	-57ms	-80ms

The main findings of Shimizu (1996) are as follows. From the VOT data across places of articulation for voiced and voiceless stop consonants:

1. The VOT of voiceless stops shows positive values, while the VOT of voiced stops shows negative values. This indicates that vocal fold vibration begins before the release for voiced stops, whereas it begins after the release for voiceless stops.
2. The VOT of the voiceless velar stop /k/ is the longest among the consonants, likely because its place of articulation is at the back of the oral cavity, which delays the onset of vocal fold vibration after the release.
3. The VOT of voiceless bilabial and alveolar stops is relatively short, suggesting that the forward placement of the articulatory point allows for quicker initiation of vocal fold vibration.

Furthermore, from the data on following vowels:

1. For voiceless stops, the height of the following vowel affects the VOT. Specifically, for the velar stop /k/, the VOT tends to be longer after high vowels (/i/ and /u/). This suggests that the elevation of the tongue during the production of high vowels may influence the timing of voicing onset.
2. For voiced stops, the variation in VOT caused by following vowels is

less pronounced and remains relatively stable compared to voiceless stops.

Furthermore, Shimizu (2010) introduced a study in which six Japanese native speakers and three American English native speakers produced words in their respective native languages (e.g., *田—打* for Japanese, *team-deam* for English), and the VOT values were analyzed. Table 4 shows the mean VOT values of Japanese initial voiced and voiceless stops in Shimizu (2010). For Japanese, the VOT values of stops are much shorter than those in English, as noted by Homma (2006). In both languages, a clear distinction between voiced and voiceless stops is observed. Additionally, in English, the VOT values increase as the place of articulation moves from bilabial to velar stops. A similar trend is observed in Japanese for velar stops.

Table 4

Mean VOT values (ms) and Standard Deviations (SD) in Japanese and English

	/p/	/t/	/k/	/b/	/d/	/g/
Japanese	41(17.1)	30(12.7)	66(12.1)	-89(28.5)	-75(32.7)	-75(27.0)
English	68(15.3)	82(18.6)	85(20.1)	-88(18.1)	-74(28.0)	-88(14.4)

1.2 Effects of L1 on L2 VOT

This section presents several studies that investigate how the VOT characteristics of learners' first languages (L1), as discussed earlier, affect the production of stop consonants in their second language (L2).

From the perspective of VOT, both Flege (1991) and Flege & Eefting (1987) provide valuable insights into the influence of L1 Spanish on L2 English phonetic acquisition. Their studies reveal that the phonological system of Spanish speakers' native language significantly impacts their ability to produce English voiceless stops /p, t, k/, as evidenced by acoustic analyses. As mentioned above, in Spanish, the phoneme /t/ is characterized by a short VOT, whereas English /t/ exhibits a much longer VOT in monolingual speakers. These differences highlight the challenges Spanish speakers face when acquiring the English /t/. Specifically, both studies demonstrate that the age of acquisition plays a crucial role in determining how closely Spanish learners approximate native English VOT values during

production.

Flege (1991) examined VOT in English /t/ among Spanish-English bilinguals. Early learners (exposed to English at ages 5-6) produced VOT values around 53 ms, similar to English monolingual norms, while late learners (exposed in adulthood) showed intermediate values around 33 ms. ANOVA revealed that monolinguals and early learners produced significantly longer VOTs than late learners ($F(2,27) = 15.2$, $p < 0.05$), with no significant difference between monolinguals and early learners. These results suggest late learners process L2 sounds through L1 phonetic categories.

Flege and Eefting (1987) investigated VOT category boundaries for /d/-/t/ contrasts. English monolinguals showed boundaries at 43 ms, Spanish monolinguals at 23 ms, and bilinguals at intermediate values (early learners: 27 ms; late learners: 29 ms). ANOVA with post-hoc tests revealed significantly longer VOT boundaries for English monolinguals compared to all other groups ($F(6,63) = 15.1$ $p < 0.01$). These findings indicate L1 influence persists even in early learners' perception. Table 5 summarizes results from both studies.

Both studies confirm that L1 significantly influences L2 phonetic acquisition, with pronounced effects among late learners. Interestingly, early learners approximate native English VOT values in production but still demonstrate residual L1 influence in perception.

Table 5

VOT Production and Perception Across Monolinguals (Spanish & English) and Bilinguals

	Spanish Monolinguals	English Monolinguals	Early Learners	Late Learners
VOT for /t/ (utterance-initial)	18 ms	51 ms	53 ms	33 ms
/d/–/t/ category boundary (perception)	23 ms	43 ms	27 ms	29 ms

Key characteristics	Short VOT	Long VOT	VOT close to English but L1 influence in perception	Intermediate VOT

Shimizu (2010) conducted an acoustic analysis examining how Korean L1 speakers produce English stops as L2, focusing on how their L1 phonological system influences this production. The study revealed that Korean, which has a three-way contrast in stops (tense unaspirated, weakly aspirated, and strongly aspirated), significantly influences L2 pronunciation patterns. Through VOT analysis, the research demonstrated that Korean speakers applied their L1 strongly aspirated category when producing English voiceless stops (/p/, /t/, /k/), resulting in appropriate long VOT values. The acoustic measurements showed specific VOT values across different places of articulation: for bilabials, voiced /b/ averaged 24.3ms and voiceless /p/ averaged 79.1ms; for alveolars, /d/ averaged 13.8ms and /t/ averaged 70.5ms; and for velars, /g/ averaged 39.3ms and /k/ averaged 101.0ms. For voiced stops (/b/, /d/, /g/), they applied their L1 tense unaspirated category, producing longer VOT values than standard English pronunciations (see Table 6).

Table 6

VOT Variations Across L1 Korean and L2 English by Place of Articulation and Sound Category

Language	**Place of Articulation**	**Sound Category**	**VOT (SD)**
Korean	Bilabial	Tense unaspirated (/p/)	19.0ms (5.7)
		Weakly aspirated (/pʰ/)	60.7ms (21.8)
		Strongly aspirated (/pʰʰ/)	95.4ms (20.2)
	Alveolar	Tense unaspirated (/t/)	17.8ms (9.6)
		Weakly aspirated (/tʰ/)	61.3ms (14.9)
		Strongly aspirated (/tʰʰ/)	90.0ms (7.3)
	Velar	Tense unaspirated (/k/)	33.6ms (12.1)
		Weakly aspirated (/kʰ/)	66.9ms (23.2)
		Strongly aspirated (/kʰʰ/)	90.3ms (7.6)

English	Bilabial	Voiced (/b/)	24.3ms (14.3)
		Voiceless (/p/)	79.1ms (30.8)
	Alveolar	Voiced (/d/)	13.8ms (28.7)
		Voiceless (/t/)	70.5ms (20.9)
	Velar	Voiced (/g/)	39.3ms (16.3)
		Voiceless (/k/)	101.0ms (27.5)

Harada (2007) investigated VOT production by L1 English children acquiring L2 Japanese in an immersion program. The acoustic analysis revealed that the immersion children produced Japanese voiceless stops /p, t, k/ with mean VOT values (66 ms) significantly longer than both monolingual Japanese children (26 ms) and their Japanese-English bilingual teachers (45 ms), while maintaining shorter values than their English VOT (88 ms). These intermediate VOT patterns demonstrate that the children have established distinct phonetic categories for Japanese and English stops, despite not achieving native-like values in either language. The findings provide empirical support for Flege's (1995) Speech Learning Model (SLM), particularly regarding the establishment of new phonetic categories when L1 and L2 sounds are sufficiently distinct. The partial overlap of VOT values indicates cross-language interference, consistent with SLM's prediction of L1 influence on L2 phonetic acquisition.

2. An Empirical Study of VOT Production for Initial Stops in English by Japanese EFL Learners

2.1 Purpose of the study

This study aims to conduct an acoustic analysis of English speech produced by Japanese learners of English, comparing VOT across different places of articulation and following vowels. It also seeks to examine how the learners' native language, Japanese, influences their L2 English production, with a focus on different proficiency levels.

2.2 Participants

The study included twenty Japanese EFL learners (ten male, ten female). Two American native English speakers (one male, one female) served as controls. The Japanese participants' English proficiency levels, as converted from their TOEIC scores to the Common European Framework of Reference

for Languages (CEFR), were distributed as follows: 10 participants at A2 level and 10 participants at B1 level.

2.3 Materials

The 18 target words consisted of one syllable, with initial voiceless /voiced stops, in three vowel contexts. The high front vowel context /i/ included the words "pea," "tee," "key" (voiceless) and "bee," "dee," "gee" (voiced). The high back vowel context /u/ comprised "pooh," "too," "coo" (voiceless) and "boo," "do," "goo" (voiced). The low back vowel context /ɑ/ included "par," "tar," "car" (voiceless) and "bar," "dar," "gar" (voiced). Table 7 lists the stimulus words used for measuring VOT in Japanese EFL learners. Words are categorized by their initial stops and following vowels.

Table 7

Stimuli Used in the VOT Experiment for Japanese EFL Leaeners

Initial Stops	Following Vowel	Word	IPA Transcription
/p/	/i:/	pea	/pi:/
/p/	/u:/	pooh	/pu:/
/p/	/ɑ/	par	/pɑr/
/t/	/i:/	tee	/ti:/
/t/	/u:/	too	/tu:/
/t/	/ɑ/	tar	/tɑr/
/k/	/i:/	key	/ki:/
/k/	/u:/	coo	/ku:/
/k/	/ɑ/	car	/kɑr/
/b/	/i:/	bee	/bi:/
/b/	/u:/	boo	/bu:/
/b/	/ɑ/	bar	/bɑr/
/d/	/i:/	dee	/di:/
/d/	/u:/	do	/du:/
/d/	/ɑ/	dar	/dɑr/
/g/	/i:/	gee	/gi:/
/g/	/u:/	goo	/gu:/
/g/	/ɑ/	gar	/gɑr/

2.4 Procedure

The target words were presented in random order, and participants were instructed to pronounce each word clearly. All utterances were recorded for subsequent acoustic analysis.

VOT measurements were conducted using Praat software (version 6.4.25; Boersma & Weenink, 2024), measuring from the release burst to the onset of periodic voicing in the following vowel (see Figure 1).

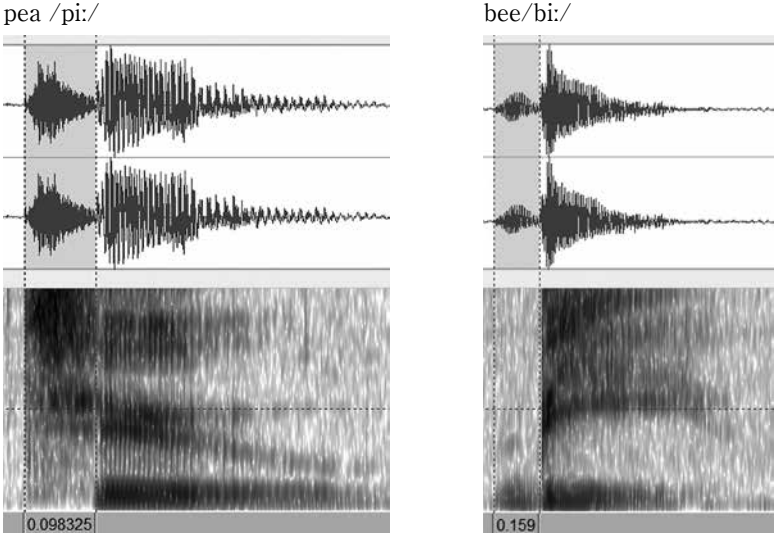

Figure 1

VOT measurement examples for "pea" /pi:/ and "bee" /bi:/ using Praat

2.5 Results

2.5.1 Descriptive Statistics of VOT by Initial Stops among Japanese EFL Learners

All statistical analyses were performed using R statistical software (Version 4.4.1; R Core Team, 2024). VOT measurements were analyzed for six initial stops (/b/, /d/, /g/, /p/, /t/, /k/) produced by Japanese EFL learners. Table 8 shows the mean VOT values and standard deviations (in seconds). Figure 2 illustrates the VOT distribution across the six initial stops.

Table 8

Descriptive Statistics for VOT Across Initial Stops among Japanese EFL Learners

	N	VOT	SD
/b/	60	-0.0222	0.0238
/d/	60	-0.0203	0.0184
/g/	60	-0.0345	0.0354
/p/	60	0.0633	0.0251
/t/	60	0.0775	0.0256
/k/	60	0.0985	0.0273

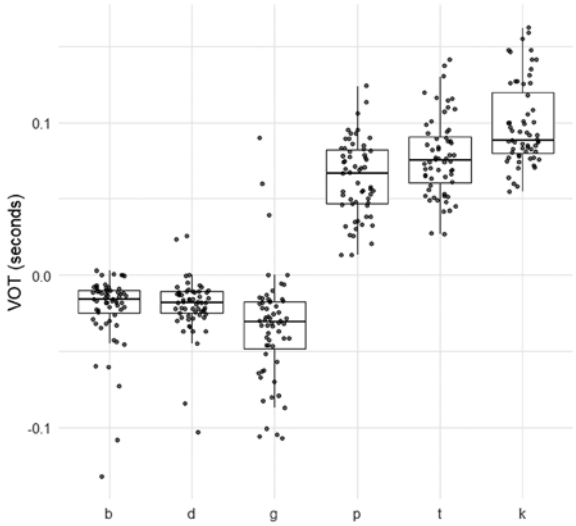

Figure 2

VOT Distribution Across Initial Stops among Japanese EFL Learners

One-way ANOVA revealed significant main effects of initial stops on VOT, $F(5, 354) = 299.20$, $p < .001$, indicating systematic differences across the six phonemes. Post-hoc comparisons using Tukey's HSD tests showed that voiceless stops (/p/, /t/, /k/) consistently had significantly longer VOTs than their voiced counterparts (/b/, /d/, /g/) (*ps* < .001). Notably, /k/ exhibited significantly longer VOTs compared to other phonemes (*ps* < .001). The findings confirm systematic distinctions in VOT among the initial stops, consistent with prior research on stop consonants in second language acquisition.

2.5.2 Descriptive Statistics of VOT across Vowel Contexts among Japanese EFL Learners

This section focuses exclusively on the VOT data for voiceless stops (/p/, /t/, /k/). This is because previous research (Shimizu, 1996) has demonstrated that the height of the following vowel affects the VOT in voiceless stops more than in voiced stops. VOT measurements were analyzed across three vowel contexts (/i/, /u/, /ɑ/). Table 9 shows the mean VOT values and standard deviations (in seconds). Figure 3 illustrates the VOT distribution across the three vowel contexts.

Table 9

Descriptive Statistics for VOT Across Vowel Contexts Among Japanese EFL Learners

	N	VOT	SD
/i/	60	0.0812	0.0319
/u/	60	0.0826	0.0235
/ɑ/	60	0.0756	0.0329

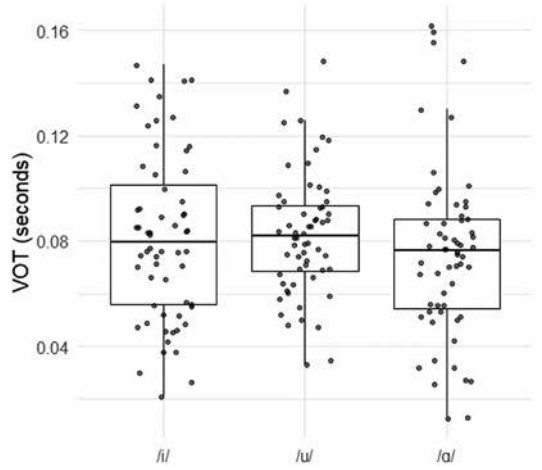

Figure 3

VOT Distribution Across vowel contexts among Japanese EFL Learners

A one-way ANOVA revealed no significant main effect of vowel context on VOT, $F(2, 177) = 0.93$, $p = .396$, $\eta^2 = 0.0104$, indicating that VOT did not differ significantly across the three vowel contexts.

2.5.3 Descriptive Statistics of VOT for Initial Stops Across Groups

VOT measurements were analyzed across three groups (A2, B1, and American) for six initial stops (/b/, /d/, /g/, /p/, /t/, /k/). Below are the mean VOT values (in seconds), with standard deviations provided in parentheses for English stops (see Table 10). Figure 4 illustrates the VOT distribution across groups and initial stops.

Table 10

Descriptive Statistics for VOT Across Groups and Initial Stops

	A2	B1	American
/b/	-0.0212 (0.0213)	-0.0233 (0.0264)	-0.0787(0.0716)
/d/	-0.0159 (0.0145)	-0.0247 (0.0210)	-0.1260(0.0571)
/g/	-0.0368 (0.0332)	-0.0322 (0.0380)	-0.1270(0.0678)
/p/	0.0680 (0.0185)	0.0586 (0.0299)	0.1210(0.0213)
/t/	0.0782 (0.0204)	0.0769 (0.0303)	0.1340(0.0182)
/k/	0.0968 (0.0272)	0.1000 (0.0278)	0.1440(0.0255)

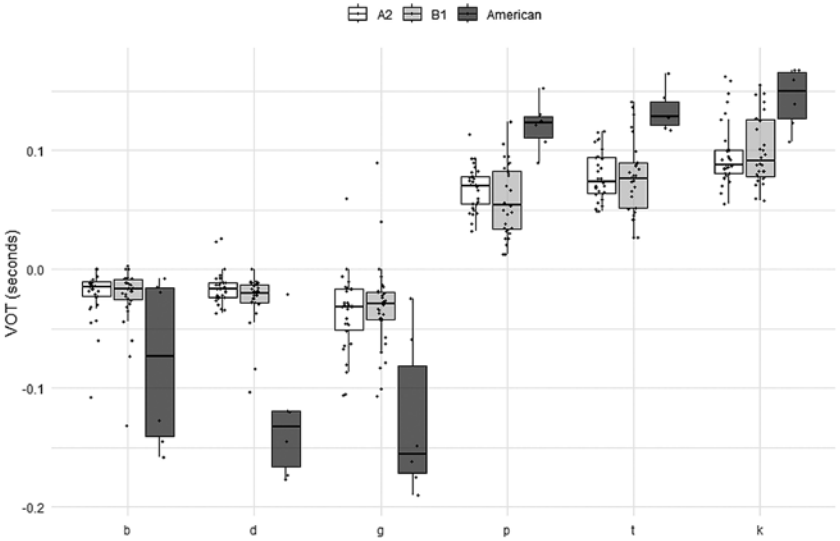

Figure 4

VOT Distribution Across Groups and Initial Stops

One-way ANOVAs revealed significant main effects of group for all initial stops. Specifically, for voiced stops, results showed /b/: $F(2, 63) = 9.31$, $p < .001$, $\eta^2 = .23$; /d/: $F(2, 63) = 55.55$, $p < .001$, $\eta^2 = .64$; and /g/: $F(2, 63) = 15.19$, $p < .001$, $\eta^2 = .33$. For voiceless stops, significant effects were observed for /p/: $F(2, 63) = 16.21$, $p < .001$, $\eta^2 = .34$; /t/: $F(2, 63) = 13.48$, $p < .001$, $\eta^2 = .30$; and /k/: $F(2, 63) = 7.70$, $p < .01$, $\eta^2 = .20$. As post-hoc comparisons, Tukey's HSD tests revealed significant differences between native speakers and both learner groups for all consonants ($ps < .001$). For voiced stops, American speakers exhibited significantly more negative VOT values than both learner groups ($ps < .001$). For voiceless stops, American speakers demonstrated significantly longer positive VOT values ($ps < .001$). No significant differences were found between the A2 and B1 groups for any consonants ($ps > .05$).

2.5.4 Descriptive Statistics of VOT Across Groups and Vowel Contexts

This section focuses exclusively on the VOT data for voiceless stops (/p/, /t/, /k/). VOT measurements were analyzed across three groups (A2, B1, and American) in three vowel contexts (/i/, /u/, /ɑ/). Below are the mean VOT values (in seconds), with standard deviations provided in parentheses (see Table 11). Figure 5 illustrates the VOT distribution across groups and following vowels.

Table 11

Descriptive Statistics for VOT Across Groups and Vowel Contexts

	A2	B1	American
/i/	0.080 (0.025)	0.082 (0.038)	0.134 (0.018)
/u/	0.085 (0.022)	0.080 (0.025)	0.138 (0.025)
/ɑ/	0.078 (0.028)	0.073 (0.037)	0.127 (0.028)

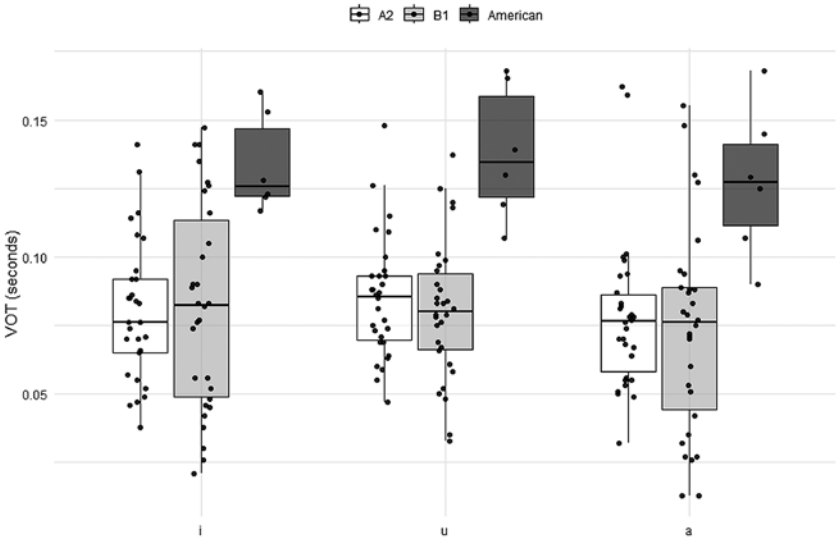

Figure 5
VOT Distribution Across Groups and Vowel Contexts

One-way ANOVAs revealed significant main effects of group for all vowel contexts. Specifically, results showed /i/: $F(2, 63) = 7.77$, $p < .001$, $\eta^2 = .20$; /u/: $F(2, 63) = 15.23$, $p < .001$, $\eta^2 = .33$; and /ɑ/: $F(2, 63) = 7.03$, $p < .01$, $\eta^2 = .18$. As post-hoc comparisons, Tukey's HSD tests revealed significant differences between American speakers and both learner groups for all vowel contexts (ps < .01). For /u/, American speakers exhibited significantly longer VOT values, with the largest differences observed (ps < .001). For /i/, similar patterns were observed (ps < .01). For/ɑ/, American speakers also produced significantly longer VOT values compared to A2 and B1 learners (ps < .01). No significant differences were found between the A2 and B1 groups for any vowel contexts (ps > .05).

2.6 Conclusions and Discussions

The findings and implications regarding the general VOT characteristics of Japanese EFL learners, as revealed by this study, are summarized as follows. (1) Voiceless stops (/p/, /t/, /k/) exhibited significantly longer VOTs compared to voiced stops (/b/, /d/, /g/) in the productions of Japanese EFL learners. Notably, the VOT for /k/ was significantly longer than for /p/ and

/t/. These findings support the universal phonetic principles described by Lisker and Abramson (1964), which state that voiceless stops are typically produced with a burst of aspiration, whereas voiced stops involve shorter or even negative VOT values. Furthermore, velar stops consistently showed longer VOT values than their labial or dental/alveolar counterparts, a pattern also observed in Japanese (Shimizu, 1996). Shimizu reported that the VOT of voiced stops in Japanese is negative, whereas the VOT of voiceless stops is positive. Specifically, the VOT of the voiceless velar stop /k/ is the longest among consonants. These findings suggest that the VOT patterns of Japanese EFL learners reflect characteristics of their native language.

(2) There were no statistically significant differences in VOT among the three vowel contexts (/i/, /u/, /ɑ/) in the productions of Japanese EFL learners. This lack of significant variation contrasts with previous studies on Japanese speakers, where vowel height has been reported to influence VOT durations (Shimizu, 1996). The discrepancy suggests that Japanese EFL learners may not yet exhibit systematic VOT variations based on vowel contexts, possibly due to insufficient L2-specific phonetic training.

The findings and implications derived from the comparison of VOT between Japanese EFL learners and native English speakers are summarized as follows.

(3) For voiceless stops, American English speakers demonstrated significantly longer positive VOT values compared to Japanese EFL learners. In contrast, no significant differences in VOT were observed between the A2 and B1 learner groups. As noted by Shimizu (2010), Japanese stops generally exhibit much shorter VOT values than those in English. This suggests that the shorter VOT values of voiceless stops produced by Japanese learners likely reflect the phonetic characteristics of their native language. Consequently, it may be concluded that VOT production does not automatically improve with general language proficiency advancement.

(4) American English speakers consistently produced longer VOTs across all vowel contexts compared to Japanese EFL learners. Among the vowel contexts, the high-back vowel /u/ showed the largest effect size ($\eta^2 = .33$), indicating that this vowel environment presents the greatest challenge for Japanese learners in terms of VOT production. The English vowel /u/ involves a complex articulatory gesture, characterized by high tongue position, tongue retraction, and lip rounding. The articulatory challenge lies in

the need for the tongue to rapidly transition to a high position following the release of the stop. For Japanese learners, whose native vowel /ɯ/ is slightly more central and lacks lip rounding, controlling such articulatory movements may be particularly difficult.

References

Boersma, P., & Weenink, D. (2024). Praat: Doing phonetics by computer (Version 6.4.25) [Computer software]. Phonetic Sciences, University of Amsterdam. https://www.fon.hum.uva.nl/praat/

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207–229.

Flege, J. E., & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics, 15*(1), 67–83.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America, 89*(1), 395–411.

Harada, T. (2007). The production of voice onset time (VOT) by English-speaking children in a Japanese immersion program. *International Review of Applied Linguistics in Language Teaching*, *45*(4), 353–378.

Homma, Y. (2006). *Nichieigo-no onkyo onseigaku [Acoustic phonetics in English and Japanese]*. Kyoto: Yamaguchi Shoten.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*(3), 384–422.

R Core Team. (2024). R: A language and environment for statistical computing (Version 4.4.1) [Computer software]. R Foundation for Statistical Computing. https://www.r-project.org

Shimizu, K. (1996). *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Tokyo, Japan: Seibido Publishing Co.

Shimizu, K. (2010). Acoustic analysis of English and Japanese stop voicing contrasts produced by Korean L2 learners. *Nagoya Gakuin University Journal of Language and Culture, 22*(1), 1–10.