Rubén PÉREZ-RAMÓN & Mariko KONDO (2022). Intelligibility of Japanese accented segments of English. In Y. Ono & M. Shimada (Eds.) *Data Science in Collaboration, Volume 5* (pp. 1–8). Tsukuba: inext Co., Ltd.

Intelligibility of Japanese accented segments of English

Rubén PÉREZ-RAMÓN[a*] & Mariko KONDO[a]
[a]*Waseda University, Japan*
*rperez.ram@gmail.com

Abstract: In this paper we present a study of segmental foreign accent, as opposed to the more traditional holistic approach to the matter of non-native speech. Particularly, we want to investigate how Japanese accented individual segments in English words affect the intelligibility of Japanese listeners ranging a number of degrees of proficiency in English. Results show that, overall, there is a significant difference in intelligibility between accented and unaccented tokens, but this difference seems to be more attributable to consonants than vowels. Additionally, the disambiguation of proficiency groups reveals that, as expected, low proficiency listeners present a lower degree of intelligibility than high-proficiency groups, but the difference in intelligibility between accented and unaccented tokens does not seem to increase with lower competence in the target language.

Keywords: foreign accent, perception, consonants, vowels, intelligibility

1. Introduction

A foreign accent (FA) arises when a non-native speaker pronounces a target language with noticeable changes in its segmental or suprasegmental features due to the interference of his or her first language (L1) (Arslan, 1997; Munro & Derwing, 1995). FA presents a speech pattern inconsistent with that of native speakers (Porretta, 2015) that may affect the perception in a range of factors such as intelligibility, comprehensibility (Munro & Derwing, 1995) or cognitive load (Jensen & Thøgersen, 2017).

Traditionally, foreign accent has received a holistic treatment on behalf of the research community, i.e., foreign accented speech has been studied as a whole (Bent & Bradlow, 2003 among others); however, some studies have focused on individual features of foreign accent such as speech rate (Munro & Derwing, 2001), duration (Tajima et al., 1997), intonation (Bengrait, 2020) or nuclear stress (Hahn, 2004). Lately, some studies have also tackled the issue of foreign accent by focusing on the production and perception of segments (Pérez-Ramón 2020; Henter, 2018; Horslund, 2021 among others).

In this study we want to analyze how mispronounced individual segments in English words affect their intelligibility. Particularly, by using acoustic manipulation, a target segment will be replaced in an English word by its most common Japanese realization. Afterwards, intelligibility in accented and unaccented tokens by Japanese listeners with a different range of English proficiency will be measured. Our first research question, therefore, would be: Are Japanese accented English words as intelligible for Japanese listeners as non-accented words?

Additionally, the aim of this type of segmental manipulation is to check for more specific cues, so our second research question concerns intelligibility differences in perceived consonants and vowels. In previous research (Pérez-Ramón 2020) it has been found that consonants and vowels entail different degrees of perceived foreign accent. We want to check in this study if there is also a difference in intelligibility, so, our second research question is: Is intelligibility dependent on the type of segment?

Finally, as previously mentioned, we will look for differences according to the proficiency in English of the listener. It is known that proficiency in the target language can modify the perception of foreign-accented cues (Thordadrottir, 2021), so our third research question would be: Is intelligibility dependent on the proficiency of the listener in English?

2. Methods

2.1 Stimuli

The generation of the experimental stimuli was performed following the splicing technique, described in (García Lecumberri, 2014) and summarized here. 12 English segments (7 consonants and 5 vowels) were selected as the target segments for this experiment. For each segment, its most common Japanese accented counterpart was identified (Table 1). A bilingual speaker with no identifiable accent either in Japanese or American English was recruited following the methodology depicted in (Pérez-Ramón, 2021) to record three English words with CVC structure for each of the target segments and their equivalent Japanese segments inserted in non-words with a similar phonetic context (e.g., for the English word *vet* [vet], the Japanese nonword *bete* [bete] was recorded, so the target [v] and the accented counterpart [b] segments were in the same [_e] position). The three words were selected according to their frequency in the English language by consulting the Google frequency corpus (Kaufman, J., 2021). Afterwards, the target segment was removed from the word via acoustic manipulation using Praat (Praat, 2021) and the Japanese segment was spliced to the rest of the word with an overlap varying between 0.1ms and 0.5ms depending on the nature of the segment (e.g., plosives demand a shorter overlap than fricatives or vowels). This way, for each English word (e.g., *vet* [vet]), a Japanese accented version was generated in which only the target segment was accented, while the rest of the word remained with the original English pronunciation (e.g., [bet]). In order to avoid a bias towards manipulated tokens, the native-like word was also generated using the splicing technique, replacing the original segment with the same segment from a different utterance of the same word.

Table 1
Target segments, Japanese accented counterparts and selected words

Native segment	Foreign segment	Words		
æ	a	half	dad	laugh
ʌ	a	luck	cup	much
ɝ:	a:	birth	hurt	work
ɪ	i	ship	kid	thick
ɑ:	o	shop	hop	top
f	ɸ	feet	fish	food
k^h	k	keep	come	call
t^h	t	tag	teach	took
l	ɾ	leave	league	long
ɹ	l	run	reach	roof
v	b	vet	van	ville
θ	s	thick	thought	thumb

2.2 Listeners

51 participants were recruited using the student platform at Waseda University. The requirements were that their first language was Japanese and that they were not bilingual in any language. They were divided in 4 groups according to their proficiency in English, from no/low proficiency to high proficiency (proficiency in English was determined by their scores in official English exams, e.g., TOEFL, TOEIC...). Background of participants can be seen in Table 2. Groups were named according to the standardization of their English proficiency according to the Common European Framework of Reference (Council of Europe, 2001).

Table 2
Demographic background of participants

	Group A1	Group B1	Group B2	Group C1
n (female)	4 (3)	18 (15)	22 (15)	7 (3)
Age	15-20: 2 21-30: 2	15-20: 7 21-30: 11	15-20: 11 21-30: 10 41-50: 1	15-20: 5 21-30: 2
Other languages (proficiency)	Chinese (int): 3	Chinese (int): 3	Korean (int): 3	Spanish (int): 1 Chinese (int): 1
Musical knowledge (n)	2	13	12	3

2.3 Task

Participants were asked to complete an intelligibility task in which they were asked to write down the word they heard. Participants were only allowed to hear individual stimuli once. On completing one token, the following one would play automatically. Tokens were grouped into two blocks by type of target segment (consonants or vowels) and randomized within each block. Each stimulus was presented twice. The order of presentation of blocks was randomized and counterbalanced across participants. Three versions of this task were generated, each one with two out of the three words; the version of the task presented was randomized and counterbalanced across participants. The final task, therefore, consisted of 96 stimuli (2 accents × 12 sounds × 2 words × 2 repetitions).

The experiment was designed using the Gorilla experiment design tool (Anwyl-Irvine et al., 2019). A personalized link was sent to each participant to complete the assignment. Participants were allowed to complete the experiment at home using their own devices, but were asked to do so in a quiet environment and using headphones. All 51 participants completed the task using a laptop or desktop computer with the exception of 2 participants, who used their smartphone devices.

3. Results

The analysis of the results was conducted using RStudio (R Core Team, 2020). Following Rogers & Davis (2009), 230 responses whose reaction time took less than 300ms or more than 5000ms were removed from the dataset.

3.1 Overall results

Figure 1
Overall percentage of intelligibility in foreign accented and native stimuli. Error bars represent ± 1 standard error. For a better visualization, y-axis shows 50%-100% intelligibility.

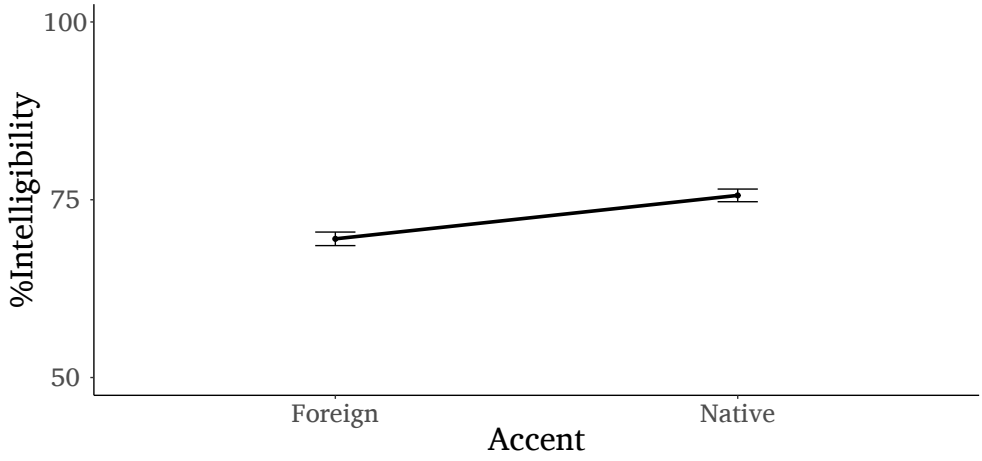
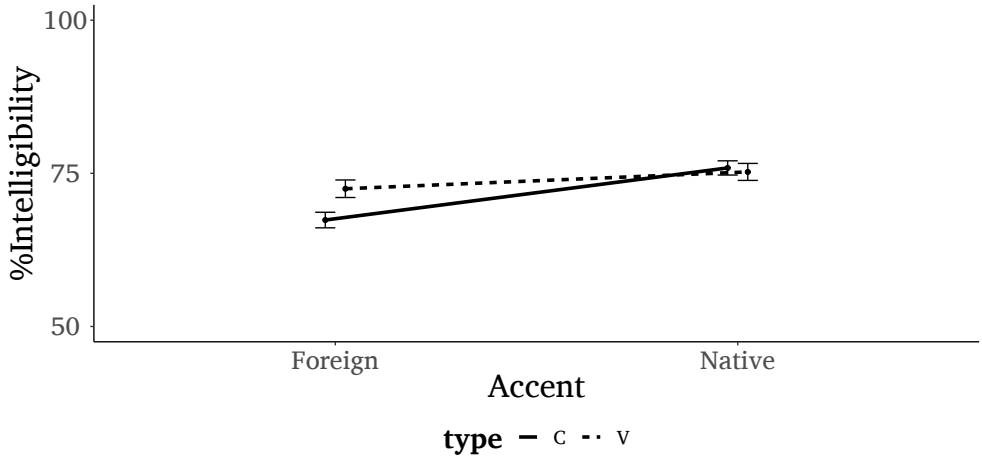


Table 3
Mixed model results for main effect of accent for the intelligibility results

Model: %intelligibility ~ accent + (1 | participant)				
Fixed effects	β	Error	t	p
Intercept (foreign accent)	0.84	0.06	14.05	< .001
native accent	0.31	0.07	4.7	< .001
Random effects	SD			
Intercept (participant)	0.28			

Overall, as expected, foreign accented tokens were less intelligible than native-like tokens (Figure 1). An important remark to make is that native-like tokens were correctly perceived only around a 75% of the times, far from the 100% expected in non-accented words. A general linear mixed model was generated in R using the package *lme4* (Bates et al., 2015) to check for significant differences between accented and non-accented tokens with individual participants as standard error, and a significant difference [$p < .001$] was found (Table 3).

Figure 2
Percentage of intelligibility in foreign accented and native stimuli by type of segment (consonant or vowel). Error bars represent ± 1 standard error. For a better visualization, y-axis shows 50%-100% intelligibility. Results for each condition have been slightly jittered to avoid over-plotting.



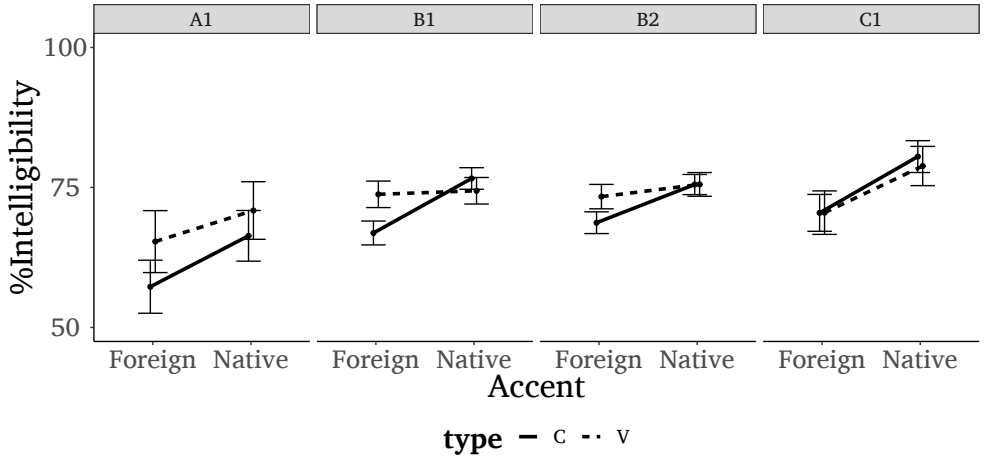
An analysis was carried out to check for differences in the results of consonants and vowels (Figure 2). This analysis revealed that a foreign accented vowel has less effect in intelligibility than a consonant. A general linear mixed model with accent and type of segment as fixed effects and participants as random effect threw a significant result for the accent factor [$p < .001$] and for the interaction of accent with type of segment [$p < .05$]. Specifically, according to a post-hoc pairwise comparison analysis

performed via the *emmeans* command in R (Lenth, 2021), an accented segment resulted in a significant loss of intelligibility for consonants, but not for vowels. In concordance with this result, a significant difference was also found between words with an accented consonant and an accented vowel, but not between native-like tokens.

Table 4
Mixed model results for main effects of accent and type of segment for the intelligibility results

Model: %intelligibility ~ accent × type of segment + (1 | participant)				
Fixed effects	β	Error	t	p
Intercept (foreign accent:consonant)	0.74	0.07	10.55	p < .001
native accent	0.43	0.09	4.94	p < .001
vowel	0.24	0.09	2.65	p < .01
native accent:vowel	-0.28	0.13	-2.08	p < .05
Random effects	SD			
Intercept (participant)	0.28			

Figure 3
Percentage of intelligibility in foreign accented and native stimuli by type of segment (consonant or vowel) for the four proficiency groups. Error bars represent ± 1 standard error. For a better visualization, y-axis shows 50%-100% intelligibility. Results for each condition have been slightly jittered to avoid over-plotting.



Finally, an analysis was carried out to check for differences between proficiency groups. Results show that, generally, all groups suffered a more pronounced loss in intelligibility when an accented consonant was inserted compared to a vowel, but this effect was less discernable for the high proficiency group. A general linear mixed model was generated with accent, type of segment and proficiency group as fixed factors and participants as random factor (Table 5). The model revealed a significant effect of accent [p < .001] and a significant interaction between accent and type of segment [p < .05]. Only marginal significance was found for type of segment [p < .1] and proficiency [p < .1]. A post-hoc *emmeans* analysis revealed that there was a significant difference in intelligibility for accented and unaccented consonants only for groups B1 and B2, but not for A1 or C1. No group showed lower intelligibility when the accented segment was a vowel. The marginal difference in proficiency arose from a similarly marginal difference between group A1 and groups B2 [p < .1] and C1 [p < .1].

Table 5
Mixed model results for main effects of accent, type of segment and proficiency for the intelligibility results. For a better visualization, only the Intercept and the significant factors are presented.

Model: %intelligibility ~ accent × type of segment × proficiency + (1 | participant)				
Fixed effects	β	Error	t	p
Intercept (foreign accent:consonant:A1)	0.29	0.23	1.29	
proficiency B2	0.5	0.25	2	$p < .05$
proficiency C1	0.58	0.29	1.99	$p < .05$
Random effects	SD			
Intercept (participant)	0.24			

4. Discussion

In this study, the issue of intelligibility of English words with a Japanese accented segment by Japanese native listeners differing in their degree of proficiency in English has been presented. Results show that, as seen in previous studies with different combinations of target language/foreign accent (Pérez-Ramón, 2020), a holistic analysis can mask crucial particularities of individual cues like, in this case, mispronounced types of segments. Furthermore, it has been shown that foreign accent affects intelligibility in a different way according to the proficiency in the target language of the listeners.

The first research question of this study concerned intelligibility in a general sense. We wanted to know if a single accented segment would motivate a decrease in intelligibility. Results show that yes, there was a significant detriment in intelligibility when a foreign accented segment was inserted. The fact that even foreign accented tokens were perceived as highly intelligible, with correct answers ~70% of the times, goes in accordance with previous research (Munro & Derwing, 1995; van Maastrich, 2016) that shows that a strong foreign accent do not always implies a similarly strong loss in intelligibility. However, the fact that the native tokens were correctly perceived only around a 75% of the times suggests that some underlaying process is taking place and preventing a higher score.

Our hypothesis is that a deeper inspection of two factors, namely type of segment and listener's proficiency in English, could throw some insights regarding this issue. This is why the second research question of this study looks deeper into the differences between foreign accented vowels and consonants. The results of the intelligibility task disambiguated by type of segment revealed that the loss in intelligibility was more notable with a mispronounced consonant than with a vowel. Some studies have found that Japanese listeners are generally more sensitive to contrasts between English vowels than consonants (Flege et al., 1998), which means that they can more easily detect a mispronounced English vowel than a consonant. Our results support this idea in the sense that all groups were generally better in correctly identify words in which a vowel was accented than a consonant. It is also known that Japanese speakers make more vowel errors in English (Kondo et al., 2015), meaning that they are deaf to the differences of certain vowels (i.e., some vowels that are classified as different by native listeners fall within the same vocalic category for Japanese listeners). This finding implies that Japanese listeners are less able to judge intelligibility of vowels.

Finally, the matter of proficiency in the target language was also tackled. It is necessary to address here a limitation of this study: the fact that proficiency groups were not balanced (i.e., more participants in B1 and B2 groups than in A1 and C1) may be masking relevant information in this issue. However, we think that the results obtained are a good indicative of the trend to be expected in a more balanced situation. Particularly, the low proficient group shown signs of less intelligibility than high proficiency groups, but the decrease in intelligibility in accented tokens as compared to the native ones does not seem to be steeper; in other words, a foreign accented segment seems to affect to a similar degree whatever the starting point is. Furthermore, the significant difference between accented consonants and vowels looks to decrease when the listener is more proficient in English, i.e., mispronounced vowels are more disruptive for native-like listeners than for non-native listeners. This finding goes in line with

the so-called *matched interlanguage intelligibility benefit* (Bent and Bradlow, 2003) in the sense that Japanese listeners would accuse less their own accent than native listeners. It is expected that proficient listeners, who are more tuned to the segmental cues of English, present a pattern more similar to that of native listeners.

5. Conclusion

The results of this study add up to the idea that holistic studies can mask the particularities of foreign accent, i.e., the individual cues from which foreign accent arises. Particularly, results suggest that a Japanese accented consonant in an English word is more detrimental in terms of intelligibility than a vowel, but this distinction seems to vary according to the listener's proficiency in English. However, further research with more balanced groups and including a native control group and listeners with other languages as L1 is required to extract more robust conclusions.

Acknowledgements

This research is supported by Grant-in-Aid for JSPS Fellows, No. 20F40079

References

Arslan, L; Hansen, J. (1997). A study of temporal features and frequency characteristics in American English foreign accent. *Journal of the Acoustical Society of America* 102(1), 28-40.

Anwyl-Irvine, A.L., Massonié J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods*.

Bengrait, N. (2020). Analysis of Algerian Students English Intonation Deviations with the Employment of Speech Analyzer and Praat Programmes. *Journal of Studies in Language, Culture and Society*, *3*(1), 27-65.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600-1610.

Boersma, P. & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.2.08, retrieved 5 February 2022 from http://www.praat.org/.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.

Derwing, T., & Munro, M. J. (2001). What speaking rates do non-native listeners prefer? *Applied linguistics*, *22*(3), 324-337.

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.

Flege, J. E., Guion, S., Akahane-Yamada, R., & Downs-Pruitt, J. (1998, June). Categorial discrimination of English and Japanese vowels and consonants by native Japanese and English subjects. In *Proceedings of the 16th International Congress on Acoustics and the 135th meeting of the Acoustical Society of America* (Vol. 4, pp. 2973-2974).

García Lecumberri, M. L., Barra Chicote, R., Pérez Ramón, R., Yamagishi, J., & Cooke, M. (2014). Generating segmental foreign accent.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL quarterly*, *38*(2), 201-223.

Henter, G. E., Lorenzo-Trueba, J., Wang, X., Kondo, M., & Yamagishi, J. (2018, April). Cyborg speech: Deep multilingual speech synthesis for generating segmental foreign accent with natural prosody. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4799-4803). IEEE.

Horslund, C. S., & Van Nostrand, P. F. (2021). Segmental error patterns in Finnish-accented English. *European Journal of Applied Linguistics*.

Jensen, C., & Thøgersen, J. (2017). Foreign accent, cognitive load and intelligibility of EMI lectures. *Nordic Journal of English Studies*, *16*(3), 107-137.

Kaufman, J. (2021). List of the 10,000 most common English words in order of frequency. https://github.com/first20hours/google-10000-english.

Kondo, M., Tsubaki, H., & Sagisaka, Y. (2015). Segmental variation of Japanese speakers' English: Analysis of "the North Wind and the Sun" in AESOP Corpus (< Feature Articles> Corpus-Based Data-Driven Speech Research in Asia). *Journal of the Phonetic Society of Japan*, *19*(1), 3-17.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, *45*(1), 73-97.
Pérez-Ramón, R., Cooke, M., & Lecumberri, M. L. G. (2020). Is segmental foreign accent perceived categorically? *Speech Communication*, *117*, 28-37.
Pérez-Ramón, R. & Kondo, M. (2021). Relationship between reaction time and nativeness rating. *Journal of the Phonetic Society of Japan*, in press.
Porretta, V. (2015). Perception and lexical processing of gradient foreign accentedness [Doctoral thesis]. University of Alberta.
R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Rogers, J., & Davis, M. (2009). Categorical perception of speech without stimulus repetition. *Proceedings of Interspeech, Brighton,* 376-379.
Russell V. Lenth (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.4. https://CRAN.R-project.org/package=emmeans.
Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, *25*(1), 1-24.
Thordardottir, E., & Hávarðarson, G. (2021). Accent, Comprehensibility, and Confidence in Second and Third Languages of Adolescents: How Do They Relate to Language Proficiency and Language Interactions? *Journal of Speech, Language, and Hearing Research*, *64*(12), 4828-4844.
van Maastricht, L., Krahmer, E., Swerts, M. (2016). Native speaker perceptions of (non-)native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication 83*, 21-33.