

# Rating nonnativeness in L1-Japanese L2-Arabic Speakers' Vowels

Yahya Aldholmi, Sara Alotaibi, Malak Alrouqi

Department of Linguistics & Translation Studies, King Saud University, Saudi Arabia

<https://doi.org/10.36505/ExLing-2021/12/0003/000476>

## Abstract

This experiment investigates whether Arabic native speakers can detect nonnativeness (i.e., accentedness) in similar L1-Japanese L2-Arabic short and long vowels. We asked two native speakers of Arabic and two Japanese learners of Arabic as a foreign language to record 48 Arabic CVCVC vs. CVCV:C word pairs at a normal speech rate. We then isolated the second vowel from each pair and instructed 33 NSAs to rate, on a 1-7 scale (1= certainly native, 7= certainly nonnative), the nonnativeness of the target vowels. Results show that, irrespective of L1-L2 segment similarity, the raters were successfully able to detect accentedness in the nonnative stimuli with no overall difference in the nonnativeness ratings for short vs. long vowels.

Keywords: duration, accentedness detection, nonnativeness rating, Japanese, Arabic

## Introduction

Previous research assessing a listener's capacity to successfully detect accentedness/nonnativeness in L2 production (e.g., Flege 1984; Flege, Munro, & MacKay, 1995; Mackay, Flege, & Imai, 2006; Park, 2013; Almohareb, 2020; Zhi & Li, 2021) has utilized experimental materials ranging from entire passages to complete sentences, fragments, whole or partial words, syllables, segments, and even portions of segments. A handful of such studies have indicated duration as a critical factor contributing to the perceived degree of nonnativeness and provide evidence that longer materials, as common wisdom suggests, supply more information to listeners attempting to detect or rate nonnativeness. Stimulus duration has been examined collectively with other factors, for example level of proficiency, and previous results show that the detection of nonnativeness in a minimal utterance produced by highly proficient L2 learners is most reliable when using stimuli containing segments that are dissimilar to those in the learners' L1 (e.g., Park, 2013). Taking into consideration the role of both *stimulus duration* and *segment similarity* in rating nonnativeness, we examine whether native speakers of Arabic (NSAs) are able to detect nonnativeness in the production of three short and long vowels ([a-a:], [i-i:], and [u-u:]), each of which has a similar version in the counterpart language) by Japanese learners of L2 Arabic (JLAs). We also ask whether long vowels can provide more information than short ones to the NSAs making these judgments.

## Methodology

### Stimuli

In pursuit of the study objective, we carefully selected 48 Arabic CVCVC vs. CVCV:C minimal or nearly minimal pairs in which the target V2 was varied ([a-a:], [i-i:], and [u-u:]). The target vowels, which exist both in Arabic and Japanese, were always neighbored with obstruents, as in /nabat/ “grew” vs. /naba:t/ “plant”, in order to reduce the impact of consonantal information on nonnativeness ratings, although most of the consonants that appeared adjacent to the target vowels exist in both languages. We instructed two (one male and one female) NSAs and two (one male and one female) JLAs to produce the list at a normal speech rate, using the carrier phrase /qul-tu/ “I said”, similar to phrases such as /katab-tu (ʔajd<sup>s</sup>-an)/ “I wrote (as well)” used in previous studies (e.g., Tsukada, 2010). We then extracted the target vowels from each pair and placed each vowel between two beep tones of a lower amplitude.

### Procedure

Thirty-three NSAs (age  $M= 27.7$ ,  $SD= 4.7$ ) were instructed to rate, on a 1-7 scale (1= certainly native, 7= certainly nonnative), the (non)nativeness of each vowel. The experiment was self-paced and the participants were allowed to listen to each trial as many times as they wished. For familiarization purposes, each participant was provided with a few practice examples, which were excluded from the actual experiment, before they commenced the task.

## Results

The overarching results indicate that while nonnativeness ratings for long vs. short vowels in the nonnative stimuli were similar ( $Median= 4$  for both) in general, the raters were successfully able to detect nonnativeness in the nonnative stimuli overall as demonstrated by a difference in their ratings ( $Median= 4$  and  $2$  for vowels produced by JLAs and NSAs, respectively). These results are supported by the main effect outcome of a repeated measures regression model in which both *stimulus (non)nativeness* and *stimuli duration* were fit as predictor factors and with *nonnativeness ratings* as a predicted factor, Wald  $\chi^2(1) = 37.8$ ,  $p < .001$ , and by the interaction effects between the two predictor factors, Wald  $\chi^2(1) = 16.4$ ,  $p < .001$ .

As exhibited in Figure 1, although most of the ratings for each condition are relatively evenly distributed over the entire scale, there are two outstanding patterns. First, the participants rated approximately 52% of the native long vowels as *certainly native* and 18% of the vowels from the same condition as *native*; that is, roughly 70% of the native long vowels were rated as (*certainly*) *native*. Second, short vowels were responded to with the *uncertain* rate (28% for nonnative short vowels and 22% for native short vowels) more than the long vowels were (13% for nonnative long vowels and 8% for native long vowels).

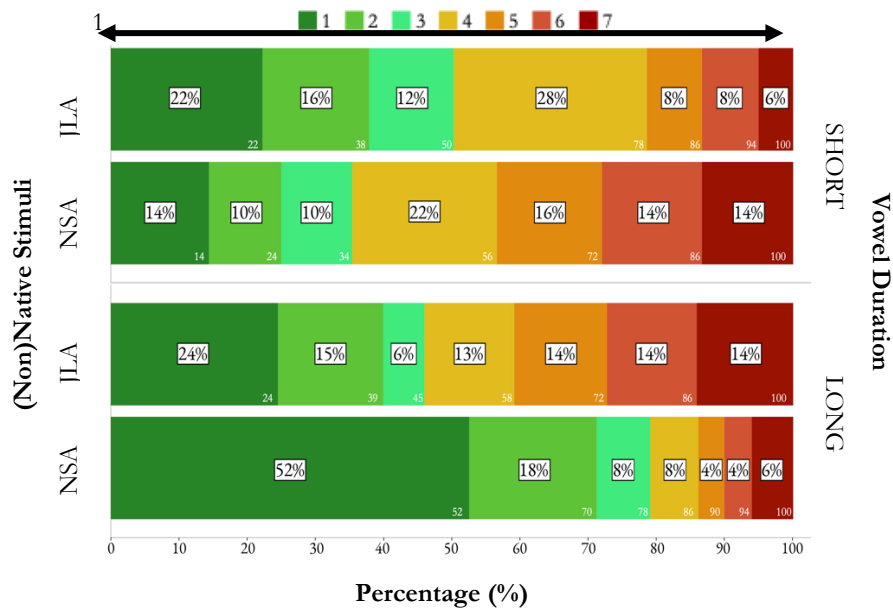


Figure 1. Percentages of (non)nativeness ratings for NSAs vs. JLAs' short vs. long vowels

## Discussion and conclusion

This study set out with two main related questions. The first probes whether NSAs can detect nonnativeness in three Arabic short vs. long vowels produced by JLAs. Although Japanese has contrastive duration and the vowels under investigation exist in the Japanese phonological repository, the current findings show that accentedness/nonnativeness can be still detected in L1-L2 similar segments. The second question asks whether long vowels facilitate nonnativeness detection more than short vowels. The findings reveal that both short and long vowels provide roughly the same magnitude of information in this respect.

Three pertinent points must be noted here. First, most previous studies that used a whole or partial segment to rate accentedness utilized consonants rather than vowels. For instance, Flege (1984) found that accentedness can be detected from a portion of a segment (30 ms of a /t/ burst). Harada (2006) instructed his Japanese raters to evaluate accentedness, focusing only on the bilinguals' singleton-geminate distinction and found out that the durational difference between the two categories was not sufficient to sound native-like. The current findings add to the body of that research; vowels that exist in both L1 and L2 provide enough information for raters to detect nonnativeness.

Second, research on the durational short-to-long ratio in Arabic vs. Japanese vowels has shown both similarities and dissimilarities. For example, Tsukada (2009) states that the short-long ratio in Arabic and Japanese is not substantial (0.43 and 0.39 ms, respectively), while other studies (e.g., Tsurutani, Tsukada, & Ishihara, 2010) state that unlike Arabic vowels, Japanese short vowels are shorter than 50% of the duration of their long counterparts. Since accentedness was detected in both short and long vowels in the current study, it is likely that the learners were unable to maintain the durational ratio for both short and long vowels (i.e., duration was not Arabic-like in both types), which resulted in accentedness/nonnativeness in production and consequently in perception.

Third, despite careful attention to controlling the design of this study, there remain some limitations. In particular, this study relies on findings from previous research. A more careful study would recruit as many native and nonnative Arabic and Japanese talkers as possible, measure the durational ratio, and then use the talkers' productions as stimuli for the accentedness rating task. This would allow a closer inspection of the source of accentedness and yield more robust results.

## References

- Almohareb, S. 2020. The Effects of rater-specific characteristics on the rating of foreign accent (Doctoral dissertation, George Mason University).
- Flege, J.E. 1984. The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, 76(3), 692-707.
- Flege, J.E., Munro, M.J., MacKay, I.R. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- Harada, T. 2006. The acquisition of single and geminate stops by English-speaking children in a Japanese immersion program. *Studies in Second Language Acquisition*, 28(4), 601-632.
- Mackay, I. R., Flege, J. E., Imai, S. 2006. Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, 27(2), 157-183.
- Park, H. 2013. Detecting foreign accent in monosyllables: The role of L1 phonotactics. *Journal of Phonetics*, 41(2), 78-87.
- Tsukada, K. 2009. An acoustic comparison of vowel length contrasts in Arabic, Japanese and Thai: Durational and spectral data. *Int. J. Asian Lang. Process.*, 19(4), 127-138.
- Tsurutani, C., Tsukada, K., Ishihara, S. 2010. Comparison of native and non-native perception of L2 Japanese speech varying in prosodic characteristics. In *Proceedings of 2010 Australasian International Conference of Speech Science and Technology* (pp. 122-125).
- Zhi, N., Li, A. 2021. Acoustic salience in the evaluations of intelligibility and foreign accentedness of nonnative vowel production. *Lingua*, 256, 103069.