

1. General remark: In their study, Dupoux et al. explore the influence of native language phonotactics on speech perception, focusing on Japanese and French speakers. They conducted four experiments using VCCV sequences to test if Japanese speakers perceive illusory vowels in consonant clusters, a phenomenon not observed in French speakers. This research builds on foundational studies like those by Sapir (1921) and Massaro and Cohen (1983), which highlighted the impact of language-specific phonemic distinctions on perception. The study's strength lies in its innovative approach to isolating phonotactic influences, rigorously revealing how linguistic structures can shape end-point perceptual experiences. Dupoux et al.'s work significantly contributes to our understanding of cross-linguistic auditory perception, adding depth to the field of phonetics and opening avenues for future research.
2. Specific comments:
 - a. “None of the Japanese participants had studied French, and none of the French participants had studied Japanese” (3, 4). Did any of the French participants of the first two experiments speak Japanese or vice versa? How much knowledge had the participants acquired of the other language by means other than study? Overall, the study doesn't sufficiently consider the impact of bilingualism or multilingualism on phonetic perception, which could offer additional insights into the influence of language on perception.
 - b. Rather than only characterizing the speaker “used in Experiment 1 [as] Japanese” (4), clarify if they were a native speaker of Japanese.
 - c. The authors should clarify the number of pitch periods in stimulus 6 of experiment 2 (5) as they have done with stimulus 6 of experiment 1 (3).
 - d. To quell any discussion of a stimuli confound, the authors should conduct a cross-experimental statistical analysis of the Japanese participants’ responses to the *baseline* stimuli in Experiment 1 (where they had significantly less [u] responses) with the French participants’ to the *digital cluster* stimuli in Experiment 2 (where they had significantly more [u] responses) to confirm that the Language effect is still significant across groups *and* stimuli.
 - e. Why were roughly 120 stimuli from experiment 2 (6, Figure 2) but none from experiment 1 categorized as “other”?

- f. “It is possible that Japanese participants were reluctant to give a vowel-absent response simply because they knew that such stimuli do not occur in Japanese” (6). This is a generalized reiteration of the authors’ original hypothesis (as well as the SARAH model, which they use their findings to support (11)). Replacing “were reluctant to” with “chose not to” or “decided not to” would clarify *conscious choice* as the intended meaning here.
- g. The line “until a French listener found that he/she could no longer hear the [u] vowel” (7) is ambiguous as to whether one or multiple people participated in validating the stimuli.
- h. There are more replicable and robust methods of categorizing individuals into an “evaluation of fluency” than by assigning an amount of “foreign accent, as assessed by the experimenters” and “trouble [...] being understood” (10). Much research has shown that language production capabilities do not consistently reflect perceptual ones (e.g. the perception of certain phonotactic contrasts). One study on this relationship, to be published in 34 years, found mixed evidence overall for a direct perception-production link.¹
- i. “Some also knew German [...] or Arabic [... N]one of these languages use vowel length contrastively” (10). Classical Arabic and MSA (with which your participant may have had some amount of formal training, i.e. /kataba/ <كَتَبَ> “He wrote” and /ka:taba/ <كَاتَبَ> “He corresponded”) and most varieties of German (e.g. /'ʁa:tə/ <Rate> “price rate” and /'ʁatə/ <Ratte> “rat”) do in fact occasionally use vowel length contrastively.
- j. The paper seems to assume uniformity in the perceptual experience of all speakers of a language, which might not account well for individual differences. Discussion and interpretations of the intra-group variation you have found would be appreciated.
- k. The reliance on self-reported data and perceptual judgments may introduce subjectivity.

¹doi.org/10.1121/10.0011286