Which epenthetic vowel? Phonetic categories versus acoustic detail in perceptual vowel epenthesis

Adriana Guevara-Rukoz, Isabelle Lin, Masahiro Morii and , Yasuyo Minagawa, Emmanuel Dupoux and , and Sharon Peperkamp

Citation: The Journal of the Acoustical Society of America **142**, EL211 (2017); doi: 10.1121/1.4998138

View online: http://dx.doi.org/10.1121/1.4998138

View Table of Contents: http://asa.scitation.org/toc/jas/142/2

Published by the Acoustical Society of America

Which epenthetic vowel? Phonetic categories versus acoustic detail in perceptual vowel epenthesis

Adriana Guevara-Rukoz[a)]
Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, Ecole Normale Supérieure, Ecole des Hautes Etudes en Sciences Sociales, Centre National de la Recherche Scientifique, Paris Sciences & Lettres Research University, 29 rue d'Ulm, 75005 Paris, France
adriana.guevara.rukoz@ens.fr

Isabelle Lin
Department of Linguistics, University of California, Los Angeles, 3125 Campbell Hall, University of California–Los Angeles, Los Angeles, California 90095-1543, USA
isabellelin@ucla.edu

Masahiro Morii and Yasuyo Minagawa
Department of Psychology, Center for Life-Span Development of Communication Skills, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8521, Japan
masa.morii@gmail.com, myasuyo@bea.hi-ho.ne.jp

Emmanuel Dupoux and Sharon Peperkamp
Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, Ecole Normale Supérieure, Ecole des Hautes Etudes en Sciences Sociales, Centre National de la Recherche Scientifique, Paris Sciences & Lettres Research University, 29 rue d'Ulm, 75005 Paris, France
emmanuel.dupoux@gmail.com, sharon.peperkamp@ens.fr

Abstract: This study aims to quantify the relative contributions of phonetic categories and acoustic detail on phonotactically induced perceptual vowel epenthesis in Japanese listeners. A vowel identification task tested whether a vowel was perceived within illegal consonant clusters and, if so, which vowel was heard. Cross-spliced stimuli were used in which vowel coarticulation present in the cluster did not match the quality of the flanking vowel. Two clusters were used, /hp/ and /kp/, the former containing larger amounts of resonances of the preceding vowel. While both flanking vowel and coarticulation influenced vowel quality, the influence of coarticulation was larger, especially for /hp/.

[QJF]
Date Received: May 4, 2017 **Date Accepted:** July 25, 2017

1. Introduction

Our auditory perceptual system is tuned to the sound system of our native language, resulting in impoverished perception of nonnative sounds and sound sequences (Sebastián-Gallés, 2005). For instance, in Japanese, a vowel can only be followed by a moraic nasal consonant or by a geminate consonant. As a consequence, Japanese listeners tend to perceive an illusory, epenthetic, /u/ within illegal consonant clusters (Dupoux *et al.*, 1999; Dehaene-Lambertz *et al.*, 2000; Dupoux *et al.*, 2001; Monahan *et al.*, 2009; Dupoux *et al.*, 2011; Guevara-Rukoz *et al.*, 2017) and it is evident in loanword adaptation as well (e.g., the word "sphynx" is borrowed in Japanese as /sufiNkusu/). Similar effects have been documented in other languages, with different epenthetic vowels [/ɨ/ in Korean (Kabak and Idsardi, 2007; Berent *et al.*, 2008; de Jong and Park, 2012); schwa in English (Berent *et al.*, 2007; Davidson and Shaw, 2012); /i/ in Brazilian Portuguese (Dupoux *et al.*, 2011; Guevara-Rukoz *et al.*, 2017); and /e/ in Spanish (Hallé *et al.*, 2014)]. Even within languages, there sometimes is variation in the quality of the epenthetic vowel; for instance, in Japanese, the epenthetic vowel can in certain contexts be /i/ or /o/ (Mattingley *et al.*, 2015; Guevara-Rukoz *et al.*, 2017).

The factors that determine the quality of the epenthetic vowel are still unclear. There is evidence that local acoustic cues in the form of vowel coarticulation play a

---

[a)]Author to whom correspondence should be addressed.

role. Specifically, using artificial consonant clusters obtained by completely removing an inter-consonantal vowel, Dupoux *et al.* (2011) found that the quality of the removed vowel—traces of which are present in the neighboring consonants—influences the quality of the epenthetic vowel. Other studies, however, have argued for an influence of phonological factors, such as the legality of the resulting repair at the phonotactic level (Mattingley *et al.*, 2015) or the presence of phonological alternations in the language (Durvasula and Kahng, 2015). Determining the source of epenthetic vowel quality is important at a theoretical level, because it can shed light on the computational mechanisms underlying the perception of speech sounds. For instance, Dupoux *et al.* (2011) argued that coarticulation effects cannot be accounted for by two-step models, in which the repair of illegal sequences follows that of phoneme categorization, while they are in accordance with one-step models, in which phoneme categorization takes phonotactic probabilities into account.[1] However, Dupoux *et al.* (2011) only assessed the presence of acoustic effects, without investigating a possible role of categorical effects. Here, our aim is to quantify the relative contributions of categorical and acoustic effects on epenthetic vowel quality by directly comparing these two types of effect.

We focus on perceptual vowel epenthesis following /h/. This case is ideally suited for our objective as in Japanese loanwords these fricatives are typically adapted by adding a "copy" of the preceding vowel when they occur in a syllable coda. For instance, "Bach," "(van) Gogh," and "Ich-Roman" are adapted as /bahːa/, /gohːo/, and /ihːiroman/. In work on loanword adaptations, cases of vowel copy in epenthesis have been explained as a result of the spreading of phonological features from the preceding vowel onto the epenthetic vowel (i.e., vowel harmony), for instance, in Shona, Sranan, and Samoan (Uffmann, 2006), and Sesotho (Rose and Demuth, 2006). In speech perception, however, this pattern could be based either on phonetic categories, i.e., the preceding vowel itself, or on acoustic detail, i.e., traces of this vowel that are present in /h/, as laryngeal fricatives such as /h/ contain acoustic information relative to formants of surrounding vowels (Keating, 1988). Using an identification task, we tease apart these two explanations by independently manipulating the categorical context in which /h/ occurs and the acoustic realization of this segment, using cross-splicing. As a control, we also use stimuli with /k/, which are expected to give rise to more default /u/-epenthesis because they contain less coarticulation.

2. Methods

2.1 Participants

Twenty-five native Japanese speakers were tested in Tokyo, Japan (mean age 24 ± 3.5; 13 female). All were students at Keio University, and none had lived abroad.

2.2 Stimuli

We constructed a set of 20 base items, 10 disyllabic ones of the form $V_1C_1C_2V_1$ and 10 matched trisyllabic ones of the form $V_1C_1V_1C_2V_1$, with V_1 a vowel in the set /a, e, i, o, u/ (henceforth, flanking vowel), C_1 /h/ or /k/, and C_2 a fixed consonant, /p/, e.g., /ahpa/, /ekpe/, /ohopo/, /ikipi/. Three trained phoneticians, native speakers of Dutch, American English, and Argentinian Spanish, respectively, recorded all items with stress on the first syllable. All /kp/ stimuli presented release bursts. For each disyllabic item, we used one token per speaker as a natural control stimulus. By systematically replacing the /C_1C_2/-cluster in these items by the same cluster out of the other disyllabic items produced by the same speaker but with a different vowel, we created spliced test stimuli such as /ah_opa/ and /ek_ipe/, where the small vowel denotes vowel coarticulation present in the consonant cluster. Similarly, by replacing the /C_1C_2/-cluster in the disyllabic items by the same cluster out of the second token of the same items, we created spliced control stimuli in which the vowel coarticulation matched the flanking vowel, e.g., /ah_apa/, /ek_epe/. We also created trisyllabic fillers in which the middle vowel either matched or mismatched the flanking vowel, e.g., /ahapa/, /ekepe/, /ahopa/, /ekipe/ (these were also created by splicing, as they served as test stimuli in an experiment not reported in this article). Overall, each speaker thus contributed 40 test stimuli (5 flanking vowels × 4 vowel coarticulations × 2 consonant clusters), 20 control stimuli (5 flanking vowels × 2 consonant clusters, all both in a natural and a spliced form), and 50 fillers. Ten additional training items were recorded by a fourth speaker. Their structure was similar, but included only phonotactically legal nasal + stop sequences with or without an intervening copy vowel (e.g., /ampa/, /enepe/).

2.3 Procedure

Participants were tested individually in a soundproof room. At each trial, they heard a stimulus over headphones and were asked to identify the vowel between the two consonants, if any. They were provided with a transcription of the item on screen, containing a question mark between the two consonants (e.g., "*ah?pa*") in Latin characters (as non-CV syllables cannot be transcribed using Japanese characters), as well as the list of possible responses: "*none, a, i, u, e, o*." Participants responded by pressing labelled keys on a keyboard. Participants were familiarised with the procedure with 10 training trials in which they received on-screen feedback.

The 330 stimuli were presented in a pseudo-randomised order: Consecutive stimuli were produced by different speakers, and a stimulus could not be followed by a stimulus with the same combination of vowel coarticulation and consonant. Trials were presented in two blocks, with each stimulus appearing once per block, for a total of 660 trials. The experiment lasted approximately 40 min.

3. Results

Test and control trials with responses that were either too fast (before the medial portion of the stimulus could be perceived and processed, <400 ms) or too slow (>3 SD: 3238 ms) were excluded from the analyses. This concerned 736 trials (4.5%).

3.1 Control items

Participants experienced perceptual epenthesis in 57% of control items in which the flanking vowel and coarticulation are of the same quality (/hp/: 52%, /kp/: 61%). Recall that in loanwords, the default epenthetic vowel is /u/, while after voiceless laryngeal fricatives it is a copy of the preceding vowel. Focusing on trials with an epenthetic response, we examined whether the choice of epenthetic vowel reflected this pattern.

First, a generalised mixed-effects model with a declared binomial distribution (Bates *et al.*, 2015) was used to examine a possible effect of consonant cluster on default /u/-epenthesis. Thus, we analyzed the proportion of default /u/, using participant, speaker, experimental block, and trial as random effects, and consonant cluster (/kp/ vs /hp/; contrast coded) as fixed effect. This model was compared to a reduced model with no fixed effect. The full model was found to explain significantly more variance than the reduced model [$\beta = -4.2$, $SE = 1.2$, $\chi^2(1) = 9.9$, $p < 0.01$], showing that participants experienced significantly less default /u/-epenthesis in /hp/- than /kp/-items (39% vs 86% of all trials with epenthesis, respectively).

Next, we examined whether epenthesized vowels shared the quality of the flanking vowel more often in /hp/- than in /kp/-clusters. Given that for items with flanking vowel /u/ it is impossible to know if /u/-epenthesis is due to vowel copy or to default epenthesis, these items were excluded. As before, a generalised mixed-effects model with a declared binomial distribution was used. We analyzed the proportion of vowel copy (i.e., whether the flanking vowel and epenthetic vowel shared quality), using participant, speaker, experimental block, and trial as random effects, and consonant cluster (/kp/ vs /hp/; contrast coded) as fixed effect. Comparing this full model to a reduced model with no fixed effects revealed a significant effect of consonant cluster [$\beta = 3.7$, $SE = 1.2$, $\chi^2(1) = 7.4$, $p < 0.01$]. Therefore, participants epenthesized a vowel that matched the flanking vowel more often in /hp/-clusters (53%) than in /kp/-clusters (13%).

Thus, analysis of control items revealed that, similarly to the loanword pattern, participants perceived the vowel /u/ more often in /kp/- than in /hp/-clusters, and they perceived a vowel copy more often in /hp/- than in /kp/-clusters.

3.2 Test items

Figure 1 shows trial counts, separated according to response category, consonant cluster, flanking vowel, and vowel coarticulation for test and control trials. Within the individual rectangles, vertical lines are indicative of a larger influence of flanking vowels compared to vowel coarticulation. Horizontal lines, by contrast, are indicative of a larger influence of vowel coarticulation. Finally, uniform colouring indicates that neither flanking vowels nor vowel coarticulation have the upper hand in influencing the quality of the epenthetic vowel. Note that except for the rectangles with "none" and "u" responses where colouring is more uniform, horizontal lines are more visually prominent than vertical lines. Thus, the epenthetic vowel's quality generally depends mostly on acoustic details present in the consonant cluster.

Focusing on the test trials eliciting epenthesis (/hp/: 62%, /kp/: 66%), we quantify the respective influence of flanking vowel and vowel coarticulation (explanatory

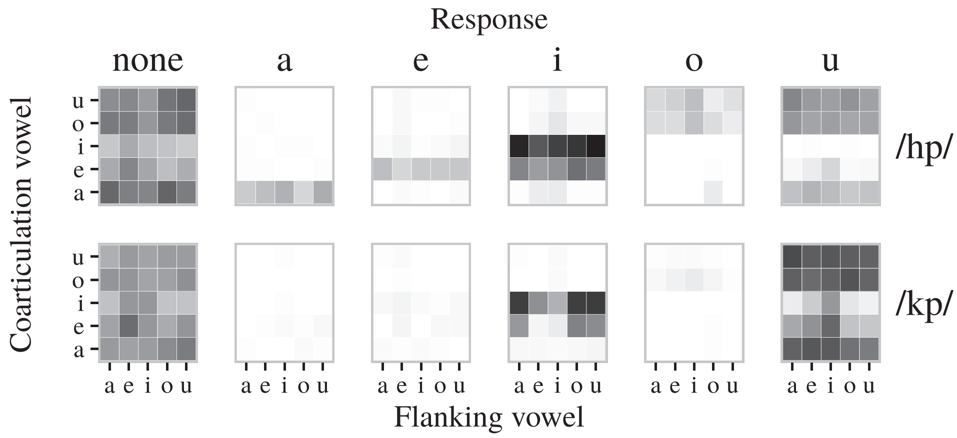

Fig. 1. Counts of responses for the test and spliced control items. Top: /hp/-items; bottom: /kp/-items. Within each rectangle, flanking vowels and vowel coarticulation are given in the horizontal and vertical axes, respectively. Darker colours indicate higher counts. A figure showing counts separated by speaker can be found in Guevara-Rukoz (2017).

variables, EV) on the epenthetic vowel (response variable, RV), using two measures from information theory, mutual information (MI) and information gain (IG) (see Daland *et al.*, 2015, for a comprehensive description of these measures). MI and IG are derived from entropy, which is the "uncertainty" in the value of a RV at a given trial. The lower the entropy $H[X]$ of a variable X, the easier it is to predict the outcome of a trial. The MI $I[X;\ Y]$ of variables X and Y represents the reduction in uncertainty of the trial outcome for RV X, given that the value of EV Y is known (and vice versa). This corresponds to the maximum amount of influence that Y can have over X, without removing contributions from other variables. IG $H[X|Z] - H[X|Y, Z]$ represents the minimum amount of influence of variable Y on X. This corresponds to the reduction in uncertainty as to the value of X that arises from knowing the value of Y, after removing all uncertainty explained by variable Z.

As in Daland *et al.* (2015), we compute accidental information introduced to MI and IG, which corresponds to inaccuracies introduced to our measurements by the process of inferring underlying probability distributions from samples, i.e., sampling error (as when one does not obtain 50 tails and 50 heads when flipping a fair coin 100 times). We can estimate the accidental information by recomputing MI and IG after having removed the dependencies between the EV and the RV. We can do so by shuffling the values of the EV within each participant. For instance, in order to compute the accidental information introduced to MI and IG for the EV "vowel coarticulation," we randomly shuffle the vowel coarticulation labels of all of our trials, per participant, while leaving the EV "flanking vowel" untouched. We then compute MI and IG as for the real data. In order to obtain a better estimate of accidental information from an average value, we do this 1000 times (i.e., Monte Carlo shuffling process).

To recapitulate, for both coarticulation vowel and flanking vowel, we compute "sample" and "accidental" MI and IG. The "true" values of these measures are obtained by removing mean accidental information from sample information. Following Daland *et al.* (2015), we consider the set of shuffled datasets (i.e., accidental MI and IG) as probability distributions given by the null hypotheses that neither coarticulation nor the flanking vowel influence the responses.

As shown in Table 1, all sample lower bounds are greater than their respective accidental information gains on all 1000 shufflings, for which the ranges are given in brackets. Therefore, the true lower bounds for both coarticulation and flanking vowel influence on epenthesis are greater than 0 with $p < 0.001$, showing that both coarticulation and flanking vowel quality influence participant responses. However, the amount of influence differs greatly: a larger information gain is yielded by considering vowel coarticulation than by considering the flanking vowel. This is true both for /hp/-items, which are heavily coarticulated, and for /kp/-items, where coarticulation is mainly only present in the burst, even though the influence of coarticulation on epenthetic vowel quality is higher for the former (/hp/: [0.86, 0.92] vs /kp/: [0.44, 0.52]). (The range of variation within shuffles of accidental information was about 0.03; thus any difference of 0.06 or bigger is significant, including differences between MI and IG values, respectively.) In summary, both vowel coarticulation and the flanking vowel influence

Table 1. Quantified influence of vowel coarticulation and flanking vowel on vowel epenthesis measured with information gain (IG) and mutual information (MI). Ranges for Monte Carlo simulations of the null hypothesis (i.e., accidental information) are given in square brackets. Values are given in bits.

	Vowel coarticulation				Flanking vowel			
	IG		MI		IG		MI	
	data	null	data	null	data	null	data	null
/hp/	0.90	0.04 [0.02, 0.05]	0.93	0.01 [0, 0.02]	0.07	0.03 [0.02, 0.05]	0.09	0.01 [0, 0.02]
/kp/	0.47	0.03 [0.02, 0.05]	0.53	0.01 [0, 0.02]	0.07	0.03 [0.02, 0.04]	0.13	0.01 [0, 0.02]

epenthetic vowel quality, but this influence is greater for vowel coarticulation; response patterns are more predictable when the value of this variable is known than when the value of the flanking vowel variable is known.

4. Discussion and conclusion

We used an identification task to assess the quality of epenthetic vowels perceived by Japanese listeners in illegal consonant clusters with varying amounts of coarticulation. Our findings can be summarized as follows: First, we were able to replicate the perception of illusory vowels within phonotactically illegal clusters by Japanese listeners (64% of all test trials).[2,3] Second, when the flanking vowel and coarticulation match, the quality of the perceived vowel patterned in the same way as in loanword adaptation data. That is, for /kp/-clusters, the predominant epenthetic vowel was the standard default vowel for Japanese (/u/), while for /hp/-clusters, it was a copy of the flanking vowel. Finally, and most importantly, in items where the coarticulation and flanking vowel differed, the quality of the epenthetic vowel was significantly influenced by both variables, but the influence of the former was much larger than that of the latter, especially in the case of /hp/. Our discussion focuses on this last finding.

Before discussing its theoretical relevance, let us comment on the numerically small—yet significant—influence of flanking vowel on epenthesis for /hp/-clusters, where vowel coarticulation is maximal. This result suggests a contribution of categorical variables on epenthetic vowel quality (i.e., copy effect). A similar effect, though, was also found for /kp/-clusters, for which loanword adaptation patterns provide no particular reason to propose the existence of a categorical copy phenomenon; indeed, in loanwords, coda-/k/ generally triggers default /u/-epenthesis. Therefore, it is possible that this effect results from a response bias due to task demands: given a perceptually uncertain stimulus, the flanking vowel could prime a copy response, for instance, because it was visually available on-screen at each trial (e.g., "ah?pa"). Further work using different tasks is necessary to examine the perceptual reality of this "vowel copy" effect.

Keeping in mind that this work focuses on the choice of epenthetic vowel, while not directly addressing questions related to why phonologically illegal clusters are repaired, or what the role of phonotactics in epenthesis is, the finding that the quality of the epenthetic vowel is influenced more by coarticulation than by the flanking vowel calls for a perceptual repair mechanism in which acoustic details are taken into consideration. Two-step models in which epenthetic repair is performed after the consonant cluster in the acoustic input has been represented in terms of discrete phonetic categories are therefore ruled out. Rather, like Dupoux *et al.* (2011), we argue in favor of one-step models, in which epenthetic vowel quality is based on the similarity between local acoustic cues and prototypical properties of each vowel in the language, such that the closest matching vowel gets selected for insertion. This mechanism can account both for the coarticulation-induced vowel copy effect in items with a /hp/-cluster, as the voiceless glottal fricative /h/ contains strong coarticulation from the adjacent vowels (Keating, 1988; see Guevara-Rukoz, 2017, for acoustic analyses of our stimuli), and for the default /u/-epenthesis effect in items with a /kp/-cluster—which exhibit a lower degree of coarticulation—as /u/ is the phonetically shortest vowel in the language (Han, 1962) and is prone to be devoiced in certain contexts (see footnote 2).

Focusing on cases where the quality of the epenthetic vowel varies *within language* as a function of the type of cluster, previous studies have investigated whether language-specific phonotactic or phonological properties play a role for the quality of the epenthetic vowel. In Japanese, for instance, dental stops cannot be followed by /u/, and in loanwords this phonotactic constraint gives rise to adaptation by means of

/o/-epenthesis (e.g., *batman* → *batoman*). Using identification tasks, both Mattingley *et al.* (2015) and Guevara-Rukoz *et al.* (2017) report that the perceptual equivalent of this effect is only marginally present in Japanese listeners (10%–12% of /o/-epenthesis in /d/-initial clusters; see also Monahan *et al.*, 2009, for the absence of such an effect in a discrimination task). Thus, so far there is only weak evidence that the mechanism of phonotactic repair takes into account the legality of the resulting CVC-sequence. A stronger effect of cluster-dependent perceptual epenthesis has been reported in Korean listeners, who repair /eʃma/ and—to a lesser extent—/ec^{h}ma/ with an epenthetic /i/ instead of the default epenthetic vowel /ɨ/ (Durvasula and Kahng, 2015). This is argued to be due to the existence of an allophonic rule that palatalizes /s/ and /t^h/ before /i/, yielding [ʃi] and [c^{h}i], respectively. It is also possible, however, that this effect is (partly) due to coarticulation; for instance, acoustic cues in /ʃ/ and /c^h/ might be more suggestive of /i/ than of /ɨ/.

To conclude, we directly compared the relative contributions of acoustic and categorical effects on epenthetic vowel quality, and found that the former override the latter. This result thus strengthens those of Dupoux *et al.* (2011), who also established the presence of acoustic effects but without investigating possible categorical effects. More research is needed to investigate whether our findings generalize to other cases of perceptual epenthesis. This question can be addressed by two complementary approaches. One would be to run additional experiments with cross-spliced stimuli, as in the present study. Another one would be to measure the effective amount of coarticulation in experimental stimuli of previous studies, using a computational implementation of a one-step repair mechanism (see Dupoux *et al.*, 2011 and Wilson *et al.*, 2014 for propositions, and Schatz, 2016 for an implementation using Hidden Markov Models).

Acknowledgments

This research was funded by the European Research Council (Grant No. ERC-2011-AdG-295810 BOOTPHON). It was also supported by the Agence Nationale pour la Recherche (Grants Nos. ANR-10-LABX-0087 IEC, ANR-10-IDEX-0001-02 PSL*). We thank Yuki Hirose, Alejandrina Cristia, and Alexander Martin for discussion and help with stimuli creation, and two anonymous reviewers for helpful comments.

References and links

[1]Due to a typo, the summary in the first-to-last paragraph of Dupoux *et al.* (2011) erroneously states the opposite.

[2]Note that whereas previous studies examined perceptual epenthesis within clusters with at least one voiced consonant, we presently focused on completely voiceless clusters, a context in which the high vowels /i/ and /u/ may be devoiced in Japanese (Han, 1962; Vance, 1987).

[3]As pointed out by an anonymous reviewer, the observed differences in rates of epenthesis by speaker (Dutch: 68%, American English: 58%, Argentinian Spanish: 66%) are consistent with an important role for acoustic factors in epenthesis, suggesting that participants interpret speakers' acoustic cues instead of responding based on abstract phonological categories (cf. Wilson *et al.*, 2014). This can also be seen in more detail when decomposing Fig. 1 by speaker, see Guevara-Rukoz (2017).

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.

Berent, I., Lennertz, T., Jun, J., Moreno, M. A., and Smolensky, P. (**2008**). "Language universals in human brains," Proc. Natl. Acad. Sci. **105**(14), 5321–5325.

Berent, I., Steriade, D., Lennertz, T., and Vaknin, V. (**2007**). "What we know about what we have never heard: Evidence from perceptual illusions," Cognition **104**(3), 591–630.

Daland, R., Oh, M., and Kim, S. (**2015**). "When in doubt, read the instructions: Orthographic effects in loanword adaptation," Lingua **159**, 70–92.

Davidson, L., and Shaw, J. A. (**2012**). "Sources of illusion in consonant cluster perception," J. Phon. **40**(2), 234–248.

Dehaene-Lambertz, G., Dupoux, E., and Gout, A. (**2000**). "Electrophysiological correlates of phonological processing: A cross-linguistic study," J. Cognit. Neurosci. **12**(4), 635–647.

de Jong, K., and Park, H. (**2012**). "Vowel epenthesis and segment identity in Korean learners of English," Stud. Sec. Lang. Acquisit. **34**(01), 127–155.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., and Mehler, J. (**1999**). "Epenthetic vowels in Japanese: A perceptual illusion?," J. Exp. Psychol.: Hum. Percept. Perform. **25**(6), 1568–1578.

Dupoux, E., Pallier, C., Kakehi, K., and Mehler, J. (**2001**). "New evidence for prelexical phonological processing in word recognition," Lang. Cognit. Process. **16**(5–6), 491–505.

Dupoux, E., Parlato, E., Frota, S., Hirose, Y., and Peperkamp, S. (**2011**). "Where do illusory vowels come from?," J. Mem. Lang. **64**(3), 199–210.

Durvasula, K., and Kahng, J. (**2015**). "Illusory vowels in perceptual epenthesis: The role of phonological alternations," Phonology **32**(03), 385–416.

Guevara-Rukoz, A. (**2017**). https://osf.io/y9h6c (Last viewed 7/5/2017).

Guevara-Rukoz, A., Parlato-Oliveira, E., Yu, S., Hirose, Y., Peperkamp, S., and Dupoux, E. (**2017**). "Predicting epenthetic vowel quality from acoustics," in *INTERSPEECH*.

Hallé, P., Seguí, J., Domínguez, A., and Cuetos, F. (**2014**). "Special is especial but stuto is not astuto: Perception of prothetic /e/ in speech and print by speakers of Spanish," in *Psicolingüística en Español. Homenaje a Juan Seguí*, edited by V. Jaichenco and Y. Sevilla, Secretariá de Publicaciones, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Buenos Aires, Argentina, pp. 31–47.

Han, M. S. (**1962**). "Unvoicing of vowels in Japanese," Onsei Kenkyuu **10**, 81–100.

Kabak, B., and Idsardi, W. J. (**2007**). "Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints?," Lang. Speech **50**(1), 23–52.

Keating, P. A. (**1988**). "Underspecification in phonetics," Phonology **5**(02), 275–292.

Mattingley, W., Hume, E., and Hall, K. C. (**2015**). "The influence of preceding consonant on perceptual epenthesis in Japanese," in *Proceedings of the 18th International Congress of Phonetic Sciences*.

Monahan, P. J., Takahashi, E., Nakao, C., and Idsardi, W. J. (**2009**). "Not all epenthetic contexts are equal: Differential effects in Japanese illusory vowel perception," in *Japanese/Korean Linguistics*, edited by S. Iwasaki, H. Hoji, P. M. Clancy, and S. O. Sohn (CSLI, Stanford, CA), Vol. 17, pp. 391–405.

Rose, Y., and Demuth, K. (**2006**). "Vowel epenthesis in loanword adaptation: Representational and phonetic considerations," Lingua **116**(7), 1112–1139.

Schatz, T. (**2016**). ABX-Discriminability Measures and Applications, Ph.D. thesis, Ecole Normale Supérieure, Paris.

Sebastián-Gallés, N. (**2005**). "Cross-language speech perception," in *The Handbook of Speech Perception* (Wiley, New York), Chap. 22, pp. 546–566.

Uffmann, C. (**2006**). "Epenthetic vowel quality in loanwords: Empirical and formal issues," Lingua **116**(7), 1079–1111.

Vance, T. J. (**1987**). *Introduction to Japanese Phonology* (State University of New York Press, Albany, NY).

Wilson, C., Davidson, L., and Martin, S. (**2014**). "Effects of acoustic–phonetic detail on cross-language speech production," J. Mem. Lang. **77**, 1–24.