Not All Epenthetic Contexts are Equal: Differential Effects in Japanese Illusory Vowel Perception*

PHILIP J. MONAHAN, ERI TAKAHASHI, CHIZURU NAKAO AND WILLIAM IDSARDI
University of Maryland

1. Introduction

Recent psycholinguistic research has emphasized the role of native language phonological knowledge in the online perception of non-native structure (Näätänen et al. 1997; Dupoux et al. 1999; Dehaene-Lambertz, Dupoux and Gout 2000; Dupoux et al. 2001; Kazanina, Phillips and Idsardi 2006; Kabak and Idsardi 2007). These findings indicate that the native language phonology strongly constrains the types of linguistic representations that can be constructed and entertained during speech perception. For example, Dupoux et al. (1999) demonstrated that Japanese listeners illusorily perceive an epenthetic [u] – more properly [ɯ] – when presented with consonant clusters that violate the phonotactics of Japanese. Thus, it appears that Japanese listeners bring their native language phonology to bear on their

* This project was supported by NIH DC R01 05660 to David Poeppel.

perception of speech, and moreover, epenthesis in Japanese loanword phonology (e.g., *maku̲donaru̲do̲* 'McDonald's') is not driven solely by articulatory constraints. A remaining question, however, is how much of the native language phonology is implicated in the perception of non-native structures. To address that question, we exploit the fact that Japanese restricts the types of vowels that can follow coronal consonants. Specifically, the mid-back vowel [o] is epenthesized following coronal consonants. The sequences *[tu] and *[ti] are illicit in Japanese, while the sequences [to], [te] and [ta] are attested. We investigate whether, in accordance with the native language phonology, Japanese listeners illusorily epenthesize [o] following coronal consonants. Our findings indicate that Japanese listeners can distinguish the strings *etoma* from *etma* and do not illusorily epenthesize [o] after coronal consonants, suggesting that only some of the native language phonology is engaged in the perception of non-native speech. There appears to be a more intricate interplay between epenthetic context and vowel category, and listeners are sensitive to both the context that requires repair and the relevant vowel category.

2. Japanese Syllable Structure

Japanese syllable structure is predominantly CV. The only permissible coda consonants in Japanese are either [+nasal], as in (1), or the first member of a geminate, as in (2). Other instances of coda consonants are illicit, as in (3).

(1)	a.	[tom.bo]	'dragonfly'
	b.	[kaN.ga.e]	'thought'
(2)	a.	[kap.pa]	'a river sprite'
	b.	[gak.koo]	'school'
(3)	a.	*[kap.ta]	
	b.	*[tog.ba]	
	c.	*[pa.kat]	(Itô 1989)

When foreign words are borrowed into Japanese, vowels are inserted in loanwords to break up consonant clusters (Itô and Mester 1995). Underlined segments represent an epenthesized vowel, as in (4).

(4)	a.	[fu.ru̲.ta.i.mu̲]	'full-time'
	b.	[su̲.fiN.ku̲.su̲]	'Sphinx'
	c.	[ku̲.ri.su̲.ma.su̲]	'Christmas'

In most cases, the vowel [u] is the vowel category epenthesized, as in (4). It has been suggested that the reason for why [u] is the most common epenthetic vowel is because it is the least sonorant among the Japanese vowels (Lovins 1973; Katayama 1998; Peperkamp and Dupoux 2003) and is susceptible to devoicing (Jaeger 1978; Vance 1987; Tsuchida 2001; Varden 1998). Recent psycholinguistic research has investigated the online implications of these constraints on Japanese phonotactics in the perception of nonnative structures.

3. Illusory Epenthesis

Dupoux et al. (1999) found that Japanese listeners reported hearing an illusory [u] vowel between two consonants (e.g., *ebzo*) even when no vowel was physically present in the stimulus. They created a continuum of nonword stimuli that ranged from items with no vowel between two consonants (e.g., *ebzo*) to items with a full vowel between the two consonants (e.g., *ebuzo*). In an offline judgment task, Japanese listeners, unlike the French controls, reported perceiving an illusory [u] more than 70 percent of the time when no vowel was present. Also, in a speeded ABX discrimination task, Japanese listeners failed to distinguish *ebzo* from *ebuzo*. The conclusion then is that Japanese listeners illusorily epenthesize [u] when perceiving an illicit consonant cluster and moreover, that speech perception is strongly influenced by phonotactic constraints of the native language.

These results have also been confirmed using electrophysiology. Dehaene-Lambertz et al. (2000) used Event-Related Potentials (ERPs) and elicited a difference in the electrophysiological response in French listeners as early 164 ms post-onset of a deviant stimulus (e.g., *ebzo* following a series of four *ebuzo* standards) in a mismatch negativity design. Japanese listeners, however, showed no such early difference in the processing of the deviant stimuli, and instead, only showed a much later and weaker effect. These findings indicate that Japanese speakers did not reliably perceive a difference between *ebzo* and *ebuzo* and that the phonotactic restrictions of the native language are implicated extremely early in the processing of nonnative speech structures.

Dupoux et al. (2001) subsequently asked whether such perceptual epenthesis was lexical in nature. To test a potential lexical account of the above findings, two sets of stimuli were created. Words in both sets had the shape CVCCV. One set (*u*-set; e.g., *sokdo*) was such that the insertion of [u] would produce a Japanese word (e.g., *sokudo* 'speed'). The second set (non-*u* set; e.g., *mikdo*) was such that the insertion of non-[u] vowels (i.e., [a], [i], [e] or [o]) would produce an existing Japanese word (e.g., *mikado* 'emperor'), but the insertion of epenthetic [u] would produce a nonword (e.g., *mikudo*). If perceptual epenthesis were driven by lexical factors, then the

items in the non-*u* set should be perceived with a non-*u* vowel (e.g., *mikado*) yielding a real-word of Japanese. On the other hand, if the effect is non-lexical in nature, words in both sets should be perceived with epenthetic [u], and only words in the *u*-set should be considered as a word. The results showed that in a lexical decision task, Japanese listeners responded to items in the *u*-set as real words but to items in the non-*u* set as nonwords. In short, it was found that non-epenthetic vowels are not perceived illusorily, even when the result of epenthesis would create a real word of Japanese. Instead, [u] seemed to be the predominant epenthetic vowel. Thus, perceptual epenthesis is not driven by potential lexical status. It should be noted, however, that Dupoux and colleagues have not systematically tested the appropriate epenthetic contexts for the vowels [i] and [o], which are productively epenthetic following palatal affricates and coronal consonants, respectively (see Section 4). Finally, using Korean, Kabak and Idsardi (2007) demonstrated that the illusory epenthesis effect is driven by constraints on syllable structure and not phonotactic sequencing restrictions. The Japanese data are compatible with either possibility.

These previous studies demonstrate that the native language phonology influences not only the production of non-native structures but also strongly constrains the online perception of speech, even in the initial stages of processing. In Section 4, we present additional facts of Japanese epenthesis that we exploit to address the question of how much of the phonology is implicated in the perception of speech.

4. Different Epenthetic Contexts

The preceding consonant in Japanese dictates the vowel category that is epenthesized to repair a syllable structure violation. All the consonants used in the previous psycholinguistic studies (e.g., [g, k, b, ʃ]) license [u] epenthesis (i.e., [gu, ku, bu, ʃu]). The high-back vowel [u], however, is not the only epenthetic vowel in Japanese. For example, *[tu] and *[du] are illicit sequences in Japanese. Thus, in loan words with a coronal obstruent (i.e., [t] or [d]) in coda position, [o] is the epenthetic vowel.

(5)	a.	'fight'	[fa.i.t**o**]	*[fa.i.tu]
	b.	'drive'	[d**o**.ra.i.bu]	*[du.ra.i.bu]
	c.	'strawberry'	[su.t**o**.ro.be.rii]	*[su.tu.ro.be.rii]
	d.	'McDonald's'	[ma.ku.do.na.ru.d**o**]	*[ma.ku.do.na.ru.du]

The high-front vowel [i] is epenthesized after palatal affricates [ʧ, ʤ], as in (6), and the high-back vowel [u] is inserted in all other environments, as was presented in (4).

(6)

a.	'catch'	[kyat.ʧ**i**]	*[kyat.ʧu]
b.	'pitch'	[pit.ʧ**i**]	*[pit.ʧu]
c.	'fudge'	[fad.ʤ**i**]	*[fad.ʤu]
d.	'range'	[ren.ʤ**i**]	*[ren.ʤu]

Thus, there are three distinct epenthetic contexts: after coronal stops (requiring [o] epenthesis), after palatal affricates (requiring [i] epenthesis) and everywhere else (requiring [u] epenthesis). Previous studies on Japanese illusory perception have only looked at cases where [u] epenthesis should occur and the Japanese listeners were found to perceive an illusory [u]. In this paper, we investigate the response properties of Japanese listeners to non-native syllable structures that would, in Japanese, induce [o] epenthesis. Do Japanese listeners perceive an illusory vowel that is appropriate for the given phonological environment, in this case [o], or do they perceive [u] by default regardless of the conditioning environment? In other words, are all epenthetic contexts equal and do Japanese listeners use all of their phonology (i.e., their knowledge of epenthetic conditioning environments) in perceiving non-native structures?

5. Experiment

5.1. Method

5.1.1 Participants

Sixteen native speakers of Japanese (mean age: 24.1 yrs) participated in the experiment. An additional 14 native speakers of English (mean age: 18.7 yrs) acted as control participants. All subjects provided informed consent and were compensated for their participation (either ¥1000 or $10). The experimental session lasted roughly 40 minutes.

5.1.2 Materials

The auditory stimuli used in the experiment were recorded by a female native speaker of Japanese who was naïve as to the purpose of the experiment. All tokens were recorded in the form [eCVma] with a full vowel present (e.g., *etoma, ekuma*). In this experiment we used the vowels [u] and [o] and the coronal consonants (i.e., [t, d]), velar consonants (i.e., [k, g]) and nasal consonants (i.e., [m, n]). The auditory stimuli were edited using Praat (Boersma 2001). Two types of stimuli were then created. The first type was of the form [eCma]. In these cases, the entire vowel from the original signal was removed, leaving only 10 ms of the stop consonant burst. Therefore, the token [etma] could have been derived from either [etoma] with [o] spliced out (henceforth, transcribed with a double strikethrough (i.e., [et~~o~~ma])) or from [etuma] with [u] spliced out (i.e., [et~~u~~ma]). The other type

of stimuli had the form [eCVma] with a full vowel. For these stimuli, we removed three pitch periods (~ 12 ms) from the vowel. This ensured that all experimental items were edited, thereby eliminating the possibility that participants would simply be responding to whether or not the stimuli had been modified.

Table 1: Items used in the experiment

		Vowels		
		[o]	[u]	None
Consonants	Coronal	*etoma*	*etuma*	*etma*
		edoma	*eduma*	*edma*
	Velar	*ekoma*	*ekuma*	*ekma*
		egoma	*eguma*	*egma*
	Nasal	*emoma*	*emuma*	*emma*
		enoma	*enuma*	*enma*

Table 1 lists all the words used in the experiment. Of the six consonants, the coronals [t, d] were directly relevant to our question. The velars [k, g] acted as control conditions to determine whether we could replicate previous findings (the '*ebzo* effect') by Dupoux and colleagues. The nasals [m, n] acted as fillers for the experiment. We also tested pairs that differed in vowel category (e.g., *etoma* vs. *etuma*), which all participants should be able to successfully distinguish. These items also acted as controls, demonstrating successful discrimination of a contrast. No cross-place or cross-voicing contrasts were tested (e.g., *[etma-ekma], *[etma-edma]), but all combinations of vowels ([o, u], no vowel) were tested (e.g., [etuma-etuma], [etuma-etoma], [etuma-etma], [etoma-etuma], [etoma-etoma], [etoma-etma], [etma-etma]).

5.1.3 Procedure

Following Kabak and Idsardi (2007), an AX discrimination task was used. Each participant was tested individually in a small room with a computer. Participants were instructed to listen to pairs of stimuli over headphones and judge whether the two stimuli were the 'same' or 'different'. If they thought the two sounds were the same, they were told to press *f* on the keyboard and *j* if different. Participants were asked to answer as quickly, yet accurately as possible. The inter-stimulus interval was 500 ms and the inter-trial interval pseudorandomly varied between 830 ms and 1250 ms. No feedback was given during the experiment. A trial timed-out if a participant did not respond within 5 seconds. The order of presentation was counterbalanced so that tokens were equally likely to be the first or second of the pair. Each pair of stimuli was repeated ten times, resulting in 480 total trials in the ex-

periment (360 test pairs, 120 filler pairs). There were an equal number of same and different pairs in the experiment, and all trials were autommatically randomized for each participant.

5.2. Hypotheses and Predictions

There are at least three different hypotheses we can formulate with regard to illusory epenthesis in Japanese. The first hypothesis (Hypothesis 1) holds that illusory epenthesis is indifferent to context, and Japanese listeners will perceive an illusory [u] regardless of the environment. Thus, when presented with a token like [etma], Japanese listeners should perceive [etuma], even though [u] is not the appropriate epenthetic vowel after coronal stops, and the sequence *[tu] is illicit in Japanese. Under this view, [u] is always the default illusory epenthetic vowel.

The second hypothesis (Hypothesis 2) states that all of native language phonology is engaged to constrain speech perception. Therefore, the category of the illusorily epenthesized vowel is dictated by the conditioning phonological environment. Therefore, after coronal stops, the vowel [o] is illusorily perceived, whereas after velars, [u] is perceived. When presented with [etma], Japanese participants are predicted to hear [etoma].

The two aforementioned hypotheses represent perhaps the most obvious and simplest possibilities. We would like to offer one final potential hypothesis (Hypothesis 3), one in which listeners are sensitive to the phonological contexts that condition the epenthetic vowel category but cannot postulate a vowel that does not share the properties of epenthetic [u] (i.e., does not undergo devoicing, is perceptually more sonorous).

Table 2: Predictions made by Hypotheses

	Stimulus	
	etoma-etma	*etuma-etma*
Hypothesis 1	DIFFERENT	SAME
Hypothesis 2	SAME	DIFFERENT
Hypothesis 3	DIFFERENT	DIFFERENT

Specifically, illusory epenthesis occurs only in the context where vowels that undergo devoicing can be epenthesized. The prediction, then, for our materials is that Japanese listeners will be unable to epenthesize [o] following the coronal consonants because [o] does not undergo devoicing and is perceptually more sonorous than [u]. Therefore, according to this hypothesis, they should be able to discriminate [etma] from both [etuma], because they are aware of the licensing conditions on [o] epenthesis, and from [etoma], because [o] is not an illusorily perceived vowel. Under this final

hypothesis, only certain aspects of the native language phonology can be implicated in online speech perception.

5.3. Results

We report the results of our experiment in A' scores. The A' statistic is a nonparametric version of d' scores. Like the d' measure, A' factors out subject response bias (Macmillan and Creelman 2005: 101). The formula for A' is presented in (7). An A' score of 1.0 represents perfect performance, while 0.5 represents chance-level performance.

(7)
$$A' = \frac{1}{2} - \frac{(H-F)(1+H-F)}{4H(1-F)} \quad if \quad H \geq F$$
$$A' = \frac{1}{2} - \frac{(F-H)(1+F-H)}{4F(1-H)} \quad if \quad H \leq F$$
where H = *hits* and F = *false alarms*

The mean A' scores for English and Japanese participants ranged between A' = 0.998 and A' = 0.583. Figure 1 shows the mean performance on each tested condition for the two groups.

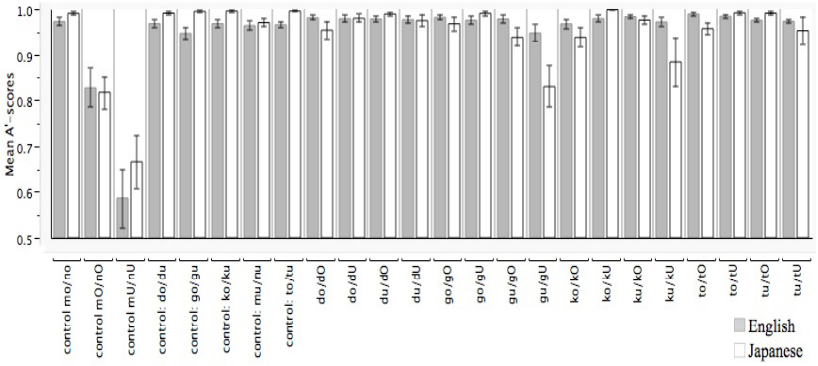

Figure 1: Sensitivity (measured in A') to all tested conditions. Error bars indicate one standard error of the mean. A capital vowel (i.e., [O]/[U]) refers to vowels not present in the stimulus (i.e., spliced from the original tokens).

A mixed-effects ANOVA (with Subject as a random effect) shows a significant main effect for the tested pair condition ($F(23,644) = 31.58$, $p < 0.0001$) but not for language ($F(1,28) = 0.11$, $p = 0.7478$) and a significant interaction between language and condition ($F(23,644) = 2.15$, $p < 0.0014$). Visual inspection of the results show a gap around A' = 0.9. All participants

performed very well with the [o]-[u] controls, as expected, with the Japanese listeners performing slightly better than the English listeners (presumably because the materials were recorded by a Japanese speaker). In addition, both groups faired worse at distinguishing [emma] from [enma]. It should also be noted that the Japanese listeners' ability to distinguish *ekema* from *ekuma* suggests that not all of the acoustic cues for the vowel were removed from the burst of the stop consonant. This is consistent with the finding that vowel cues are reliably present in as little as 10 ms of a stop burst extracted from a CV syllable (Blumstein and Stevens 1980).

Planned comparisons: We first report participants' performance on the velar conditions. As shown in Figure 2, Japanese listeners showed significantly poorer discrimination between *eguma/egema* (A' = 0.830) and *ekuma/ekema* (A' = 0.882) than English listeners (A' = 0.948, ΔA' = 0.12; $t(28) = -2.29$, $p < 0.01$; A' = 0.972, ΔA' = 0.09, $t(28) = -1.52$, $p < 0.05$, respectively). In other words, Japanese listeners could not discriminate [ekuma] and [ekma] as well as English listeners, replicating original *ebzo* effect (Dupoux et al. 1999). This finding confirms the idea that Japanese speakers illusorily perceive an epenthetic [u] to repair syllable constraint violations.

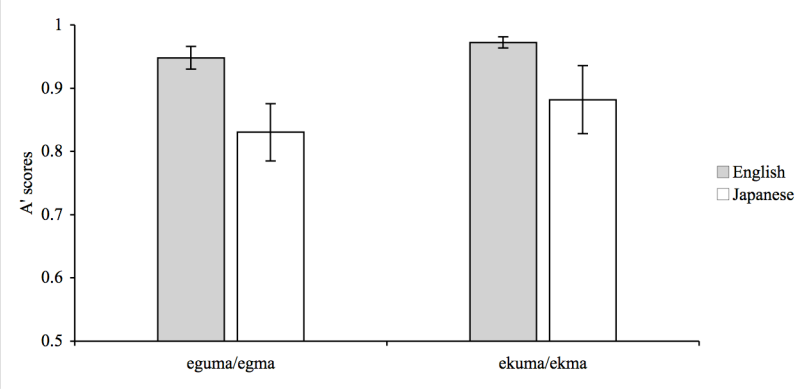

Figure 2: Sensitivity (measured in A') to the presence or absence of the vowel [u] following velar consonants. Error bars indicate one standard error of the mean.

In contrast Japanese participants successfully discriminated between velar consonants with and without [o], *egoma/egema* (A' = 0.967) and *ekoma/ekema* (A' = 0.937). This result replicates Dupoux et al. (2001) in that illusory epenthesis only occurs with [u] and not [o] in contexts where [u] is the appropriate epenthetic category. Here, Japanese listeners can be

interpreted to have perceived [ekma] as [ekuma] explaining why they could discriminate [ekma] and [ekoma]. No statistical difference from English listeners was found (*egoma/egɯma:* A' = 0.982, ΔA' = 0.015, $t(28) = -0.87$, $p = 0.39$; *ekoma/ekɯma:* A' = 0.968, ΔA' = 0.03, $t(28) = -1.26$, $p = 0.22$).

Our results thus far replicate previous findings (Dupoux et al. 1999; Dupoux et al. 2001). Our main question in this paper, however, is whether Japanese listeners illusorily perceive an epenthetic [o] in phonological environments where it is the appropriate epenthetic vowel. As Figure 3 shows, in the coronal conditions, Japanese subjects successfully discriminated between *edoma/edɯma* (A' = 0.953) and there was no significant difference between the Japanese and English listeners (A' = 0.981, ΔA' = 0.03; $t(28) = -1.33$, $p = 0.19$). Although Japanese listeners showed significantly poorer discrimination on the pair *etoma/etɯma* than English listeners (Japanese: A' = 0.955; English: A' = 0.988, ΔA' = 0.03, $t(28) = -2.27$, $p < 0.05$), it is difficult to conclude that Japanese listeners were unable to discriminate the two, since their A' was near ceiling. The reason for the statistically significant difference is likely due to the low variance shown by English speakers, who were at near perfect performance. These results, then, suggest that Japanese listeners can distinguish [etoma] from [etma], and that Japanese listeners do not perceive illusory [o] in the environments where [o] should be epenthesized, contrary to what was predicted by Hypothesis 2.

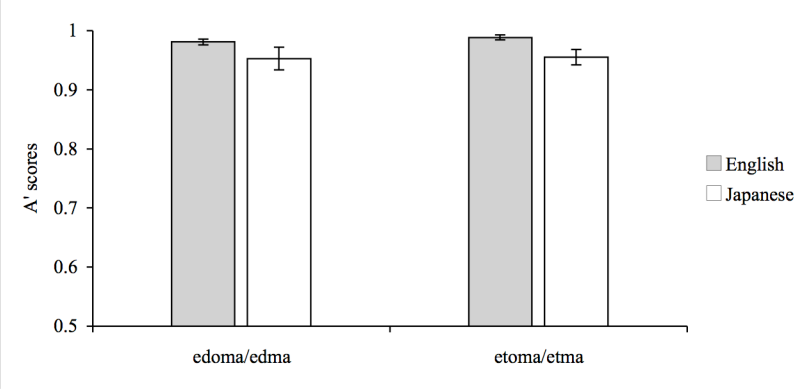

Figure 3: Sensitivity (measured in A') to the presence or absence of the vowel [o] following coronal consonants. Error bars indicate one standard error of the mean.

Additionally, Hypothesis 1 predicts that Japanese speakers perceive [etma] as [etuma]. However, Japanese subjects also showed excellent discrimination of *eduma/edɯma* (A' = 0.974) and *etuma/etɯma* (A' = 0.95), as shown in

Figure 4. There was no significant difference between Japanese and English participants' performance on *eduma/ed~~u~~ma* and *etuma/et~~u~~ma* (A' = 0.975, ΔA' < 0.001; $t(28) = -0.02$, $p = 0.98$; A' = 0.974, ΔA' = 0.02, $t(28) = -0.73$, $p = 0.47$, respectively). This suggests that, with coronals, [u] is not perceptually epenthesized either, contrary to the predictions of Hypothesis 1.

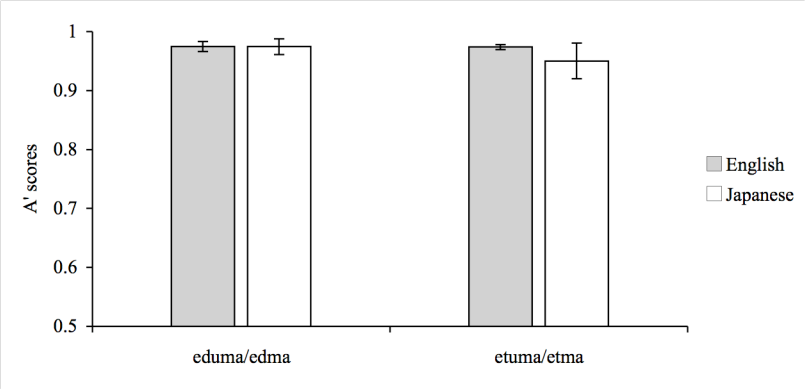

Figure 4: Sensitivity (measured in A') to the presence or absence of the vowel [u] following coronal consonants. Error bars indicate one standard error of the mean.

As Hypothesis 3 predicted, Japanese speakers appear to have perceived neither illusory [o] nor [u] following coronal consonants. Even though Japanese participants' performance on *edoma/ed~~o~~ma* and *etoma/et~~o~~ma* may have been slightly poorer than that of the English participants, the important point is that Japanese participants' performance on the coronal condition is significantly better than their performance on the velar condition with epenthetic [u], as shown in Figure 5. Specifically, collapsing across voicing (no effects of voicing were found), Japanese speakers performed significantly better at discriminating coronals with epenthetic [o] (*e[d,t]oma/e[d,t]ma*) than velars with epenthetic [u] (*e[g,k]uma/e[g,k]ma*; ΔA' = 0.1; $t(15) = 3.15$; $p < 0.01$).

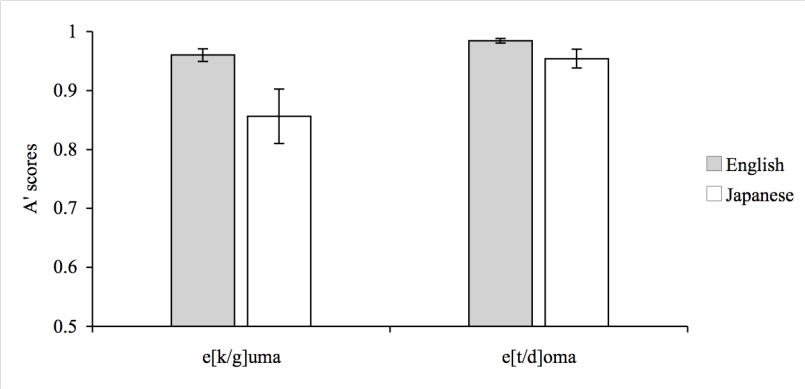

Figure 5: Sensitivity (measured in A') to the presence or absence of the epenthetic [u] or [o] following velar and coronal consonants, respectively. Error bars represent the standard error of the mean.

In short, following coronal consonants, Japanese listeners do not illusorily epenthesize [u], the most common epenthetic vowel in Japanese, contra Hypothesis 1. Moreover, Japanese listeners do not illusorily epenthesize [o], even when the native language phonology requires that be the epenthetic vowel category, contra Hypothesis 2. Instead, Japanese listeners appear to be sensitive to the conditioning environments for epenthesis in online speech perception but are unable to epenthesize the appropriate epenthetic vowel in some cases (in particular, the mid-back vowel [o]), even when their native language phonology dictates.

6. Discussion

The results from our experiment demonstrate that Japanese listeners do not illusorily epenthesize [o] after coronal stops. It seems that, unlike [u], the vowel [o] cannot be perceptually epenthesized even when it is the appropriate vowel category according to the native language phonology and is the vowel used in loanwords with coronal consonants. What these findings suggest is that Japanese listeners are aware of the conditioning contexts for the types of epenthetic vowel categories but are unable to epenthesize the appropriate vowel category in all contexts. In other words, not all of the native language phonology is implicated in online speech perception.

We might ask why we were unable to replicate the canonical *ebzo* findings in contexts where [o] is the appropriate epenthetic vowel. First, recall that the explanation for illusory [u] epenthesis in Japanese was that (a) it is least sonorant among the Japanese vowels and (b) it is susceptible to devoicing. With [o], however, these conditions are not met. First, unlike [u],

[o] does not undergo devoicing. Therefore, it might be the case that only vowels that undergo devoicing can be perceptually epenthesized. Second, [o] is higher than [u] in the sonority hierarchy. That is, [o] is more sonorous and thus may be too perceptible to be illusory (i.e., the lack of [o] would be too noticeable).

What might the Japanese listener be doing when he/she perceives a consonant sequence (e.g., *etma*) that violates the syllable structure of Japanese but requires the epenthesis of the mid-back vowel [o]? First, we suggest that the Japanese listener parses the string into three syllable-sized units [e.tV.ma]. They realize that [o] is the appropriate epenthetic vowel category, but it is unable to be illusorily epenthesized. At this point, the identity of the vowel is unknown, but Japanese listeners know that there must be a vowel between [t] and [m], because [t] cannot be parsed as a coda consonant. An 'unknown' vowel is assigned to nucleic position in the second syllable. In trying to determine the category of the vowel, [o] is the appropriate category following [t]. However, [o] cannot be perceived illusorily (see above). The candidates [u] and [i], both of which undergo devoicing and are common epenthetic vowels, may be considered. However, neither [tu] nor [ti] is an acceptable sequence in Japanese. In Japanese loan words, the nature of the consonant [t] is usually changed to accommodate such sequences (/tu/ → [tsu]; /ti/ → [ʧi]). But in the case of [etma], there is no evidence that [t] has changed to [ts] or [ʧ]. Therefore, the lack of evidence for consonant mutation prohibits those possible percepts (i.e., *[etuma], *[etima]). Thus, when presented with [etma], Japanese speakers are predicted to perceive something like [etVma] and will be able to distinguish it from both [etuma] and [etoma]. Thus, we can conclude that the knowledge of their phonological grammar constrains the types of representations Japanese listeners can construct, but only sub-parts of the phonology are actually implicated.

It should be noted, however, that even in the original studies on Japanese illusory vowel perception (Dupoux et al. 1999), Japanese listeners were presented with non-native structures that would not induce the presence of a voiceless vowel. In all the items in Experiments 1 and 2 in Dupoux et al. (1999), the two adjacent consonants were voiced. In Japanese, vowel devoicing only occurs between voiceless consonants (Han 1962; McCawley 1968; Vance 1987). In other words, the trigger or the context of the devoicing rule (i.e., between voiceless segments) do not seem to matter, while the target of the devoicing rule (i.e., only high vowels can be devoiced in Japanese) does matter in speech perception. This asymmetry explains why Japanese listeners illusorily perceive [u] in *ebzo* but not [o] in *etma*. Thus, it seems that, even in the original studies, Japanese listeners were only using sub-parts of their native language phonology to constrain

the grammatical representations they were constructing during online speech perception. One prediction that can be made if this were true is that Japanese listeners should perceive an illusory [i] in items like *etʃma* and *edʒma*. This is because [i] is productively epenthesized after palatal affricates and it is a high vowel that undergoes devoicing. Japanese listeners that we have consulted have this intuition.

7. Conclusion

In conclusion, our findings indicate that particular portions of the native language phonology are implicated in online speech perception, while other aspects are not. Specifically, we found that in environments where [u] is not the appropriate epenthetic vowel (i.e., after coronal stops), no illusory vowel perception was found. Thus, Japanese listeners seem to be sensitive to both the phonological contexts that require repair and the relevant vowel category. To validate the proposal that Japanese listeners are constructing partial representations (i.e., a trisyllabic template, see Hypothesis 3), additional experimentation is required. What remains evident, however, is that only some of the phonology is recruited for online speech perception, and future research is required to identify which pieces of the phonology are engaged.

References

Blumstein, S.E. and K. Stevens. 1980. Perceptual Invariance and Onset Spectra for Stop Consonants in Different Vowel Environments. *Journal of the Acoustical Society of America* 67: 648-662.

Boersma, P. 2001. Praat, A System for Doing Phonetics by Computer. *Glot International* 5: 341-345.

Dehaene-Lambertz, G., E. Dupoux and A. Gout. 2000. Electrophysiological Correlates of Phonological Processing: A Cross-Linguistic Study. *Journal of Cognitive Neuroscience* 12: 635-647.

Dupoux, E., K. Kakehi, Y. Hirose, C. Pallier and J. Mehler. 1999. Epenthetic Vowels in Japanese: A Perceptual Illusion? *Journal of Experimental Psychology*: *Human Perception and Performance* 25: 1568-1578.

Dupoux, E., C. Pallier, K. Kakehi and J. Mehler. 2001. New Evidence for Prelexical Phonological Processing in Word Recognition. *Language and Cognitive Processes* 16: 491-505.

Han, M.S. 1962. Unvoicing of Vowels in Japanese. *Onsei no Kenkyuu* 10: 81-100.

Itô, J. 1989. A Prosodic Theory of Epenthesis. *Natural Language and Linguistic Theory* 7: 217-259.

Itô, J. and A. Mester. 1995. Japanese Phonology. *The Handbook of Phonological Theory*, ed. J. Goldsmith, 817-838. Oxford: Blackwell.

Jaeger, J. 1978. Speech Aerodynamics and Phonological Universals. *Proceedings of the 4th Annual Meeting of The Berkeley Linguistics Society*, eds. J.J. Jaeger, A.C. Woodbury, F. Ackerman, C. Chlarello, O.D. Gensler, J. Kingston, E.E. Sweetser, H. Thomson and K.W. Whistler, 311-329. Berkeley, CA: BLS.

Kabak, B. and W.J. Idsardi. 2007. Perceptual Distortions in the Adaptation of English Consonant Clusters: Syllable Structure or Consonantal Constraints? *Language and Speech* 50: 23-52.

Katayama, M. 1998. *Optimality Theory and Japanese Loanword Phonology*. PhD Dissertation, University of California, Santa Cruz.

Kazanina, N., C. Phillips and W.J. Idsardi. 2006. The Influence of Meaning on the Perception of Speech Sounds. *Proceedings of the National Academy of Sciences* 103: 11381-11386.

Lovins, J. 1973. *Loanwords and the Phonological Structure of Japanese*. PhD Dissertation, University of Chicago.

Macmillan, N. A. and C. D. Creelman. 2005. *Detection Theory: A User's Guide*. 2nd ed. Mahwah: Lawrence Erlbaum Associates.

McCawley, J. 1968. *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Näätänen, R., A. Lehtokoski, M. Lennes, M. Cheour, M. Houtllainene, A. Iivonen, M. Vainio, P. Alku, R.J. Iimoniemi, A. Luuk, J. Allik, J. Sinkkonen and K. Alho. 1997. Language-Specific Phoneme Representations Revealed by Electric and Magnetic Brain Responses. *Nature* 385: 432-434.

Peperkamp, S. and E. Dupoux. 2003. Reinterpreting Loanword Adaptations: The Role of Perception. *Proceedings of the International Congress of Phonetic Sciences* 15: 367-370.

Tsuchida, A. 2001. Japanese Vowel Devoicing: Cases of Consecutive Devoicing Environments. *Journal of East Asian Linguistics* 10: 225-245.

Vance, T. 1987. *An Introduction to Japanese Phonology*. Albany, NY: SUNY Press.

Varden, J.K. 1998. *On High Vowel Devoicing in Standard Modern Japanese: Implications for Current Phonological Theory*. PhD Dissertation, University of Washington.