Nat Lang Linguist Theory (2019) 37:825–868
https://doi.org/10.1007/s11049-018-9423-2

On the relation between speech perception and loanword adaptation

Cross-linguistic perception of Korean-illicit word-medial clusters

Robert Daland[1] · **Mira Oh**[2] · **Lisa Davidson**[3]

Received: 22 August 2016 / Accepted: 25 May 2018 / Published online: 20 August 2018

Abstract Loanword adaptation has been claimed to provide a unique window onto the relation between speech perception and the phonological grammar. This paper focuses on whether the 'illusory vowel' effect—in which the presence/absence of a vowel is poorly discriminated within an illicit cluster—is sufficient to explain why vowel epenthesis is the preferred repair for medial clusters in Korean loanword adaptation. A cross-linguistic discrimination experiment revealed a causative role of the stop release burst (or other audible frication noise) in the perception of an illusory vowel; in some cases, perception alone explains vowel epenthesis in loanword adaptation. A follow-up, identification experiment showed that Koreans' perceptual similarity judgements do not match up with the adaptation pattern for stop-nasal clusters (e.g. pakna), although they do for fricative-stop and stop-stop clusters (e.g. paska, pakta). This finding is problematic for a purely perceptual account of loanword adaptation. The paper sketches a Bayesian account of Korean speech perception that integrates top-down phonotactic likelihood and bottom-up acoustic match and is able to explain the experimental results. It closes with some speculation on the role of the Preservation Principle versus perception in loanword adaptation.

Keywords Speech perception · Loanword phonology · Bayesian · Korean · Phonotactics

✉ M. Oh
mroh@chonnam.ac.kr

R. Daland
r.daland@gmail.com

L. Davidson
lisa.davidson@nyu.edu

1 Department of Linguistics, UCLA, Los Angeles, USA

2 Department of English, Chonnam National University, Gwangju, South Korea

3 Department of Linguistics New York University, New York, USA

1 Introduction

The phonological structure of the native language has a profound and lasting effect on the speech production and perception systems. Decades of research have established that monolinguals and low-proficiency bilinguals produce and perceive consonants in terms of native language phonetic categories (Goto 1971; Miyawaki et al. 1975; Best 1991; Bradlow et al. 1997; Flege 1995; Kuhl 2004). Although native language effects are best understood for phonetic categories in word-initial and other salient positions, it is also clear that native language *phonotactics* affect language behavior. For example, English words may begin with [sl] and [t^hɹ] (*slow*, *try*) but no other combination of these segments. Pitt and McQueen (1998) found that the absence of word-initial *[sɹ] and *[t^hl] biases listener perception: an ambiguous sound [?] tends to be perceived as *r* when it occurs in [t^h?a] (i.e. *tra* not *tla*), but as *l* when the acoustically identical sound occurs in [s?a] (i.e. *sla* not *sra*). Hallé et al. (1998) have shown an analogous effect in French, which allows word-initial *kl*/*gl* but not *tl*/*dl* clusters; French listeners are prone to mishear [tla] as [kla] and [dla] as [gla].

This kind of perceptual distortion might be thought of as adaptive, since monolingual listeners have no need to identify speech sound sequences that cannot occur in their native language. However, this is a problem that listeners may face in loanword adaptation, when a borrower adapts a word from the source language to accommodate the phonological and phonetic demands of the borrowing language. For this reason, loanword adaptation has been regarded as a fruitful tool for probing the relationship between speech perception and native language phonology (e.g. Silverman 1992; Kang 2003; Yip 2006; Peperkamp et al. 2008).

A notable fact is that loanwords containing phonotactically illicit sequences are often repaired by vowel epenthesis, even when the borrowing language repairs the 'same' sequence a different way in native words (e.g. Paradis and LaCharité 1997: 384, fn. 6). This is illustrated in (1), using the Korean native and non-native medial consonant clusters that are the focus of this paper:

(1)

	cluster	example	native phonological processes
a.	stop-stop	/nat^h-ke/ [nat͡k'ɛ] 'a piece'	faithful / laryngeal neutralization
b.	stop-nasal	/kuk-min/ [kuŋmin] 'nation'	nasality assimilation
c.	fric-stop	/pus-t^{h}oŋ/ [put͡t^{h}oŋ] 'brush case'	neutralization (laryngeal, manner)
	cluster	**example (NAKL 1991)**	**loanword adaptation pattern**
d.	stop-stop	*cha*[pt]*er*$_E$ ∼>[1] [c^{h}ɛp͡t^{h}ʌ]$_K$	faithful / laryngeal neutralization
e.	stop-nasal	*pi*[kn]*ic*$_E$ ∼> [p^{h}ik^{h}i͡nik͡]$_K$	vowel epenthesis
f.	fric-stop	*di*[sk]*ette*$_E$ ∼> [tisik^{h}ɛt͡]$_K$	vowel epenthesis

[1] We use the tilde-angle bracket sequence ∼> to indicate loanword adaptation. Subscripts indicate the source/borrowing language: (*E*)nglish, (*K*)orean, (*J*)apanese; however, the subscript will be omitted when the source/borrowing language is clear in context. All loanword data are from a loanword corpus created by the National Academy of the Korean Language (NAKL 1991), unless explicitly noted otherwise.

Illegal medial consonant clusters in native Korean words are normally repaired by coda neutralization and/or assimilation (1abc). However, when the analogous cluster occurs in a loanword source form, it may be adapted with an epenthetic vowel (1ef).

The specific research question of this paper is where the epenthetic vowel in (1ef) 'comes from.' This paper investigates the hypothesis that the epenthetic vowel originates in speech perception, i.e. that Korean adapters perceive an 'illusory vowel' in words like *picnic* and *diskette*, and that this illusory vowel perception is a consequence of expectations generated by the phonotactic and phonetic structure of the Korean language. These specific hypotheses figure into a broader research program identifying the relation between native language exposure, speech perception, and loanword adaptation.

In the next section, we review several literatures forming the theoretical and empirical backdrop for the present study. Our review of the Korean loanword adaptation facts turns up a disagreement in the literature as to the actual loanword adaptation facts: Kabak and Idsardi (2007) claim that Koreans adapt medial stop-nasal clusters with the native nasality assimilation process in (1b), while Boersma and Hamann (2009) claim that they are adapted with vowel epenthesis as in (1e). To determine whether a speech perception theory explains the loanword adaptation facts, it is first necessary to know what those facts are. Therefore, the first empirical section of this paper consists of a loanword corpus study investigating both stop-stop and stop-nasal clusters (1de). As it turns out, the loanword adaptation pattern is categorical as shown in (1); Kabak and Idsardi (2007) didn't take into account the effect of morphological structure (Oh 2012).

Our review of the illusory vowel effect turns up various lines of argumentation that a subphonemic property—the presence of an audible stop burst release or other frication noise—plays a powerful, causative role in the perception of an illusory vowel. The second empirical section of the paper reports the results of two perception experiments testing the hypothesis that loanword adaptation can be explained as identifying loanwords with the perceptually closest, phonotactically licit form. Participants were exposed to 'source' forms containing the 3 medial cluster types shown in (1) (e.g. *pakta*, *pakna*, *paska*) and potential 'repair' forms (coda neutralization/assimilation, or epenthesis) that could yield a licit Korean form. Experiment I assessed whether Korean (and American) listeners could discriminate source forms from various repairs in an AX discrimination task. Experiment II evaluated Korean listeners' perceptual similarity judgments with an ABX task in which they were asked to indicate whether the illegal source form sounded more like the native/assimilatory repair or the loanword/epenthetic repair. The presence of audible burst/frication was manipulated in the stop-initial clusters; the results of both experiments suggest a causative role of burst/frication noise in the illusory vowel effect. *Contra* hypothesis, for medial stop-nasal clusters Korean listeners did not identify the loanword/epenthetic repair as most perceptually similar to the illegal source form (i.e. [p^hakn̚a] ∼> [p^haŋna], *[p^hak^{h}ɨna]). These results suggest a nuanced picture, wherein the illusory vowel effect explains some but not all of the categorical loanword adaptation pattern in (1). The paper closes with a Bayesian model of speech perception in Korean which integrates phonotactic and phonetic factors, and a discussion of the relation between speech perception and loanword adaptation.

2 Background

2.1 The illusory vowel effect

In a seminal study, Dupoux and colleagues showed that phonotactically illicit consonant clusters may be perceived with an epenthetic vowel (Dupoux et al. 1999). Their test language, Japanese, allows for only two kinds of word-medial clusters: geminates (e.g. *natto* 'fermented soybeans'), and nasal-obstruent sequences (e.g. *kendo* 'sword-fighting'). Thus, a sequence like *ebzo* is phonotactically illegal in Japanese, while the sequence *ebuzo* is the perceptually closest, legal form. Dupoux and colleagues presented Japanese listeners with pairs like *ebzo*∼*ebuzo* in a speeded discrimination task, and found that performance was near chance, which they interpreted as evidence that Japanese listeners perceived an illusory vowel in the *ebzo* items. Subsequent work (e.g. Dupoux et al. 2011), has refined and expanded our knowledge of this phenomenon by pointing to the role of language-specific phonetics and coarticulatory cues in shaping which illusory vowel is perceived. However, the fundamental point that illegal clusters are often perceived with an intervening vowel has stood the test of time. Accordingly, we will use the term *illusory vowel effect* operationally, to refer to poor discrimination for the presence/absence of a vowel in otherwise matched contexts.

Kabak and Idsardi (2007) hypothesized that the illusory vowel effect arose from syllable structure phonotactics, rather than from sequential phonotactics per se. They conducted an AX discrimination study with Korean listeners that was closely modeled after Dupoux et al.'s (1999) [ebzo]∼[ebɯzo] study, but included a wider array of medial clusters:

(2) *Structure and results of Kabak and Idsardi's (2007) AX study on illusory vowel effect in Korean*

Condition	Phonotactics	Example	Comment
control	coda: *legal*	A: [pʰakt̚ʰa]	A has unreleased stop
	syllable contact: *legal*	X: [pʰakʰʊtʰa]	before aspirated stop
legal coda	coda: *legal*	A: [pʰakn̚a]	A has unreleased stop
	syllable contact: *illegal*	X: [pʰakʰʊna]	before nasal
illegal coda	coda: *illegal*	A: [pʰactʰa]	A has (released) palatal
	syllable contact: *illegal*	X: [pʰacʰɪtʰa][2]	before aspirated stop

Kabak and Idsardi (2007) found that Korean listeners exhibited good discrimination in control trials, which is expected since the presence/absence of a vowel is contrastive in Korean. They found generally poor discrimination of illegal coda trials, 'replicating' the illusory vowel effect in Korean (there was one exception, the [pʰag̚ma]∼ [pʰagʊma] item, which we will discuss later). The crucial finding pertained to the legal coda trials, which contain an illegal syllable contact (but whose

[2]The 'different'/B stimuli followed the Korean loanword adaptation pattern. Nonprevocalic palatal consonants are customarily adapted by epenthesizing a high front vowel; elsewhere, it is a high mid vowel that is epenthesized instead.

syllables themselves are licit). In these cases Kabak and Idsardi found good discrimination, and they accordingly interpreted this as evidence that the illusory vowel effect is specifically driven by syllable structure violations. The primary focus of Kabak and Idsardi's paper was on speech perception, but they did also briefly address the relation between speech perception and loanword adaptation. They claimed that the loanword adaptation pattern is essentially isomorphic to their perception results. In particular, they found that Korean listeners did not exhibit the illusory vowel effect in stop-nasal clusters, and argued that Korean loanwords do not exhibit an epenthetic vowel when the source form contains a medial stop-nasal cluster. This claim is at odds with the data point we showed in (1e); and we will address that discrepancy later. Prior to that, it will prove useful to examine illusory/epenthetic vowels in other contexts besides medial clusters.

An especially interesting aspect of Korean loanwords is that word-final stops may be adapted variably, as shown in (3):

(3) *Lexical variability in Korean adaptation of word-final stops*

a.	epenthesis always	*start* ~>	[sit^{h}at^{h}ɨ]
b.	epenthesis never	*pack* ~>	[p^{h}ɛk̚]
c.	variable epenthesis	*make* ~>	[mɛik^{h}ɨ]~[mɛik̚]

This variability is interesting because Korean phonotactics allows word-final stops (cf. [pat̚] 'field', [pap̚] 'rice', [sak̚] 'payment'). Therefore, unlike most other cases of vowel epenthesis in loanword adaptation, the presence of the epenthetic vowel is not forced by phonotactic accommodation to the borrowing language. Kang (2003) accordingly hypothesized that the variability in adaptation might originate from phonetic variation in the borrowing words. She used the TIMIT corpus as a sample of typical English pronunciations and investigated several phonetic properties. The crucial result for the present paper is that the likelihood of an audible release burst was strongly correlated with the presence of an epenthetic vowel in the adapted form. Kang proposed that the audible release burst in coda positions contributed causatively to Korean listeners' perception of an (illusory) vowel, which was then incorporated into the loanword form. One rationale for Kang's proposal is that coda segments in Korean are obligatorily unreleased (Shin 2011). It follows that for Korean listeners, the presence of an audible release burst is a reliable and predictive cue for being an onset/prevocalic segment.

There have been several subsequent studies suggesting that audible stop release bursts may contribute to the illusory vowel effect in cross-linguistic speech perception and/or loanword adaptation. For example, Peperkamp et al. (2008) began by noting that word-final nasal consonants are adapted into Japanese differently, depending on the source language. French word-final nasals are usually adapted into Japanese with epenthesis (e.g. *douane* ~> [dɯannɯ]). In contrast, English word-final nasals are usually adapted into Japanese without epenthesis (cf. *origin* ~> [oɾidʑin]). The authors hypothesized that the difference arises because French final nasals are typically produced with a strong consonantal release, while English final nasals are typically pronounced with a weak or inaudible consonantal release. To test this hypothesis, they had both English and French speakers produce forms like [mɑbiˈjun], and asked

Japanese listeners to identify the closest matching Japanese form from a list of several options, differing crucially in the presence/absence of a final vowel. The results showed that "[t]he presence of a final vowel in the responses was best predicted by a single variable, the relative duration of the release multiplied by its intensity, in other words the spectral energy contained within the release of the nasal consonant... This variable alone accounted for 61% of the variation" (p. 145). A related finding is reported in Wilson and Davidson (2013). They investigated English speakers' imitations of phonotactically illicit word-initial clusters produced by Russian talkers (for whom onset clusters like [bn] and [bd] are licit). Wilson and colleagues manipulated non-contrastive acoustic cues such as the duration and amplitude of the stop burst in the initial stop in the cluster. They found that longer release durations led to the production of significantly more epenthetic vowels, while lower amplitude bursts led to alternative repairs, such as consonant change and deletion. This variation must be attributed to acoustic variation in the stimuli, rather than unfamiliarity with producing the sequences, since there was within-cluster variation and the relevant stimuli were equally unpracticed/illicit. Yun (2016) argued that the acoustic property of intensity rise predicts the site of perceptual vowel epenthesis; one salient trigger of an intensity rise is a release burst or frication noise. Related results have also previously been found for Korean. De Jong and Park (2012) presented Korean learners of English who had little exposure to spoken English with VC syllables in both a syllable counting task and an identification task. Results showed that the participants were more likely to categorize a stimulus item as containing two syllables or identify a final vowel for VC tokens containing a fricative or a notable stop burst. Thus, there is clear support for the idea that stop burst releases play a causative role in the illusory vowel effect in cross-linguistic speech perception, when the release burst is contextually inappropriate according to the listener's native language.

This conclusion suggests an alternative interpretation rather than phonotactically illicit sequences or illegitimate syllable contact (Dupoux et al. 1999; Kabak and Idsardi 2007): perhaps [+noisy] segments explain the illusory vowel effect in previous studies, where [±noisy] indicates the presence/absence of an audible release burst or other frication noise. The [noisy] status and reported discrimination in Dupoux et al. (1999) and Kabak and Idsardi (2007) are shown for various clusters in (4).

(4) *Previous results on illusory vowel effect, by* [*noisy*] *status and discrimination*

Study	Condition	Example	Burst/frication	Discrimination
Dupoux et al. 1999	main	[ebzo]~[ebuzo]	[+noisy]	poor
Kabak and Idsardi 2007	control	[pʰakt̚ʰa]~[pʰakʰʊtʰa]	[−noisy]	good
Kabak and Idsardi 2007	legal coda	[pʰakn̚a]~[pʰakʰʊna]	[−noisy]	good
Kabak and Idsardi 2007	illegal coda /gm/	[pʰagm̚a]~[pʰagʊma]	[−noisy]	good
Kabak and Idsardi 2007	other illegal coda	[pʰactʰa]~[pʰacʰɪtʰa]	[+noisy]	poor

It turns out that in Kabak and Idsardi's (2007) stimuli, coda illegality was almost fully confounded with the presence/absence of an audible release burst. There was

one exception, the /gm/ cluster, which was unreleased but not a legal coda. For this cluster, Korean listeners nonetheless exhibited good discrimination. Thus, (4) shows that previous results *are* consistent with the hypothesis that the illusory vowel effect is caused in whole or in part by the presence of audible burst/frication noise in a contextually inappropriate position.

We have given several arguments in the preceding paragraphs that lead naturally to this interpretation, but it has yet to be confirmed directly in a controlled experiment where the presence/absence of a release burst is directly manipulated. In the experimental sections of this paper, we will do exactly that. However, prior to that, we review previous studies on the Korean adaptation pattern for the medial clusters in (1), beginning with Boersma and Hamann's paper, "Loanword adaptation as first language phonological perception."

2.2 Loanword adaptation as first language phonological perception

Boersma and Hamann (2009) offer a general theoretical proposal that loanword adaptation is driven entirely by native language perception. In support of this general proposal, they provide a detailed theoretical account of Korean loanword adaptation. According to Boersma and Hamann, the contrast between native phonology (1a–c) and loanword phonology (1d–f) falls out from their approach without any special provision for loanwords. In particular, they claim that the epenthetic vowel in borrowed stop-nasal and fricative-stop clusters arises from a mismatch between the Korean listener's perceptual expectations and structural/acoustic properties of the source form.

Boersma and Hamann's analysis is couched in Bidirectional OT, a variant of OT which includes both a UR-SR mapping as well an analogous mapping between SR and the acoustic form AF. The name Bidirectional indicates that the same grammar regulates both production and perception. A key component of this theory is *cue constraints*, which regulate the /SR/↔[AF] mapping in the same way that faithfulness constraints regulate the |UR|↔/SR/ mapping. For example, the cue constraint */+nas/↔[_] is violated whenever a nasal segment in the SR (+nas) corresponds to silence (_) in AF. This cue constraint is what blocks the (silent) stop closure in English *picnic* from being parsed as /p^{h}iŋ.nik/ by Korean listeners. The more perceptually faithful parse, /p^{h}ik.nik/, is ruled out by the undominated markedness constraint SYLLCON (Syllable Contact Law: no sonority rise from coda to onset). The remaining parse /p^{h}i.k^{h}ɨ.nik/ violates the cue constraint */ɨ/↔[_] ('central high vowels should not correspond to acoustic silence'); this is the optimal parse because */ɨ/↔[_] is low-ranked in Korean. Thus, given the AF for English *picnic*, [_^{ph}ɪ̯$^{k̚}$_.nɪ̯$^{k̚}$_],[3] Boersma and Hamann's theory explains how Korean listeners recover the SR /p^{h}i.k^{h}ɨ.nik/ and the corresponding UR |p^{h}i.k^{h}ɨ.nik|.

A unique property of this theory is that it can yield 'epenthesis' in perception, but 'assimilation' in production. When the grammar is supplied with the UR |p^{h}ik$^{(h)}$nik|,

[3]The underscores correspond to stop closures, while superscript segments indicate the formant transition. Aspiration is indicated by ^{h} as usual. For further details on AF/notation, the reader can consult Boersma and Hamann (2009).

the faithful SR is ruled out by undominated SYLLCON, as before. The epenthetic candidate /p^{h}i.k$^{(h)}$i̯.nik/ is ruled out by undominated DEP-V, while the deletion candidates /p^{h}i.k$^{(h)}$ik/ and /p^{h}i.nik/ are ruled out by mid-ranking MAX-C. The best remaining option is to violate the low-ranked manner faithfulness constraint IDENT[NAS], yielding the nasality assimilation SR /p^{h}iŋ.nik/. This SR is then realized by the AF [_^{ph}iŋ.nɪ̚^{k}_].

Boersma and Hamann's theory straightforwardly accounts for the adaptation of [+noisy] sequences like *diskette*. They assume an undominated cue constraint which forces audible frication noise to be interpreted as such. Then, just as with *picnic*, the [+noisy] (fricative) segment is blocked from being parsed as a coda segment, and low-ranking */i̯/↔[_] is violated to yield an epenthetic parse in perception. In contrast, the underlying features that could yield a [+noisy] segment are neutralized in the coda in production.

For stop-stop sequences like *chapter*, they predict a different outcome from stop-nasal and fricative-stop clusters. Boersma and Hamann assume that, like *picnic*, the AF of such loanwords includes a [−noisy] initial stop. In this case, the 'faithful' SR /c^{h}ɛp.t^{h}ʌ/ is recovered. The essential difference is that this item does not violate the Syllable Contact Law (i.e. Korean allows coda segments to be of equal sonority as the following onset, just not lower). Therefore, unlike with *picnic*, the 'faithful' SR is not blocked, so there is no need to consider the epenthetic parse.

The core of Boersma and Hamann's account lies in the 'intermediate' SR representation: epenthesis is ruled out in the production direction because the (UR↔SR) faithfulness constraint DEP-V is undominated; epenthesis in the perceptual direction is tolerated because the corresponding (SR↔AF) cue constraint */i̯/↔[_] is low-ranked. For stop-nasal clusters, the undominated structural markedness constraint SYLLCON dooms the faithful parse (*picnic* ∼> */p^{h}ik.nik/), while for fricative-stop clusters, an undominated cue constraint (*/+asp/↔[_]) does the same. Only in stop-stop clusters, where Boersma and Hamann assume a [−noisy] initial stop, do they predict faithful adaptation. In this way, their theory admirably accounts for both the native Korean phonology in the production direction (1a–c) and the loanword phonology (1d–f) as native language perception.

At least one aspect of Boersma and Hamann's account receives external support. Their account crucially assumes that */i̯/↔[_] is low-ranked, which means that a phonological vowel can be realized as silent, and conversely silence can be parsed as a high vowel. As it turns out, Korean exhibits an optional high vowel devoicing process, which is likely after sibilants and aspirated stops (Mo 2007). Koreans are therefore accustomed to hearing [Ci̯] sequences in which many of the acoustic signatures of a vowel are suppressed. This fact will assume great significance in our discussion of the illusory vowel effect; in the Bayesian account we describe later, the availability of high vowel devoicing is what allows phonotactics and phonetics to converge on an (illusory) vowel parse.

However, Boersma and Hamann's account rests on two assumptions that do not have consistent support in the literature. The first is that the Korean adaptation system for medial clusters is in fact as shown in (1). Kabak and Idsardi (2007) are in agreement with Boersma and Hamann that stop-stop clusters are adapted without epenthesis, while fricative-stop clusters are adapted with epenthesis. Where they dis-

agree is stop-nasal clusters; Kabak and Idsardi claim that the native nasality assimilation process applies rather than vowel epenthesis (*picnic* → [pʰiŋ.nik̚]). To evaluate any theory relating speech perception and loanword adaptation, it is essential to know both sets of facts. Therefore, we begin with a loanword corpus study. To anticipate briefly, Kabak and Idsardi's (2007) summary of the facts was incorrect owing to a morphological confound which was unknown at the time of their study; the loanword data are exactly as Boersma and Hamann claimed and we reported in (1).

The second assumption that Boersma and Hamann made is that the English pronunciation for medial stop-stop and stop-nasal clusters (*chapter*, *picnic*) never includes an audible stop burst release. In this case, Boersma and Hamann seem to adopt an assumption based on what was known at the time from the primary literature. However, a recent study of spontaneous speech suggests that American speakers produce audible release bursts in both types of cluster about 30% of the time (Davidson 2011). If an audible release burst in medial clusters contributes to the illusory vowel effect as per Boersma and Hamann's account, the most straightforward prediction would be that medial stop-stop clusters are adapted variably (as with word-final stops; Kang 2003). It is important to know whether this is true; so the loanword corpus study in the next section also focuses on stop-stop clusters. After controlling for the morphological confound, it appears that stop-stop adaptation is approximately categorical as shown in (1), *despite* the non-negligible rate of audible stop release bursts in the source forms.

The account of Boersma and Hamann (2009) falls into a more general class of theories addressing the relation between speech perception and loanword adaptation. These accounts have the common property that they identify loanword adaptation with the most perceptually similar, phonotactically licit form in the adapter's native language. As Boersma and Hamann are careful to point out, phonological perception is about identification rather than discrimination. Thus, for example, the fact that Koreans adapt English *picnic* with vowel epenthesis does not imply that Koreans are unable to discriminate medial [k̚n] from medial [kʰɨn]—consistent with Kabak and Idsardi (2007), who found that Koreans exhibit good discrimination of exactly this contrast. Rather, the adaptation of *picnic* with epenthesis is supposed to tell us that Koreans identify [kʰɨn] as more similar to [k̚n] than any other permissible medial sequence—in particular the [ŋn] sequence that results when Koreans apply their native phonology to an underlying medial *kn* sequence. More generally, 'perceptual accounts' like Boersma and Hamann all predict that if (illegal) sequence X is adapted as (legal) A rather than (equally legal) B, it must be because X is more similar to A than X is to B. This prediction follows directly from the core assumption of perceptual accounts, that loanword adaptation is fundamentally driven by speech perception as shaped by the native phonological grammar. In Experiment II, we test this prediction with an ABX task, in which Korean listeners are presented with a potential loanword (nonce word containing an illicit sequence), and asked whether it sounds more like an epenthetic or an assimilatory repair. For example, listeners might hear *pa*[ŋn]*a* ∼ *pa*[kʰɨn]*a* ∼ *pa*[k̚n]*a*. Since Koreans adapt medial *kn* with epenthesis, the prediction is that they should perceive *pakna* as more similar to the epenthetic repair *pa*[kʰɨn]*a* than to the assimilation repair *pa*[ŋn]*a*.

To anticipate briefly, we find that the perceptual account is falsified in exactly this case. The collection of results in this paper therefore paints a nuanced picture of

the relation between speech perception and loanword adaptation. In some cases, the illusory vowel effect is sufficient to explain vowel epenthesis in loanword adaptation (e.g. in *diskette* the epenthetic vowel is forced by perception). However, in other cases the loanword adaptation system is markedly more categorical than perception alone would suggest (cf. Park and de Jong 2008; de Jong and Cho 2012). Some cases of vowel epenthesis cannot originate from the illusory vowel effect alone, and in other cases adapters do not epenthesize where an illusory vowel might be expected. In the Discussion section, we give a Bayesian account of speech perception which shares numerous key insights with Boersma and Hamann's account, but differs with respect to the identification of stop-stop and stop-nasal clusters above. We conclude with some thoughts on the nature of the relation between speech perception and loanword adaptation.

2.3 The adaptation of medial clusters in Korean loanwords

We turn now to establishing the loanword facts. As noted above, there is disagreement in the literature. Boersma and Hamann (2009) present the loanword adaptation facts as we showed them in (1), but Kabak and Idsardi (2007) claim differently: they suggest stop-nasal clusters are adapted with nasality assimilation rather than vowel epenthesis:

> p. 31: For example, loans such as *walnut*, *Telnet*, *Big Mac*, and *Pacman* are sometimes produced as having undergone the lateralization and nasalization rules where applicable by Korean speakers (e.g., *wa*[l.l]*ut*, *te*[l.l]*et*, *bi*[ŋ.m]*ac*, *Pa*[ŋ.m]*an* etc.)
> p. 47: The way loan words containing [k.m] and [l.n] are adapted in Korean (e.g., *Pa*[ŋ.m]*an* from *Pacman*; see Kang, 1996) gives us every reason to expect that native phonological rules may affect perception.

In the next major section, we conduct a loanword corpus study to resolve this discrepancy. Prior to this, we review Oh's generalization on the adaptation of morphologically complex words, which we will show plays into it.

Oh (2012) investigated the adaptation of morphologically complex sequences, such as noun-noun compounds and phrasal verbs. She argued that these forms are adapted on a component-by-component basis, after which regular Korean phonological processes apply. For example, *all in*$_E$ is adapted as /ol-in/ because *all* is adapted as /ol/ and *in* is adapted as /in/. Resyllabification and liquid allophony then apply, yielding the observed mapping *all in*$_E$ $\sim>$ [orin]$_K$. According to the alternative hypothesis that *all in*$_E$ is adapted as a single component, the predicted adaptation would be *[ol:in]$_K$, since medial English *l*s are customarily adapted into Korean geminate [l:].

An even more compelling argument comes from the adaptation of phrasal verbs like *make up* and *back up*. Oh found that the verb in the complex form is adapted with a final epenthetic vowel if and only if a final epenthetic vowel occurs in the adaptation of the simplex form:

(5)

	simplex form and adaptation	complex form and adaptation
epenthesis	*start* ~> [sɨtʰatʰɨ]	*start up* ~> [sɨtʰatʰɨʌp̚]
no epenthesis	*back* ~> [pɛk̚]	*back up* ~> [pɛgʌp̚]
variation	*make* ~> [mɛikʰɨ], [mɛik̚]	*make up* ~> [mɛikʰɨʌp̚], [mɛikʰʌp̚]

It follows from this theory that English compounds containing stop-nasal clusters across the compound boundary will be subject to nasality assimilation:

(6)

compound		Korean UR		Korean SR
Big Mac	~>	/pik-mɛk/	→	[piŋmɛk̚]
Pacman	~>	/pʰɛk-mɛn/	→	[pʰɛŋmɛn]

Note that the forms in (6) were the ones that Kabak and Idsardi (2007) invoked as evidence that stop-nasal clusters are adapted with the native, assimilatory process rather than with epenthesis. However, in light of Oh's generalization, there is a morphological explanation for why these particular loanwords were adapted with nasality assimilation: it is because they are both compounds. The practical import of Oh's theory is that loanwords which are likely to be parsed as morphologically complex by Korean adapters should be discarded from analyses aimed at establishing purely phonological adaptations. Therefore, the next section carefully distinguishes unimpeachably simplex forms from those that are potentially complex.

3 Corpus study: Adaptation of medial noncontinuant clusters

To investigate the adaptation pattern of medial noncontinuant clusters, we performed a search of the NAKL loanword list. This corpus consists of 2462 loanwords derived from a loanword list published by the National Academy of the Korean Language (NAKL 1991). The loanwords were collected from six daily newspapers and nine magazines published in Korea in 1990 (NAKL 1991:159–237). The English pronunciation of the source form was supplied via automatic lookup from the CELEX database for about 1800 English pronunciations (DISC transcription); the remaining English pronunciations were manually supplied by the first author following the existing CELEX conventions. All searches were performed using the Linux grep utility.

As the NAKL list is partly prescriptive, we supplemented the forms it found with Google counts. For example, to obtain counts for *doctor*, we searched https://www.google.co.kr/ for "닥터 doctor". The Hangeul form 닥터/takt^{h}ʌ/ is the (non-epenthetic) form that is listed in NAKL. The English form *doctor* is included in the search to remove spurious hits (Oh 2012). We also obtained hit counts for the corresponding epenthetic form, i.e. by searching "다크터 doctor". The search results for stop-stop clusters with an intervening syllable boundary in the English form are shown below in (7) and (8), divided by whether the form could plausibly be adapted component-by-component (or is otherwise suspect, cf. *vodka*):

(7) *Stop-stop clusters in simple forms: Near-categorical absence of epenthesis*

English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])	English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])
active	액티브	–	401.0	nectar	넥타	–	26.0
adapter	어댑터	–	494.0	octave	옥타브	–	61.0
bacteria	박테리아	–	140.0	optimist	옵티미스트	–	7.5
chapter	챕터	0.5	249.0	optimization	옵티마이제이션	–	9.0
collector	컬렉터	–	512.0	peptide	펩타이드	–	129.0
connector	커넥터	–	526.0	perspective	퍼스팩티브	–	42.7
director	디렉터	–	355.5	projector	프로젝터	–	362.0
directory	디렉터리	–	256.0	rugby	럭비	–	201.0
	디렉토리	–	332.0				
doctor	닥터	–	456.0	scripter	스크립터	–	144.0
electric	일렉트릭	–	361.0	sector	섹터	–	101.0
electron	일렉트론	–	9.7	specter	스펙터	–	472.0
hectare	헥타르	–	3100.0	spectrum	스펙트럼	–	323.0
helicopter	헬리콥터	–	352.0	tractor	트렉터	–	271.0
instructor	인스트릭터	–	190.0	victory	빅토리	–	169.0
napkin	냅킨	3.6	156.0				

(8) *Stop-stop clusters in morphologically complex forms: Variability*

English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])	English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])
cocktail	칵테일	0.6	436.0	*knock down*	넉다운	–	49.4
football	풋볼	–	471.0	*popcorn*	팝콘	–	354.0
necktie	넥타이	7.7	3010.0	*quick dos*	퀵 도스	–	22.3
				ragtime	랙타임	10.4	25.7
ad balloon	애드벌룬	4.6	–	*sidecar*	사이드카	32.1	–
feedback	휘드백	1.1	–	*vodka*	보드카	228.0	0.1
	피드백	398.0	69.2				

Evidently, the 29 forms shown in (7) are adapted without epenthesis either categorically or nearly so. For the 11 potentially complex forms in (8), NAKL indicates that 4 are adapted with epenthesis: *ad balloon*, *feedback*, *sidecar*, and *vodka*. These are the only 4 forms for which Google gave higher hit counts for the epenthetic form, and in all besides *feedback* the ratio is nearly categorical.[4] NAKL

[4] An anonymous reviewer notes that these cases are compatible with Oh's generalization, e.g. *football* is adapted without epenthesis because *foot* is. Note further that *vodka* is of uncertain provenance—it may have been loaned directly from Russian into Korean, or it may have been loaned from English, or it may

indicates that the other plausible complex forms (*cocktail*, *football*, *necktie*, *knock down*, *popcorn*, *quick dos*, and *ragtime*) are adapted without epenthesis, and Google counts suggest this is categorically or near-categorically observed, except for *ragtime*.

From these forms, we make the following generalizations:

- stop-stop clusters in morphologically simplex forms are adapted without epenthesis (7)
  - even when the initial stop is voiced, which occurs just once, in *rugby*
  - (but see fn. 5 for non-conforming data that does not come from NAKL)
- stop-stop clusters in morphologically complex forms may be adapted variably (8)
  - these cases appear to follow Oh's generalization

The same process is repeated with stop-nasal clusters, yielding the forms in (9) and (10):

(9) *Stop-nasal clusters in simple forms: Always adapted with epenthesis*

English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])	English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])
badminton	베드민턴	359.0	–	litmus	리트머스	32.2	–
	배드민턴	360.0	–				
brinkmanship	브링크맨십	0.1	–	magnesium	마그네슘	219.0	–
cadmium	카드뮴	46.0	–	magnet	마그네트	68.4	–
commitment	코미트먼트	0.6	–	partner	파트너	516.0	–
	커미트먼트	18.8	–				
dogma	도그마	368.0	–	picnic	피크닉	365.0	–

(10) *Stop-nasal clusters in complex forms: Variability*

English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])	English	NAKL	Kilohits (w/ [ɨ])	Kilohits (no [ɨ])
nickname	닉네임	–	356.0	oatmeal	오트밀	132.0	–

Out of these 12 forms, NAKL indicates the simplex forms are adapted with medial epenthesis, and the Google hit counts are in complete agreement. However, *nickname* is plausible as a component-by-component adaptation, as is *oatmeal* (which is adapted with epenthesis), and again the Google counts categorically agree with the NAKL form. Thus, there are 10 cases of simplex stop-nasal clusters adapted with epenthesis, and 0 with nasality assimilation. The following generalizations can be made:

- stop-nasal clusters in morphologically simplex forms will be adapted with epenthesis (9)
  - even when the initial stop is voiceless, e.g. *litmus*
- stop-nasal clusters in complex forms may be adapted variably

have been loaned indirectly through Japanese. Moreover, *vodka* is the diminutive of *voda* 'water', so it is potentially complex. For these reasons, we regard *vodka* as an unreliable data point.

- in these cases, which are rare, adaptation seems to follow Oh's generalization

In short, the simplex adaptation pattern is exactly as shown in (1), and exactly as Boersma and Hamann (2009) assumed.[5] The adaptation pattern that Kabak and Idsardi (2007) report was based on a subset of complex forms, prior to the publication of Oh's theory of complex form adaptation. We turn to the online perception of these clusters.

4 Experiment I: [+noisy] segments cause the illusory vowel effect

As reviewed earlier, there are several reasons to expect that audible release bursts or other frication noise might play a causative role in the illusory vowel effect. However, this claim has not been directly tested in a laboratory experiment, where [±noisy] status was manipulated orthogonally to other variables. The closest previous study was Kabak and Idsardi (2007), which tested discrimination of vowel presence/absence; but in that experiment, stops were unreleased while fricatives and affricatives contained frication noise. Experiment I is closely modeled on the experiment in Kabak and Idsardi (2007), but it is not a direct replication. The key similarity is that Experiment I is a cross-linguistic AX discrimination task: both Korean and American listeners are given a Korean-licit A form and a Korean-illicit X form, and asked to say whether they are the same or different. One difference is that it focuses more narrowly on stop-stop, stop-nasal, fricative-stop clusters. Another difference is that the presence/absence of a stop burst is manipulated within the same cluster type. A final difference is that trials assess not only discrimination of the loanword/epenthetic repair (*pa*[kʰin]*a* ∼ *pakna*) but also discrimination of the native/assimilatory repair (*pa*[ŋn]*a* ∼ *pakna*).

4.1 Items

The source items vary factorially in the nature of the medial cluster, the place of articulation of one or both consonants in the medial cluster, and in the presence/absence of burst/frication noise in the initial consonant of the cluster. For each source item, there are two corresponding repair items, corresponding to vowel epenthesis and

[5] We are grateful to Michael Kenstowicz for providing verification and additional data from a prescriptive guide published by the Korean Broadcasting Society (KBS 1987). He reports that its data are largely in agreement with our generalizations. Specifically, he replicated the absence of epenthesis in the stop-stop forms we listed in (7) (except *chapter, instructor*, and *perspective*, which were not listed), and found the additional forms *alcaptan*, *cactus*, *lactose*, *lepton*, and *actin*. He reports one simplex stop-stop form that does contain epenthesis: *Victoria*. He replicated the stop-nasal findings for *badminton*, *cadmium*, *litmus*, *magnesium*, *magnet*, *partner*, and *picnic*; and reports epenthesis in the following stop-nasal forms: *acme*, *chutney*, *fitment*, *grapnel*, *admittance*, *admiralty*, *apartment*, and *magma*.

An anonymous reviewer provides the following observations. Simplex stop-stop: *Edgar* and *Baghdad* both adapted with epenthesis. Complex stop-stop: *bitcoin*, *notebook*, *Dogville*, *log book*, *drag hunt*, *dogfight* adapted with epenthesis; *Big Data* and *podcast* adapted without epenthesis. The reviewer states that these complex forms follow Oh's generalization. The reviewer further suggests that complex stop-nasal forms are predominantly adapted with epenthesis, e.g. *Rogue Nation, dragnet*, and *eggnong*. We confirmed with Google counts (Jan. 15, 2018) that *rogue, drag*, and *egg* are preferentially adapted with epenthesis, meaning that these forms too are predicted by Oh's generalization.

Table 1 Stimulus items are given in the columns marked source, epenth(esis), and assim(ilation). The initial four columns indicate the medial cluster, phonotactic status of the source item in Korean, presence/absence of a release burst, and manner

Cluster	Legal	Noise	Type	Source	Epenth.	Assim.
/pt/		−noisy	stop-stop	[$p^h ap\tilde{t}^h a$]	[$p^h ap^h it^h a$]	[$p^h att\tilde{t}^h a$]
		+noisy		[$p^h apt^h a$]	[$p^h ap^h it^h a$]	[$p^h att^h a$]
/pk/	* (place)	−noisy	stop-stop	[$p^h ap\tilde{k}^h a$]	[$p^h ap^h ik^h a$]	[$p^h ak\tilde{k}^h a$]
		+noisy		[$p^h apk^h a$]	[$p^h ap^h ik^h a$]	[$p^h akk^h a$]
/tp/	* (place)	−noisy	stop-stop	[$p^h at\tilde{p}^h a$]	[$p^h at^h ip^h a$]	[$p^h ap\tilde{p}^h a$]
		+noisy		[$p^h atp^h a$]	[$p^h at^h ip^h a$]	[$p^h app^h a$]
/tk/	* (place)	−noisy	stop-stop	[$p^h at\tilde{k}^h a$]	[$p^h at^h ik^h a$]	[$p^h ak\tilde{k}^h a$]
		+noisy		[$p^h atk^h a$]	[$p^h at^h ik^h a$]	[$p^h akk^h a$]
/kp/		−noisy	stop-stop	[$p^h ak\tilde{p}^h a$]	[$p^h ak^h ip^h a$]	[$p^h ap\tilde{p}^h a$]
		+noisy		[$p^h akp^h a$]	[$p^h ak^h ip^h a$]	[$p^h app^h a$]
/kt/		−noisy	stop-stop	[$p^h ak\tilde{t}^h a$]	[$p^h ak^h it^h a$]	[$p^h att\tilde{t}^h a$]
		+noisy		[$p^h akt^h a$]	[$p^h ak^h it^h a$]	[$p^h att^h a$]
/pn/	* (nasal)	−noisy	stop-nasal	[$p^h ap\tilde{n}a$]	[$p^h ap^h ina$]	[$p^h amna$]
		+noisy		[$p^h apna$]	[$p^h ap^h ina$]	
/tn/	* (nasal)	−noisy	stop-nasal	[$p^h at\tilde{n}a$]	[$p^h at^h ina$]	[$p^h anna$]
		+noisy		[$p^h atna$]	[$p^h at^h ina$]	
/kn/	* (nasal)	−noisy	stop-nasal	[$p^h ak\tilde{n}a$]	[$p^h ak^h ina$]	[$p^h aŋna$]
		+noisy		[$p^h akna$]	[$p^h ak^h ina$]	
/fk/	* (coda, seg)	+noisy	fric-stop	[$p^h afk^h a$]	[$p^h afik^h a$]	[$p^h ak\tilde{k}^h a$]
/θk/	* (coda, seg)	+noisy	fric-stop	[$p^h a\theta k^h a$]	[$p^h a\theta ik^h a$]	[$p^h ak\tilde{k}^h a$]
/sk/	* (coda)	+noisy	fric-stop	[$p^h ask^h a$]	[$p^h asik^h a$]	[$p^h ak\tilde{k}^h a$]
/ʃk/	* (coda)	+noisy	fric-stop	[$p^h a\int k^h a$]	[$p^h a\int ik^h a$]	[$p^h ak\tilde{k}^h a$]
/xk/	* (coda, seg)	+noisy	fric-stop	[$p^h axk^h a$]	[$p^h axik^h a$]	[$p^h ak\tilde{k}^h a$]
/ns/		−noisy	control	[$p^h ansa$]	[$p^h anisa$]	–
/ls/		−noisy	control	[$p^h alsa$]	[$p^h alisa$]	–

place/nasality assimilation, respectively. For example, the source item [$p^h atk\tilde{}^h a$] has an epenthesis repair item [$p^h at^h ik^h a$] and a place assimilation repair item [$p^h ak\tilde{k}^h a$]. The full set of items is given in Table 1, followed by a more detailed description.

In Table 1, the source and repair forms are listed in the final three columns. The preceding columns indicate relevant information: the source cluster being tested (leftmost), whether the source form is licit in Korean, and if not, the source of its illegality. Note that there are four distinct sources of ill-formedness in the source items. One comes from a phonotactically illicit stop-nasal sequence. These items contain '* (nasal)' in the legal column. One comes from a failure to obey place assimilation; these are marked '* (place)' in the legal column. For example, the /tk/ row is '* (place)' because coronal stops may assimilate in place before dorsal stops in Korean, but the source form does not. Some rows are marked '* (coda)'; these correspond to illegal coda items in the terminology of Kabak and Idsardi (2007)—the

initial consonant of the cluster is one which can never occur in coda position in Korean. Finally, some rows are marked '* (coda, seg)'; these source items all have fricative-stop clusters with fricatives that do not occur in Korean (namely /fk/, /θk/, /xk/).

Some control clusters (/ns/, /ls/) are included. This was done so as to include some trials that should be easy for all listeners. The place of articulation for cluster-initial stops and fricatives was varied factorially, so as to increase the total number of items. The place-of-articulation manipulation in fricative-stop clusters was intended to test the hypothesis that the presence of burst/frication noise itself is sufficient to trigger the percept of an illusory vowel in Korean listeners, despite being hosted on a segment which does not occur in Korean. Place of articulation was also varied in the final stop of stop-stop cluster items; note that the assimilation repair items blindly copy the place of C2 onto C1, yielding assimilation in the stimuli even in clusters where place assimilation is not reported in Korean or English (e.g. /kt/).[6] The main function of the place articulation manipulation in stop-stop clusters was to increase the number of distinct test tokens for each cluster type. We had also hoped to test the effect of English/Korean place assimilation processes on perception; we will briefly discuss in the Results section why the present study does not allow for a meaningful test. All stops in onset positions were aspirated; this was done because aspirated stops are the normal adaptation for English voiceless stops (also, Korean lenis stops are obligatorily tensified after obstruents; aspirated C2 stops avoid this). Similarly, coronal nasals were used in the stop-nasal stimuli to avoid confounds with the place assimilation process. However, in fricative-stop items, coda/manner neutralization and place assimilation were applied: the fricative was neutralized to a plain ([−noisy]) stop, and since the following stop is velar, it also assimilated in place, e.g. source [p^{h}ask^{h}a] → repair [p^{h}akk̚^{h}a]. A velar stop was used because it is otherwise unclear what the coda neutralization process would yield for these fricatives; the native neutralization process reduces /s/, /h/, and alveopalatals to [t̚] in coda, but not labial or velar fricatives since Korean does not possess these segments. In contrast to Kabak and Idsardi (2007), we omitted voiced stops and fricatives, largely to limit the number of hypotheses/factors being tested.

4.2 Recording and resynthesis

A single token of each phonetically distinct surface string in Table 1 was produced by a phonetically trained, male Korean-English bilingual in a clear speech style. The talker was born in Korea, but was immersed in North American English from the age of 5, and has native or native-like proficiency in both languages. To control prosody,

[6] Ahn (1998:99–107) states that noncontinuants (oral and nasal stops) may assimilate in place to following noncontinuants. However, Ahn reports that "[t]his kind of process occurs universally in casual speech rather than in careful speech or in dictation form" (p. 100). Furthermore, the place assimilation process is conditioned by place of articulation: coronals tend to assimilate before both labials and dorsals (/sin+pal/ → [simbal] 'shoes'; /tat+ki/ → [takk̚*i] 'closing'), while labials only assimilate before dorsals (/cip+kaps/ → [cikk̚*ap] 'price of a house'), and dorsals do not assimilate (/kaŋ+mul/ → [kaŋmul] 'river water', *[kammul]). Thus, the noncontinuant place assimilation is of a rather different character than the nasality assimilation and coda neutralization processes, which apply obligatorily within Accentual Phrases and without regard to the following consonant's place of articulation.

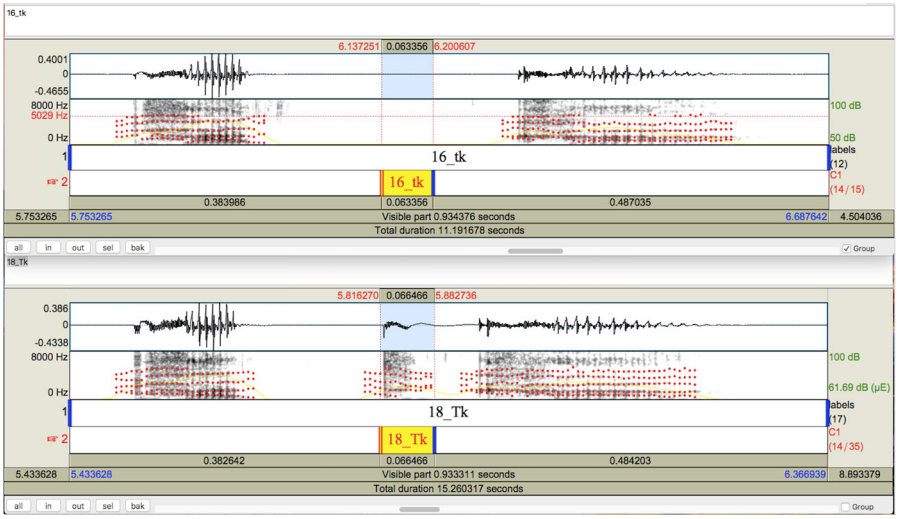

Fig. 1 Waveforms and spectrograms of both variants of the /tk/ cluster. *Top*: [p^{h}atk̚^{h}a] (−noisy). *Bottom*: [p^{h}atk^{h}a] (+noisy)

tokens were digitally edited/resynthesized after recording using Praat. The pitch contour across the initial [a] vowel was resynthesized to have a linear fall from 119 to 116 Hz, while the pitch contour across the final [a] vowel was resynthesized to have a linear fall from 82 to 67 Hz. These values represent the aggregate average contour across all tokens. For stimuli containing a medial vowel, the talker was required to produce a full [ɨ]. The pitch and duration of these medial vowels were not altered. In addition, the VOT of the initial syllable [p^{h}a] was digitally shortened (or in a few cases lengthened) to approximately 65 ms, if it was not already in the range 60–70 ms. This was the average for this speaker, and is typical for aspirated stops produced by Korean-English bilinguals (Daland and Oh 2011). Two native speakers of American English and one native speaker of Korean, all phonetically trained, inspected each token before and after resynthesis; tokens which were not deemed acceptable by any author were re-recorded and re-resynthesized.

Post hoc measurements of the burst release were obtained as follows. Praat's annotation tools were used to demarcate the burst release in the [+noisy] variants of each stop-initial 'source' stimulus. The burst release was extracted and measured for total energy, power, and intensity. Note that some of the energy may have originated from background noise, and this may have been different between different tokens because the duration of the measured region naturally differed slightly between tokens. To control for this, we 'yoked' the burst release to a corresponding temporal region in the corresponding [−noisy] token by vertically aligning the two sounds on screen. An example is shown in Fig. 1. The energy, power, and intensity of the burst release were defined as the difference (or ratio) between the measured energy of the [+noisy] region and that of the yoked [−noisy] region. These measurements are reported in Appendix A, and are used in a post hoc analysis below.

4.3 Lists

Stimulus pairs were formed by selecting the source stimulus from a row in Table 1, and another stimulus from the same row (epenthesis repair, assimilation repair; or the same stimulus for a 'same' trial), yielding 73 distinct stimulus pair types, not distinguishing order of the source of the repair form. Order of the source and repair form was counterbalanced within and across participants using a split-half design. That is, the critical ('different') pairs were subdivided into two equal subsets 1 and 2; the number of stop-stop, stop-nasal, and fricative-stop items was equal in each subset, and the repair types for each manner combination were counterbalanced across subsets, but the subdivision was otherwise random. In List A, subset 1 items occurred in the source-repair order and subset 2 items occurred in the repair-source order. In List B, subset 1 items occurred in the repair-source order, while subset 2 items occurred in the source-repair order. Thus, each List contained the same 'different' stimuli, but with items in opposite orders, and each List contained the same 'same' trials, with one repetition of each stimulus pair (73 trials in each list).

4.4 Participants

Korean listeners. A total of 37 participants were recruited from Chonnam National University in the Jeolla province of South Korea. Participants were English majors in the first or second year of study, with English spoken proficiency generally in the Intermediate range as informally assessed by the authors. The results from 7 participants were discarded because they lived abroad in an English-speaking country for 6 months or longer. Data from an additional participant was discarded owing to an abnormal response pattern (>25% 'different' responses to stimuli that were the same).

American listeners. A total of 43 fluent speakers of American English were recruited from the Linguistics/Psychology Subject Pool at UCLA and filled out a language background survey (adapted from Sundara and Polka 2008). Listeners were classified as English monolinguals (12F, 6M) if they had been exposed continuously to English since birth, reported limited or no proficiency in any language besides English, had never lived abroad for longer than a month, and had no more than 3 years of academic study of any foreign language. Listeners were classified as native English speakers (9F, 4M) if they had been exposed continuously to English since the age of 7 or earlier, had no previous exposure to Korean, but otherwise failed to meet the criteria for monolinguals (e.g. a Chinese-American bilingual). The remaining 12 participants' data was discarded owing to significant exposure to Korean.

4.5 Procedure

The Korean participants were tested en masse in a large, quiet room with individual workstations and sound-canceling headphones. The American participants were tested individually in a small, quiet room with an individual workstation and sound-canceling headphones. Verbal and written instructions were given in English to the

American group and in Korean to the Korean group. Otherwise, the experimental protocols were the same.

Participants were randomly assigned to List A or B, and the experiment was administered using Praat's MFC (Multiple-Forced Choice) Experiment protocol. The instructions indicated that listeners would hear two words, and should decide whether they were the same word, or different words. Participants were instructed to press 's' on the keyboard if they thought the two were the same word, and press 'k' if they thought they were different; they were also allowed to use the mouse to click on response boxes on the screen, marked "same 같음(s)" and "different 다름(k)". The main experimental block was preceded by a practice block, including 8 practice trials. The practice trials were the same across List A and List B, and consisted of the following items in the indicated order:

1. [pʰapˀsa]-[pʰamsa]
2. [pʰapˀsa]-[pʰapˀsa]
3. [pʰalla]-[pʰara]
4. [pʰacna]-[pʰacna]
5. [pʰalla]-[pʰalla]
6. [pʰanna]-[pʰaŋna]
7. [pʰacna]-[pʰaci̇na]
8. [pʰanna]-[pʰanna]

Feedback was given on practice trials (but not test trials), to help participants calibrate their expectations and to reduce variability/learning at the beginning of the test block. During the main experimental block, participants heard one instance of each test item, with the order of base and repair counterbalanced according to the List (A or B). Participants received a break after the practice block, after trial 32, and after trial 64.

In the main experimental block, trials began immediately after the preceding keypress or mouse click. A stimulus pair was played, preceded by 1500 ms of silence, and with 800 ms between the two items. Participants were given an unlimited time to respond, but were told in advance to respond as quickly as possible without sacrificing accuracy.

4.6 Results

All statistics and plots were obtained using R 3.0.2 (R Core Development Team 2014). Discrimination is reported by cluster type in Figs. 2, 3, and 4 (Fig. 2: stop-stop; Fig. 3: stop-nasal; Fig. 4: fricative-stop). The illusory vowel effect is apparent in all three figures (see captions).

The data are analyzed separately by cluster type. This was done both because of the unbalanced design (fricatives are inherent [+noisy]), and because the assimilation repair differs as a function of cluster type, rendering assimilation comparisons of dubious theoretical import. For each cluster type, the responses to 'different' trials were submitted to logistic mixed-effects regression, using the lme4 package. 'Same' trials were not included in statistical analysis, since the rate of 'different' responses to 'same' trials was so low as to introduce substantial sampling error (this is why logistic mixed-effects regression was used in preference to d').[7]

[7] Within-subjects d' is not reliable because the number of observations per condition is small (e.g. burstless stop-nasal: 3 same trials, 3 epenthesis trials). By-items d' also suffers from sparsity issues, owing to the

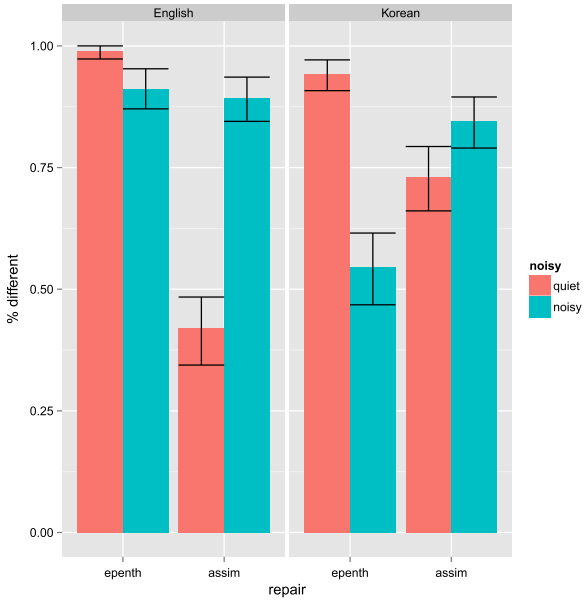

Fig. 2 Percentage of 'different' responses for stop-stop trials, subdivided by listener language (English: left; Korean: right), repair type (x-axis labels), and [noisy] level (paired bars: [+noisy] on the right). Error bars represent standard error of the mean, as estimated by the ggplot package's method. The illusory vowel effect is evident here in comparing the Korean epenthesis trials, [+noisy] vs. [−noisy]

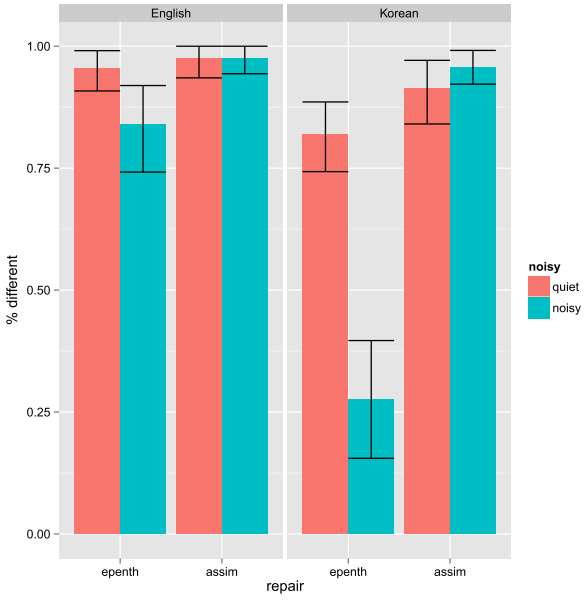

Fig. 3 Percentage of 'different' responses for stop-nasal trials, subdivided by listener language (English: left; Korean: right), repair type (x-axis labels), and [noisy] level (paired bars: [+noisy] on the right). Error bars represent standard error of the mean, as estimated by the ggplot package's mean_cl_boot method. The illusory vowel effect is evident here in comparing the Korean epenthesis trials, [+noisy] vs. [−noisy]

very low false alarm rate. It was typical for most 'same' items to receive a single 'different' response from an entire (Korean or American) population. Thus, individual items might have a sample false alarm rate of 1/31 for English and 1/29 for Korean; at these low numbers, the difference between the sample false alarm rate and the true false alarm rate can be large (cf. $z(1/31) \approx -1.849$; $z(0/31) = -\infty$). The sparsity issue can be partially ameliorated by aggregating over subjects by condition. For example, there are a total of 155 {English, fric-stop, epenthesis} 'different' trials, of which 132 received 'different' responses;

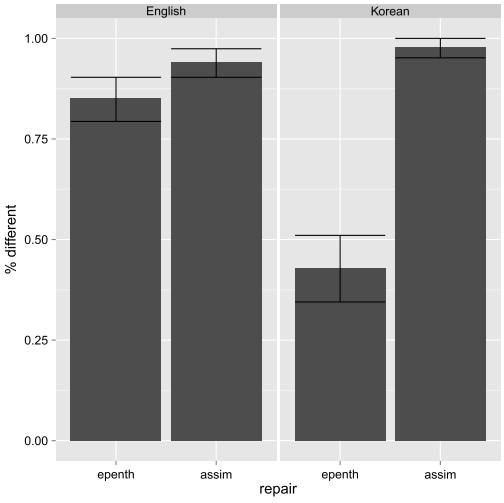

Fig. 4 Percentage of 'different' responses for fricative-stop trials, subdivided by listener language (English: left; Korean: right) and repair type (x-axis labels). Error bars represent standard error of the mean, as estimated by the ggplot package's mean_cl_boot method. The illusory vowel effect is evident here in comparing Korean epenthesis trials vs. American epenthesis trials

Stop-stop. In the stop-stop model, the dependent variable was the (log-odds of a) 'different' response. The analysis was conducted with dummy (contrast) coding, meaning that one level was specified as a baseline/reference level (indicated here by underlining), and coefficients represent the difference from that level. The simple fixed effects were *language* (English, Korean), *repair* (epenthesis, assimilation) and *noise* ([−noisy], [+noisy]); the fixed effects included a full factorial design, i.e. with the full set of interaction terms *language* × *repair* × *noise*. The random effects structure included random intercepts for *listeners* and *items*, with a random slope for language (per item—each item might have a baseline 'different' rate for each language) and a random slope for *repair* × *noise* (per listener—each listener might have a baseline 'different' rate for each repair type, both for [+noisy] and [−noisy] clusters). This is the 'maximal' well-defined random effects structure for this dataset (e.g. there could not be a *language* slope for listeners, because each listener belonged to only one language group). Significance is reported with respect to the reference case of English listeners discriminating epenthesis trials with [−noisy] items.

The simple effect of *language* was not significant, meaning that in the reference case Koreans did not respond 'different' more or less often than American listeners. The simple effect of *repair* was highly significant: all listeners were significantly less

there were also 155 {English, fric-stop, no repair} 'same' trials, of which 3 received 'different' responses, yielding $d' = z(132/155) - z(3/155) \approx 3.11$. The d' values obtained in this way are given below:

Korean	epenth	assim	English	epenth	assim
stop-stop	3.33/1.51	2.36/2.42	stop-stop	3.96/2.72	1.46/2.61
stop-nasal	2.49/1.40	2.94/3.72	stop-nasal	3.54/2.71	3.79/3.69
fric-stop	1.64	3.86	fric-stop	3.11	3.64

In stop-initial items, the [-noise] value is presented first, followed by a '/' and then the [+noisy] value. Since it was necessary to aggregate across both subjects and items to obtain these numbers, we do not know how to calculate a p-value. Indeed, the ability to straightforwardly assess significance is one of the reasons the logistic regression is reported in the main text, in preference to these d' values. In any case, the qualitative patterns are the same.

likely to respond 'different' to [−noisy] assimilation trials than epenthesis trials ($\beta = -8.58$; $z = -4.15$; $p < .001$). There was also a simple effect of noise: all listeners were significantly less likely to respond 'different' for [+noisy] items: ($\beta = -4.09$; $z = -1.99$; $p < .05$), but this effect was specific to epenthesis items, as evident from a *noise* × *repair* interaction: all listeners were significantly more likely to respond 'different' for [+noisy] assimilation items ($\beta = 7.03$; $z = 3.28$; $p < .005$). Finally, a *repair* × *language* interaction indicated that Korean listeners were advantaged at assimilation trials relative to American listeners ($\beta = 4.30$; $z = 2.28$; $p < .05$).

Stop-nasal. The same kind of analysis was conducted for stop-nasal items as for stop-stop items. There was no main effect of *language*, *repair*, or *noise*. There was a marginal but numerically large *language* × *noise* interaction: Koreans were especially disadvantaged on [+noisy] items ($\beta = -5.03$; $z = -1.87$; $p < .08$). As with stop-stop items, this effect was specific to epenthesis items, as evident from a *language* × *noise* × *repair* interaction: Koreans were specifically advantaged at [+noise] assimilation items ($\beta = 7.68$; $z = 2.02$; $p < .05$).

Fricative-stop. The same kind of analysis was conducted for fricative-stop items as for stop-nasal items, except that the fixed effects did not include noise or interaction terms (since fricatives are inherently [+noisy]). For this model, there was a simple main effect of *language*: Koreans were less likely overall to respond 'different' ($\beta = -3.61$; $z = -3.95$; $p < .001$). There was also a *language* × *repair* interaction: Koreans were more accurate in the discrimination of the assimilation repair ($\beta = 4.75$; $z = 3.31$; $p < .001$) as compared to the epenthesis repair.

Analysis of burst release and place effects? As noted earlier, the stimuli included multiple places of articulation. The primary purpose of this manipulation was to increase stimulus variety, as the experiment is quite repetitive. Note that it was logically possible that place of interaction would interact with the phenomena of interest—but the experiment was not designed to investigate place effects, because there was only one [+noisy] and one [−noisy] 'source' token per 'cluster'. Figure 5 indicates the variability between stop-initial clusters in triggering the illusory vowel effect.

The measure of burst energy used in Fig. 5 was inspired by Peperkamp et al. (2008), who found that the energy in the consonantal release of a word-final nasal was a strong predictor of whether Japanese listeners identified a following vowel. We investigated various other measures, such as the power difference, intensity difference, energy ratio, and power ratio. The results were qualitatively similar in all cases—/pn/ and /pk/ were outliers to the right, there was no systematic relation between C1 place of articulation, burst release/energy, and illusory vowel perception.

Evidently, burst loudness does not play the same, powerful conditioning role here as it did in Peperkamp et al.'s experiment. The energy of the burst releases varied over about an order of magnitude, with the loudest being the /p/ in the /pk/ and /pn/ clusters. However, these two items induced low and moderate rates of illusory vowel perception. The /kn/ cluster, whose [+noisy] token induced near-categorical rates of illusory vowel perception, had one of the weakest release bursts. Thus, there is no systematic relationship between the loudness of the burst and the rate of vowel epenthesis in these 9 tokens.

There are only 3 tokens per C1 place, and the token-to-token variation is large. An analogous lack of place effect was documented in Hwang (2011:27, Fig. 3.5), who

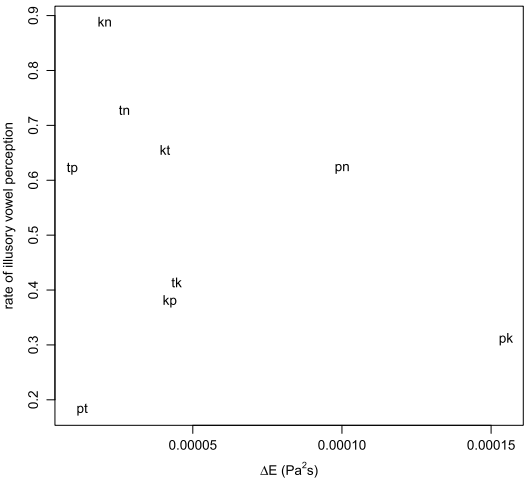

Fig. 5 Stimulus propensity to trigger illusory vowel perception, plotted against energy of the burst release. The x-axis represents the difference in energy between the burst release in a [+noisy] token, and a 'yoked' region of closure in the corresponding [−noisy] token—in other words how much louder the burst release was than the absence of a burst release (see Appendix A). The y-axis represents the proportion of 'same' responses to epenthesis trials for the [+noisy] variant by Korean listeners, an operational measure of illusory vowel perception

compared Koreans' rate of vowel perception in [bn]∼[bɨn] vs. [gn]∼[gɨn]. A salient point of comparison is that Hwang's perception study also used 1 source token for each of these two continua—so that the general effect of place cannot be deconfounded from phonetic variation specific to their source token. In short, there may or may not be a place of articulation effect (on burst release energy, and/or illusory vowel perception), but if there is, the effect size is small in comparison to the token-level variation in the stimuli here and in Hwang (2011).

4.7 Discussion

In many cases, discrimination was near-ceiling (e.g., >87.5%). For example, American listeners exhibited greater than 90% 'different' responses for stop-stop epenthesis trials, regardless of whether they were [+noisy] or [−noisy], as expected from the fact that the presence/absence of a vowel is contrastive in English. Similarly, Koreans exhibited greater than 90% 'different' responses for stop-nasal assimilation trials, regardless of whether they were [+noisy] or [−noisy].

The illusory vowel effect. A significant generalization that can be made about the results is that Koreans exhibited notably below-ceiling discrimination in all [+noisy]/epenthesis conditions. For example, in stop-stop items, Koreans were roughly on par with Americans (near-ceiling) at discriminating [−noisy] epenthesis trials (e.g. [pʰakt̚ʰa]∼[pʰakʰɨtʰa]). However, unlike Americans, Koreans exhibited a dramatic loss in discrimination of epenthesis trials including a [+noisy] release burst (e.g. [pʰaktʰa]∼[pʰakʰɨtʰa]). An analogous, though numerically even more dramatic drop is evident in the Korean [+noisy] stop-nasal epenthesis items. Since most other

phonetic properties of these stimuli were controlled, the stop burst is the presumptive cause for the drop in discrimination. Finally, Korean discrimination was well below 50% for fricative-stop epenthesis trials, but discrimination was over 75% in other fricative-stop cases (Korean assimilation; English assimilation and epenthesis). Thus, in all cases where Koreans heard [+noisy] epenthesis trials, discrimination was dramatically lower than in 'neighboring' conditions, whether different by repair, language, or noisiness.

It is worth noting that this holds equally for segments which are illegal in Korean (/f/, /x/). This fact is consistent with the idea that the [+noisy] feature is sufficient to trigger the illusory vowel effect on its own. However, it is also possible that this effect is mediated by native language perception. For example, the English labiodental fricative /f/ is standardly adapted into Korean as an aspirated labial stop /p^h/, so it is conceivable that listeners treated [$p^h afk^h a$]~[$p^h afik^h a$] discrimination trials as [$p^h ap^h k^h a$]~[$p^h ap^h ik^h a$] trials. But under either story, direct or indirect, the [+noisy] acoustic feature contributes to the illusory vowel effect.

Below-ceiling place discrimination. A notable aspect of the results was below-ceiling discrimination of the place assimilation items (stop-stop, assimilation repair). American listeners exhibited numerically much worse discrimination than Koreans. This language difference was not significant in the statistical model, and was instead explained by "ganging" of the *noise* × *repair* and *language* × *repair* interactions. However, the American and Korean populations taken as a whole do respond differently to this class of stimuli: pooling over participants and items, American listeners made 78 'different' responses out of 186 trials, while Korean listeners made 127 'different' responses out of 174 trials. A binomial test ($78/186 < (78+127)/(186+174) < 127/174$) indicates that American listeners exhibited significantly worse discrimination of place assimilation for [−noisy] items. This finding is consistent with Chang and Mishler (2012), who found that Korean listeners exhibit better perception of the place of articulation for word-final stops than American listeners do.

Thus, the stop-stop assimilation data reveal two important points. First, both American and Korean populations both exhibited below-ceiling discrimination of stop place assimilation when there was no release burst. Second, between these two populations, Korean listeners exhibited somewhat better discrimination than American listeners did, a finding which is compatible with a growing body of evidence suggesting that sensitivity to particular cues (such as stop burst releases) is modulated by language-specific experience (Abramson and Tingsabadh 1999; Hallé et al. 1999; Cho and McQueen 2006; Bohn and Best 2012; Chang and Mishler 2012). For example, Chang and Mishler suggest that because Korean stops are obligatorily unreleased word-finally (but stop place is still contrastive in that position), Koreans are highly practiced at extracting place of articulation without relying on the stop burst release. In contrast, because stop bursts are variably released in American English, Chang and Mishler expect that American listeners rely more heavily on stop burst information, and exhibit correspondingly worse perception of stop place when there is no release burst. The same kind of story may explain the current results: the main difference is that here, the stops are word-medial, and therefore presumably subject to greater rates of assimilation (and therefore offer less potential for contrast, on average). This

finding is also consistent with gestural timing research which has argued that stop-initial consonant sequences in languages like English can be subject to substantial overlap or reduction of the first consonant, which can have perceptual consequences similar to the phonological process of place assimilation (e.g. Byrd 1992; Jun 1996; Chen 2003). This may result in English listeners performing poorly on the assimilation trials for stop-initial sequences in this study.

Note that the results were not predicted by the direction of assimilation. For example, coronal stops may assimilate before dorsal stops in both English and Korean, while dorsals are not supposed to assimilate before coronals. If there is a perceptual asymmetry, it should be that the frequently assimilated case *pa*[kk͡ʰ]*a* ∼ *pa*[tk͡ʰ]*a* is more difficult to discriminate than the 'illegal' assimilation *pa*[tt͡ʰ]*a*∼*pa*[kt͡ʰ]*a*. However, for the specific tokens in this study, the asymmetry went the other way (although it went the predicted direction for coronal vs. labial). In short, our data are most consistent with the hypothesis that the within-place between-token variation is much larger than the between-place variation.

In summary, Experiment I replicated and refined previous work on the illusory vowel effect. Korean participants exhibited drastically reduced discrimination for the presence/absence of a following vowel just in the case that the preceding consonant was [+noisy], i.e. contained an audible stop burst release or other high-frequency frication noise. Thus, the novel empirical contribution of Experiment I is to causally link [+noisy] sounds with the illusory vowel effect, consistent with findings for word-final vowel perception in de Jong and Park (2012). Ultimately, it is surely the case that cross-linguistic speech perception is influenced by a number of factors; we do not mean to suggest that [+noisy] is the only one. Rather, we interpret these data as suggesting that when English source forms contain a [+noisy] sound, Korean loanword adapters are strongly encouraged by their native grammar to interpret that as an onset consonant which must be followed by a vowel. It follows from this that many instances of vowel epenthesis in loanword adaptation can be explained by the illusory vowel effect, namely sequences beginning with a fricative that cannot occupy a coda position in the borrowing language.

Despite this, the results of Experiment I suggest a mismatch between the speech perception and loanword adaptation in the adaptation of stop-nasal clusters, and possibly in stop-stop clusters as well. If the presence of an audible stop burst release is a primary determinant of the illusory vowel effect, then we might expect variability in the loanword adaptation pattern to mirror phonetic variability in the source language, of the same kind that Kang (2003) documented in word-final stops. The spontaneous production data that Davidson (2011) analyzed showed that Americans produce an audible stop release burst about 30% of the time for the initial stop in medial stop-stop and stop-nasal clusters. The innocent expectation should be that about 30% of stop-stop and stop-nasal clusters are adapted with epenthesis. However, the loanword adaptation pattern we documented in the previous section was much more categorical than this—it appears that simplex stop-nasal clusters are adapted categorically with epenthesis, while simplex stop-stop clusters are adapted near-categorically without epenthesis (see fn. 5). Then the mismatch is that some loanwords ([−noisy] stop-nasal) are adapted with epenthesis even though adapters seem able to hear that there is not a vowel in the source form while some loanwords ([+noisy] stop-stop) are

adapted without epenthesis even though adapters are expected to perceive an illusory vowel.

As Boersma and Hamann (2009) point out, loanword adaptation involves parsing a (possibly illegal) source form in terms of (legal) native categories. Therefore, loanword adaptation is an identification task rather than a discrimination task. It is possible for a listener to *hear* the difference between two sounds [x] and [y] in a context, and nonetheless identify them with the same phonological label /x/ (e.g. with modest training American listeners can discriminate prevoicing in phonologically voiced stops; Maye and Gerken 2000). Then perhaps the loanword adaptation data actually do reflect perception. That is, while Experiment I is informative about the phonetic basis of the illusory vowel effect, it does not directly evaluate the perceptual similarity relations that condition loanword adaptation.

A general tenet of perceptual accounts of loanword adaptation, including Boersma and Hamann (2009), is that the loanwords are adapted with the perceptually closest form that is phonotactically licit. Then, the fact that Korean speakers adopt *chapter* without epenthesis is supposed to indicate that the absence of epenthesis makes for a better perceptual match (for stop-stop adaptations) than the presence of epenthesis, while the fact that Korean speakers adopt *picnic* with epenthesis is taken to indicate the presence of epenthesis makes for a better perceptual match than the logical alternative, which for stop-nasal adaptations would be nasality assimilation. The logic of perceptual accounts suggests that Korean listeners' perceptual similarity judgments must be as schematized in (9):

(9)

a. stop-stop	similarity(*pakta*$_{E}$, [p^{h}akt̚^{h}a])	>	similarity(*pakta*$_{E}$, [p^{h}ak^{h}ɨt^{h}a])
b. stop-nasal	similarity(*pakna*$_{E}$, [p^{h}ak^{h}ɨna])	>	similarity(*pakna*$_{E}$, [p^{h}aŋna])
c. fric-stop	similarity(*paska*$_{E}$, [p^{h}asɨk^{h}a])	>	similarity(*paska*$_{E}$, [p^{h}akk̚^{h}a])

These predictions are the focus of Experiment II.

5 Experiment II: Perceptual similarity judgments partially mismatch loanword adaptation

The goal of Experiment II was to assess the perceived similarity between English-like, phonotactically-illicit source forms and their logically possible adaptations into Korean. The individual stimulus items were the same as in Experiment I, with one crucial difference. Instead of a discrimination task with AX *source-repair* trials, Experiment II was an identification task, with ABX *repair*$_1$-*repair*$_2$-*source* trials. On an individual trial, listeners heard two possible loanword adaptations, corresponding to an epenthetic repair and a native/assimilation repair (the A and B forms, counterbalanced). Finally, they heard the illicit source form (the X form), and judged whether it was more similar to the epenthetic repair or the assimilation repair.

5.1 Items

The same sound files were used as in Experiment I. However, their organization into trials (134 total) was different.

Stop-stop. For each [−noisy] stop-stop cluster, there were two critical trials: one in which the epenthetic repair occurred before the assimilation repair, and one in which the epenthetic repair occurred after the assimilation repair. In addition, there were four control trials in which the X form was identical to either the epenthesis or assimilation repair. These varied factorially in which repair occurred first (A), and whether the X form was identical to A or B. For illustration, the complete set of trials for the /kt/ cluster is given below:

- [pʰakʰɨtʰa]∼[pʰatt̚ʰa]∼[pʰakt̚ʰa] *critical trial* (*X illicit*), *epenthesis repair first*
- [pʰatt̚ʰa]∼[pʰakʰɨtʰa]∼[pʰakt̚ʰa] *critical trial* (*X illicit*), *assimilation repair first*
- [pʰakʰɨtʰa]∼[pʰatt̚ʰa]∼[pʰakʰɨtʰa] *control trial, epenthesis repair first*, $X = A$
- [pʰakʰɨtʰa]∼[pʰatt̚ʰa]∼[pʰatt̚ʰa] *control trial, epenthesis repair first*, $X = B$
- [pʰatt̚ʰa]∼[pʰakʰɨtʰa]∼[pʰatt̚ʰa] *control trial, assimilation repair first*, $X = A$
- [pʰakt̚ʰa]∼[pʰakʰɨtʰa]∼[pʰakʰɨtʰa] *control trial, assimilation repair first*, $X = B$

For each [+noisy] stop-stop cluster, there were the same 6 trials, except that both the X form and the assimilation repair form contained a stop-burst release (36 [+noisy] trials = 6 clusters ∗ 6 trials/cluster, 36 [−noisy] trials).

Stop-nasal. For each [−noisy] stop-nasal cluster, there were two critical trials (epenthesis repair first, assimilation repair first) and four control trials, exactly paralleling the stop-stop trials. For [+noisy] stop-nasal clusters, there were two critical trials. However, since the nasality assimilation repair is the same whether the stop is released or not, the control trials for [+noisy] stop-nasal clusters would be the same as for the [−noisy] case; thus, these control trials were not repeated (24 trials).

Fricative-stop. For each fricative-stop cluster, there were again six trials: two critical trials involving a phonotactically illicit X form, differing only in whether the epenthesis repair occurred before or after the assimilation repair, and four control trials, exactly parallel to the control trials for stop-stop and stop-nasal clusters (30 trials = 5 places ∗ 6 trials/place).

Control. For the /ns/ and /ls/ clusters, there were four trials each. These varied factorially in whether the epenthesis form occurred before or after the source form, and whether the X form was identical to the A form or the B form (8 trials = 2 clusters ∗ 4 trials/cluster).

5.2 Participants

Korean participants were recruited just as in Experiment I, from an English Phonetics course offered during the summer session at Chonnam National University in the South Jeolla province. There was a total of 17 Korean participants.

5.3 Procedure

The procedure was similar to Experiment I, except that the participants were instructed as follows: "You will hear three sounds. Say whether the last sound is more similar to the first sound or the second sound." These instructions were presented in Korean by a research assistant.

Table 2 Raw counts of listener judgements for critical trials in Experiment II. Row headers indicate cluster type of the X form. Column headers indicate whether the listener judged the X form to be more similar to a comparable epenthesis repair, or a native/assimilation repair (i.e. on each trial listeners chose one or the other)

	−noisy		+noisy	
	Epenth.	Assim.	Epenth.	Assim.
stop-stop	25 (12%)	179 (88%)	145 (71%)	59 (29%)
stop-nasal	18 (18%)	84 (82%)	88 (86%)	14 (14%)
fric-stop	–		153 (90%)	17 (10%)

The experiment was administered using Praat's MFC protocol. The sounds were presented over headphones at a comfortable listening level with an ISI of 300 milliseconds. Participants clicked on boxes marked either "first" or "second." Item order was randomized for each participant, but otherwise there was no between-subjects counterbalancing, since the full set of items was administered once to each participant. Participants were not instructed to respond as quickly as possible; however, they were also not given the option to replay the stimulus triple, so there was no special benefit to hesitating. Participants were given breaks after the 50th and 100th trial.

5.4 Results

Prior to analysis, the control trials were inspected informally. Recall that in control trials, listeners were presented with two licit Korean nonce forms, followed by an X form that is identical to one of the two forms (e.g. [p^hak^{h}it^ha]∼[p^hatt^ha]∼[p^hak^{h}it^ha]). Since these trials involve licit, contrastive stimuli, near-ceiling performance was expected. Indeed, the participants responded correctly to 92% of these trials, with individual accuracy ranging from 77% to 99%. Three individuals had accuracies in the range 77%–80%, and the remaining 14 were in the range of 87.5%–99%. The data from these lower-accuracy individuals was retained in the main analysis, on the basis of the following facts. (i) The errors in these lower-accuracy individuals appeared to be relatively consistent across time, i.e. it is not the case that they became fatigued after 100 trials and made most errors in the remaining 34 trials. (ii) The qualitative pattern on critical test trials was the same whether lower-accuracy individuals were included or not. The control trial data were not otherwise included in analysis.

Table 2 shows the raw counts for which repair was chosen as a function of cluster type and presence/absence of frication noise.

The data were submitted to mixed-effects logistic regression, with the (log-odds of) choice of repair (epenthesis or assimilation) as the dependent variable. Owing to the unbalanced design (fricatives are inherently [+noisy]), these data were again analyzed on a by-cluster-type basis. In all cases, [−noisy] was chosen as the reference levels for *noise*.

Stop-stop. *Noise* was the sole fixed effect; random intercepts were included for *listener* and *item*, each with random slopes for *noise*. The intercept was positive

($\beta = 2.23$; $z = 7.17$; $p < .001$), indicating a significant preference for the assimilation repair on [−noisy] stop-stop items, and the main effect of *noise* was negative and of a greater magnitude ($\beta = -3.40$; $z = -6.88$; $p < .001$), indicating a nearly opposite level of preference for the epenthesis repair on [+noisy] stop-stop items.

Stop-nasal. The same kind of analysis was conducted for stop-nasal items as for stop-stop items. The intercept was positive ($\beta = 1.54$; $z = 5.93$; $p < .001$), indicating a significant preference for the assimilation repair on [−noisy] stop-nasal items, and the main effect of *noise* was negative and of a greater magnitude ($\beta = -5.24$; $z = -5.16$; $p < .001$), indicating a strong preference for the epenthesis repair on [+noisy] stop-nasal items.

Fricative-stop. As fricatives are inherently noisy, these items included no fixed effects, and only random intercepts for *listeners* and *items*. The intercept was negative ($\beta = -3.05$; $z = -5.45$; $p < .001$), indicating a strong preference for the epenthesis repair.

5.5 Discussion

The results can be summarized as follows: in all and only the [+noisy] items, Koreans overall tend to judge phonotactically illicit items as more perceptually similar to the epenthesis repair that they apply in loanword adaptation. The magnitude of the preference varies across the three classes of stimuli: epenthesis is preferred over place assimilation in [+noisy] stop-stop items in 71% of trials, over nasality assimilation in [+noisy] stop-nasal items in 86% of trials, and over place assimilation (plus manner neutralization) in [+noisy] fricative-stop items in 90% of trials. Conversely, Koreans tend to judge phonotactically illicit [−noisy] items as more perceptually similar to the assimilatory repair that is applied in their native phonology. In [−noisy] stop-stop items, place assimilation is preferred over epenthesis in 88% of trials, and in [−noisy] stop-nasal items, nasality assimilation is preferred over epenthesis in 82% of trials. In short, the epenthetic repair preserves perceptual similarity better for illicit [+noisy] items, while the native/assimilatory repair preserves perceptual similarity better for illicit [−noisy] items.

The finding that [+noisy] stimuli are judged as more similar to the epenthetic repair is consistent with the interpretation we offered for the results of Experiment I: [+noisy] segments trigger the illusory vowel effect. This aspect of the results offers converging evidence for this interpretation.

However, the results are not entirely consistent with the predictions of perceptual accounts of loanword adaptation. Recall the prediction in (9) that stop-nasal clusters should be judged as more perceptually similar to the epenthetic repair (regardless of stop burst releasedness), while stop-stop clusters should be judged as more perceptually similar to the assimilation repair (regardless of burst releasedness). That was not the pattern that was found. Instead, burst releasedness does condition the similarity judgments: when the initial stop has an audible burst release, Koreans judge it as more similar to the epenthetic/loanword repair, and otherwise they prefer the assimilatory/native repair.

(10) *Schematic summary of results*

	cluster	*predicted repair preference*	*observed* [+*noisy*]	*observed* [−*noisy*]
a.	stop-stop	assimilatory	epenthetic	assimilatory
b.	stop-nasal	epenthetic	epenthetic	assimilatory
c.	fric-stop	epenthetic	epenthetic	–

There are two aspects of this pattern that we view as fatal to a purely perceptual account of loanword adaptation.

The first fact is that [−noisy] stop-nasal clusters are judged as more perceptually similar to the nasality assimilation repair, *similarity*(*pa*[kn̰]*a*, *pa*[ŋn]*a*) > *similarity*(*pa*[kn̰]*a*, *p*^{h}*a*[k^{h}ɨn]*a*). If loanword adaptation straightforwardly reflected perceptual similarity judgments, we would sometimes expect the adaptation pattern *picnic*$_{E}$ ∼> *pi*[ŋn]*ik*, **pi*[k^{h}ɨn]*ik*. Note that Boersma and Hamann's (2009) account does predict vowel epenthesis for stop-nasal clusters, so it correctly predicts the adaptation facts. But in their account, this happens because the undominated Syllable Contact Law constraint blocks the SR /piknik/, and the undominated cue constraint *[+nas]↔[_] blocks the stop closure in *picnic*$_{E}$ from being parsed as a nasal. Therefore, their account fails to make the correct prediction for perceptual similarity. We conclude that in this case, Koreans exhibit a categorical loanword adaptation pattern that is different from their own, apparently equally categorical perceptual similarity judgments. Koreans do not adapt [−noisy] stop-nasal loanword to the most perceptually similar form that is phonotactically licit.

The second fact is that [+noisy] stop-stop clusters are judged as more perceptually similar to the epenthetic repair. If loanword adaptation straightforwardly reflected perceptual similarity judgments, we would sometimes expect the adaptation pattern *chapter*$_{E}$ ∼> *cha*[p^{h}ɨt^{h}]*er*, **cha*[pt̰]*er*. Note that again, Boersma and Hamann's (2009) account makes the correct prediction for loanword adaptation in stop-stop clusters because it crucially assumes that this kind of cluster is always [−noisy]. It is straightforward to show that perceptual epenthesis is predicted for [+noisy] items under their account. Thus, Boersma and Hamann (2009) make the correct prediction for perceptual similarity, but fails to explain why epenthesis is blocked in adaptation of [+noisy] stop-stop loanwords. In the General Discussion, we will return to these two points.

A final, related point of interest is that Korean listeners judged *pa*[kn̰]*a* as more similar to *pa*[ŋn]*a* than to *pa*[k^{h}ɨn]*a*, i.e. in accord with the native nasality alternation. Our suspicion is that if American listeners were to participate in Experiment II, they would judge *pa*[kn̰]*a* as more similar to the epenthetic repair *pa*[k^{h}ɨn]*a* than to the nasality assimilation repair *pa*[ŋn]*a*. If our suspicion turns out to be correct, this would be powerful evidence for the hypothesis that native language alternations affect overt perceptual similarity judgments even in cases where listeners are able to discriminate all of the relevant items. We leave this possibility for future research.

6 General discussion

In this section, we summarize the key findings of the paper, and then take up two issues for further discussion. The first issue is how the experimental findings in this paper might be integrated into a larger theory of speech perception; we summarize how these results might be accommodated by a Bayesian model which views speech perception as a process of reverse inference, along the lines proposed in Wilson and Davidson (2013) and Durvasula and Kahng (2015). The second issue we take up is the relation between speech perception and loanword adaptation; the present results are inconsistent with the hypothesis that loanword adaptation reflects speech perception alone. We propose a hybrid theory which admits a role for a particular specification of Paradis and LaCharite's (2002) Preservation Principle, and then speculate on the dynamic mechanisms which might underlie the Preservation Principle.

6.1 Summary of key findings

The illusory vowel effect refers to a pattern of discrimination in which listeners exhibit poor discrimination for the presence/absence of a vowel that intervenes between a consonant cluster that is phonotactically illicit in their native language, e.g. poor discrimination of *ebzo* ~ *ebuzo* by Japanese listeners (Dupoux et al. 1999). This paper presents the results of two perception experiments focused on the illusory vowel effect in Korean, whose goal was to rigorously test the hypothesis that the loanword adaptation pattern derives straightforwardly from speech perception, as formalized most concretely in Boersma and Hamann (2009). The experiments focused on word-medial stop-stop (e.g. *patka*), stop-nasal (e.g. *pakna*), and fricative-stop (e.g. *paska*) clusters. The stimuli were structured as follows. For every cluster type, there is a (potentially phonotactically illicit) source form, a corresponding assimilation repair (reflecting how the source form might be altered to be pronounced according to native Korean phonological processes), and an epenthetic repair (reflecting how the source form might be adapted if it were a loanword from English). A crucial manipulation was the presence/absence of a burst release or other frication noise in the initial consonant of the source cluster ([+noisy]/[−noisy]); this kind of frication noise does not occur in Korean codas, but occurs variably in English stops that serve as the input for loanword adaptation.

The first, cross-linguistic AX discrimination experiment tested the extent to which Korean and American listeners could discriminate source forms (A) from the corresponding assimilation or epenthesis repair forms (X). Near-ceiling discrimination was observed in both populations in many cases. However, Koreans exhibited notably poor discrimination of the epenthesis repair for [+noisy] source forms (the illusory vowel effect). An additional finding was below-ceiling discrimination of place assimilation for [−noisy] source items in all listeners, although Koreans exhibited better discrimination than Americans.

The second, ABX identification experiment assessed Korean listeners' perceptual similarity judgments. In critical trials listeners were asked to judge whether the source form (X) sounded more similar to the corresponding native/assimilation repair, or the loanword/epenthesis repair (A, B). The pattern of results was clear: for

[−noisy] source forms, Koreans exhibited a strong preference for the assimilatory repair, corresponding to native Korean phonological processes; however, for [+noisy] source forms, Koreans exhibited a strong preference in the opposite direction, corresponding partially to the loanword adaptation pattern.

A final empirical contribution of this paper is careful documentation of the loanword adaptation pattern in these cases. With the exception of morphologically complex items, which can be adapted in a component-by-component fashion (Oh 2012), loanword corpus data suggest that medial stop-stop clusters are adapted near-categorically without epenthesis, while medial stop-nasal clusters are adapted near-categorically with epenthesis. This pattern is claimed in some previous papers without extensive documentation (e.g. Boersma and Hamann 2009), but appears to be misunderstood in other papers (e.g. Kabak and Idsardi 2007). In the interest of providing a more substantiated base to work from, we documented this pattern here using the NAKL loanword database and Google counts.

In summary, the results presented here document a powerful influence of stop release burst and other frication noise on the illusory vowel effect, in which listeners of a language exhibit poor discrimination for the presence/absence of a vowel in the midst of a consonant cluster that is phonotactically illicit in their native language. There is corpus evidence suggesting release bursts condition perceptual vowel epenthesis word-finally in Korean (Kang 2003); theories of cross-linguistic speech perception which assume attention to this cue (Wilson et al. 2014); theoretical accounts of Korean loanword adaptation which presuppose release bursts condition vowel epenthesis (Boersma and Hamann 2009); and behavioral studies suggesting that burst releases condition epenthetic vowel perception in other languages (Peperkamp et al. 2008); but to date it has not been demonstrated in the laboratory that release bursts condition perceptual vowel epenthesis in Koreans' perception of word-medial clusters. This result provides one more piece of empirical evidence for Dupoux et al.'s (1999) general proposal that listeners do sometimes perceive a vowel which is not 'there' in the acoustic signal. However, the present results refine Dupoux et al.'s (1999) claim by specifically isolating the presence of a release burst or other frication noise as a powerful causative factor in the illusory vowel effect (for Korean speakers).

6.2 Toward a quantitative model of non-native speech perception

To what extent is the illusory vowel effect driven by top-down/phonotactic expectations, versus mere acoustic matching? We will argue here that perception crucially involves both, in a principled way. Our approach is couched in Bayesian terms (Wilson and Davidson 2013), and integrates the proposal of Durvasula and Kahng (2015) that speech perception may be viewed as a form of inference, where the task is to recover the lexical representation.

Contemporary theories of speech perception are generally in agreement that perception is driven both by top-down expectations and bottom-up acoustic properties (e.g. McClelland and Elman 1986; Boersma and Hamann 2009). In certain Bayesian approaches to perception, the phonological parse that is assigned is the one that maximizes posterior likelihood given the acoustic input (e.g. Wilson and Davidson 2013).

This likelihood has two terms, one representing the likelihood of the phonological parse (the 'prior,' i.e. top-down expectations), and one representing the likelihood of the acoustic data being generated by the phonological parse (i.e. acoustic match). This kind of Bayesian approach formally encodes the hypothesis that speech perception should be sensitive both to acoustic details in the signal, and to phonotactic expectations derived from the listener's previous language experience. The distinction between phonotactically licit and illicit sequences is captured by assigning 'penalties,' whose effect on the log-probability scales linearly with the degree of ill-formedness (e.g. as in Hayes and Wilson 2008). As a result, sequences that are very ill-formed according to a language's phonotactic grammar can be assigned such large penalties that the probability of observing them is essentially zero. The consequence for the listener is that even if the observed sequence is an excellent acoustic match for an ill-formed parse, they will assign a more well-formed parse. We will adopt the proposal of Durvasula and Kahng (2015) that the 'parse' the listener wishes to recover consists of a lexical representation (i.e. a UR).

Let us begin with the word-medial fricative-stop sequence, *pasta*, as perceived by an idealized Korean listener. For clarity, we will adopt the following notation: >X< refers to the *acoustic* sequence produced by a talker and heard by the listener, while /X/ and [X] refer to potential phonological and phonetic parses the listener might assign to >X<. Then assume the listener has heard >pasta<. Strident fricatives are robustly cued by broadband high frequency energy that is internal to the fricative (Wright 2004); thus, the [+noisy] feature in >pasta< makes for an excellent acoustic match to any parse associating it with [s]. However, native speaker experience with Korean will have provided very strong evidence that [st^h] sequences simply do not occur within Korean words. The probability of an (/st^h/, [st^h]) parse is proportional to the product of these two terms: Pr(/st^h/,[st^h] | >st<) = Pr(/st^h/,[st^h]) × Pr(>st< | [st^h]) / *const* (where *const* is an arbitrary constant required to satisfy the axioms of probability, but which does not affect the *relative* likelihood of different parses). Since the former term is very nearly 0, the *pa*[st^h]*a* parse must be assigned very low probability *despite* providing an excellent acoustic match. One alternative parse is [tt̚^{h}], i.e. the [+noisy] acoustic feature is associated to an unreleased coda stop (a stop closure, in other words, silence). Here, although the phonotactic likelihood of [tt̚^{h}] is high, the acoustic match is exceptionally low (because stridency is known to be robustly cued). Thus, Pr(/st^h/, [tt̚^{h}] | >st<) = Pr(/st^h/, [tt̚^{h}]) × Pr(>st< | [tt̚^{h}]) / *const* is also very low, but for the opposite reason (poor acoustic match rather than phonotactic ill-formedness). Finally, let us consider the epenthetic parse *pa*[sɨt^h]*a*. Unlike medial [st^h], word-medial [sɨt^h] is phonotactically well-formed, so Pr([sɨt^h]) is high. But, just as in the [st^h] parse, the [+noisy] feature is an excellent acoustic match to [s]. Then the value of Pr(>st< | [sɨt^h]) boils down to whether the sequence >st< provides a good acoustic match to [sɨt^h], despite the absence of an overt vowel in the acoustic sequence. As shown in Mo (2007), high-vowel devoicing occurs quite frequently after lenis fricatives in Korean (around 70% of the time in her data), possibly because they are aspirated. Therefore, a 'silent' vowel may be more likely than not in this position: >st< does provide a good acoustic match to [sɨt^h]. Now both the phonotactic/top-down and acoustic/bottom-up terms in Pr(/sɨt^h/,[sɨt^h] | >st<) = Pr(/sɨt^h/,[sɨt^h]) × Pr(>st< | [sɨt^h]) / *const* are relatively high-probability. Faced with

>pasta<, the epenthetic parse *pa*[sɨt^{h}]*a* is superior to the assimilatory parse *pa*[tt̚^{h}]*a* owing to the poor acoustic match between >s< and [t̚], and the epenthetic parse is superior to the faithful parse *pa*[st^{h}]*a* owing to the phonotactic ill-formedness of [st^{h}] in Korean. Note that this analysis is essentially the same as that offered in Boersma and Hamann (2009), except that the present analysis is couched in Bayesian/probabilistic terms, rather than the bidirectional variant of OT that Boersma and Hamann employ.

Now let us apply similar reasoning to the perception of *pa*[kn̚]*a*. The faithful parse *Pr*(/kn/,[kn̚] | >kn̚<) is assigned a large penalty owing to the phonotactically ill-formed sequence [kn̚]. The assimilatory parse *Pr*(/kn/,[ŋn] | >kn̚<) is phonotactically licit, but does not offer an especially good acoustic match. Boersma and Hamann (2009) attribute this specifically to a poor acoustic match between sonorant [ŋ] and the (silent) stop closure associated with >k̚<. Finally, the epenthetic parse *Pr*(/k^{h}ɨn/,[k^{h}ɨn] | >kn̚<) is phonotactically acceptable. The question is, to what extent is this a good acoustic match? Mo (2007) indicates that high-vowel devoicing after aspirated consonants is almost as frequent as after the lenis fricative. Then the only defect of the [k^{h}ɨn] parse for >kn̚< is that the expected stop burst release is inaudible in the stimulus >kn̚<; owing to its general significance in onset positions the release burst should be robustly cued in Korean. Now the fact that Koreans prefer the assimilatory repair for [−noisy] stop-nasal items can be given a probabilistic interpretation: Koreans prefer the acoustic mismatch of >k̚< to [ŋ] over the acoustic mismatch of >k̚< to [k^{h}ɨ̥] with a certain degree of likelihood. Of course, when there is an audible stop burst release, >kn< will be an excellent match to [k^{h}ɨ̥n] because in this case the burst release is well-cued. This straightforwardly explains why the identification results in Experiment II exhibit a categorical 'flip' when there is a release burst. The same kind of explanation holds for stop-stop and stop-nasal clusters, except that in stop-stop clusters, the silence of a stop closure is always a good acoustic match for a stop parse.

In order for this approach to be fully predictive for the results in Experiments I and II, we must be able to link probabilities with behaviors. We propose that the following *linking assumptions* provide an excellent way to derive perceptual predictions from probabilistic models:[8]

- discrimination of two acoustic sequences will be poor when there is a unique phonotactically licit parse which provides a good acoustic match to both sequences
- discrimination will be good otherwise
- identification corresponds to the highest likelihood parse, regardless of whether it includes an excellent acoustic match

The implication of these points is that listeners will be able to discriminate two acoustically similar sound sequences if the best phonotactic parse differs between them, *or* they have the same most likely phonotactic parse but one provides a significantly better acoustic match than the other (cf. Best 1991, 1995; Best and Tyler 2007).

[8]To our knowledge, these assumptions have not been laid out explicitly in print. This is likely because the research traditions of non-native speech perception and stochastic phonology have had relatively little contact to date. However, we feel these assumptions are the natural way to connect the predictions of any probabilistic model of sound sequences with the results of speech perception experiments.

For example, a parse which implies a release burst will provide a discriminably worse match to a sequence which does not contain such a release burst. However, a parse which implies a vowel will *not* provide a discriminably worse acoustic match to a sequence in which vowel devoicing is licit—thereby explaining the illusory vowel effect in Experiment I, as well as why the perceptual similarity results in Experiment II were determined by the presence of the [+noisy] acoustic feature.

In summary, these examples illustrate how a Bayesian model of speech perception might integrate top-down phonotactic expectations based on language experience with bottom-up acoustic cues so as to account for both the discrimination and identification facts reported in Experiments I and II. The key is the proposal that poor discrimination of two acoustic sequences will arise just in case they have the same most likely phonological parse, *and* both acoustic sequences constitute roughly equal perceptual matches. In the present case, the prevalence of high vowel devoicing in Korean after fricatives and aspirated consonants implies that an (epenthetic) vowel in the midst of a voiceless obstruent-noncontinuant cluster will generally provide a good acoustic match (e.g. [pʰasɨtʰa] is a good acoustic match to >pʰastʰa<) despite the absence of overt vowel cues, so the illusory vowel effect is predicted in this case. However, good discrimination is predicted for both >pʰakn̊a<~>pʰakʰɨna< and >pʰakn̊a<~>pʰaŋna<, since the silence of the stop closure is a poor acoustic match to both a nasal and a stop+vowel parse. Thus, in these examples, we have described a theory of how phonotactic expectations and acoustic mismatch may work together to derive both the discrimination and identification of non-native sequences according to native phonotactics. This approach was inspired by Boersma and Hamann (2009), but goes beyond their account in several ways. One is that the current account is probabilistic, and is therefore well-suited to handle the variability that is ubiquitous in perceptual experiments. However, in principle it is straightforward to render Bidirectional OT stochastic—so this dimension of contrast is not theoretically crucial. The other, significant dimension of difference is that our account explicitly links the output of a probabilistic model with behavior in both discrimination and identification experiments. By design, Boersma and Hamann's account explicitly identifies loanword adaptation with native speech perception. Therefore, it cannot reconcile the perceptual similarity results of Experiment II—in which Korean listeners judge [kn̊] as more similar to /ŋn/—with the loanword pattern, in which [kn̊] is identified as /kʰɨn/. Boersma and Hamann's grammar can presumably be made to predict either outcome, but it cannot predict one outcome for loanword adaptation and the other for native speech perception, because the core premise of their account is that loanword adaptation is driven by (and therefore identical to) native speech perception. By positing that [kn̊] *is* perceptually closer to [ŋn] than [kʰɨn] for Korean listeners, our account straightforwardly explains how Korean listeners can hear the difference between [kn̊] and [ŋn] (Experiment I), yet identify [kn̊] with /ŋn/ (Experiment II). The remaining mystery is how Korean listeners can identify [kn̊] with /ŋn/ in Experiment II, yet identify [kn̊] with /kʰɨn/ in loanword adaptation. We address this issue in the next section.

6.3 On the relation between speech perception and loanword adaptation

We conclude by discussing the topic contained in the title of this paper: what is the relation between speech perception and loanword adaptation? The research strategy we employed in this paper was to test various predictions of the hypothesis that loanword adaptation more or less directly reflects speech perception, as formalized most concretely for Korean in Boersma and Hamann (2009).

Experiments I and II converge in suggesting that when Korean adapters hear a [+noisy] consonant, they are strongly biased to perceive a following vowel if there is not one there already. This finding is able to explain why vowels are epenthesized after fricatives in loanword adaptation: it suggests that vowel epenthesis is essentially forced by perception in these cases. It is further consistent with Kang's (2003) conclusions regarding the role of release bursts in the adaptation of English word-final stops. Our findings, along with those of de Jong and Park (2012) who show similar effects for [+noisy] environments in VC# syllables, demonstrate a clear link between speech perception and loanword adaptation. Just as Kang (2003) proposed, variability in whether the source stop is [+noisy] should predict variability in the perceptual parse, which translates straightforwardly into whether an epenthetic vowel is introduced in the loanword.

The variability of adaptation in the word-final environment is not consistent with the 'phonological' account of loanword adaptation advanced by Paradis, LaCharite, and colleagues. Their theory includes two core tenets: consistent/phonemic mapping, and the Preservation Principle. Consistent/phonemic mapping is the idea that a source phoneme should be invariably mapped to the same borrower phoneme. For example, English voiceless stops are generally adapted as aspirated stops into Korean. However, word-final voiceless stops are variably adapted: either as an aspirated stop followed by an epenthetic vowel, or as a lax stop (the aspiration contrast is neutralized in isolation forms, but can be distinguished by the addition of a case particle). This is a violation of consistent/phonemic mapping whose motivation is apparently to maximize phonetic/perceptual similarity. Korean offers another example of inconsistent mapping whose outcome is maximizing phonetic/perceptual similarity: initial /s/$_{\mathrm{E}}$ is adapted as lax [s]$_{\mathrm{K}}$ when it originates from a cluster, but as tense [s*]$_{\mathrm{K}}$ when it originates from a singleton onset, cf. *spy*$_{\mathrm{E}}$ ~> [sɨp$^{\mathrm{h}}$ai]$_{\mathrm{K}}$, *sort*$_{\mathrm{E}}$ ~> [s*ot$^{\mathrm{h}}$ɨ]$_{\mathrm{K}}$ (Oh 1996; Kim and Curtis 2002). English /s/ is shorter when in an onset cluster, and this length difference is the apparent basis for the differential mapping pattern. In short, Korean offers abundant evidence that consistent/phonemic mapping may be violated to increase phonetic/perceptual similarity.

The other tenet of the phonological account of loanword adaptation is the Preservation Principle. The Preservation Principle is a loanword-specific principle whose function is to explain why vowel epenthesis is preferred to native repairs in loanword adaptation. It is defined in Paradis and LaCharité (1997:384) as follows: "Segmental information is maximally preserved within the limits of the Threshold Principle [which prohibits phonologically large modifications]." This formulation is regrettably imprecise. For example, in some languages vowels may contrast with zero (cf. *cough*, *coffee*); the presence/absence of a segment is indis-

putably 'segmental information.' Vowel epenthesis in loanword adaptation can neutralize vowel∼zero contrasts in the source language, and it is not difficult to find examples of this in Korean loanword adaptation specifically. However, since Paradis and LaCharité intended the principle to explain why vowel epenthesis is the preferred repair for illicit phonotactics, the simple presence/absence of a source vowel cannot be the kind of segmental information they mean to protect. Rather, the generous interpretation is that the principle prevents segmental features that are contrastive in the source form being neutralized in the borrowed form.

This more precise formulation of the Preservation Principle is able to explain the loanword adaptation pattern in (1), repeated with minor enhancement in (11):

(11)

	cluster	example (NAKL 1991)	repair	alternative	PP says
a.	stop-stop	*cha*[pt]*er*$_E$ ∼> *cha*[p̚t͡ʰ]*er*$_K$	assimilatory	*cha*[p^{h}ɨt^h]*er*$_K$	p∼>p̚ok
b.	stop-nasal	*pi*[kn]*ic*$_E$ ∼> *pi*[k^{h}ɨn]*ic*$_K$	epenthesis	*pi*[ŋn]*ic*$_K$	*k∼>ŋ
c.	fric-stop	*di*[sk]*ette*$_E$ ∼> *di*[sɨk^h]*ette*$_K$	epenthesis	*di*[kk̚^{h}]*ette*$_K$	*s∼>k̚

In (11a), the 'assimilatory' adaptation (/p/$_E$ ∼> /p^h/$_K$ → [p̚]$_K$) is acceptable (even though the aspiration feature in the Korean adaptation is neutralized) because aspiration is not contrastive in English. Note that the adapted stop is obligatorily unreleased, whereas it may be variably released in English; but since stop releasedness is not contrastive in English, it isn't protected by the Preservation Principle. In (11b), vowel epenthesis enables the adapted form to contain a velar stop rather than a velar nasal. In English, stop∼nasal is a contrastive manner specification even in the coda position; failure to epenthesize would therefore cause Korean phonology to neutralize a segmental feature that is positively specified in the source form. And this holds regardless of stop burst releasedness. The story is the same for (11c): /s/$_E$ ∼> /s/$_K$ → [k̚]$_K$ would result from neutralizing the manner features of a coda fricative (and assimilating in place before a dorsal), but in English fricative∼stop is a contrastive manner specification even in the coda position. Vowel epenthesis allows the adapted form to have the same manner features as the source. Therefore, the Preservation Principle correctly predicts the adapted form should preserve positively specified features of the source form that would have been neutralized by native Korean phonology. The variability of adaptation in word-final stops is compatible with this version of the Preservation Principle, as is the differential adaptation of English /s/ as Korean tense or lax, since it does not destroy positively specified English features, it merely employs/introduces a phonological feature that is *not* contrastive in English.

However, it turns out that even the Preservation Principle is not fully exceptionless in Korean loanword phonology: the sequences /nr/$_E$, /ln/$_E$, and /nl/$_E$ do not induce vowel epenthesis in loanwords. In Korean phonology, /ln/ and /nl/ are typically realized as [ll]: e.g., /k^{h}al-nal/→[k^{h}allal] 'edge of a knife', /san-lim/ [sallim] 'forest'. These sequences are adapted as either [ll] or [nn] in loanwords, e.g. *Henry* ∼> [henni]∼[helli], *on-line* [onnain]∼[ollain]. Because /n/ and /l/ contrast in English, the Preservation Principle wrongly predicts vowel epenthesis to block assimilation, e.g. *Henry* ∼> *[henɨri], *on-line* ∼> *[onɨllain]. We will refrain from speculating

further on this point, except to note that one prominent strain of thought is that the site and type of a repair is determined by the magnitude of the 'perceptual distortion' (Fleischhacker 2005). Vowel epenthesis may introduce a bigger perceptual distortion in a sonorant-sonorant cluster than in the obstruent-noncontinuant clusters studied here (Yun 2016).

Another interesting mismatch between speech perception and loanword adaptation is reported by Shinohara et al. (2011), who investigated the adaptation and perception of syllable-initial stops from various languages into Korean. Of particular relevance here is the contrast between Korean perception of Japanese stops, and Korean loanword adaptation of the same stops. They show that voiceless, word-initial stops are unfailingly adapted as lenis stops (p. 1463, Table 1), but near-categorically identified as aspirated stops in speech perception (p. 1471, Fig. 2a). Crucially, comparison with Japanese voiced stops shows an anti-Preservation effect: Japanese initial voiced stops are adapted as Korean lenis stops (so that the Japanese voicing contrast is neutralized in adaptation), while Japanese initial voiced stops are variably identified as lenis or fortis.

In summary, a more explicit refinement of the Preservation Principle is compatible with the data presented here on Korean adaptation of medial clusters that begin with voiceless obstruents—but it does not explain much that is still mysterious. The perceptual account explains *why* there is variability in the adaptation of word-final stops, but the Preservation Principle does not. Neither account explains why the same kind of phonetic variability results in consistent adaptation in word-medial stops, but variable adaptation in word-final stops.

Taken together, the present results paint a nuanced view of the relationship between speech perception and loanword adaptation. Loanword adaptation is in part driven by speech perception, since adapters must perceive the forms they are adapting. This explains why subphonemic details—in particular, audible release burst/frication noise—are sufficient to force perceptual epenthesis (e.g. the illusory vowel effect in *di̲s̲kette*), and it naturally follows that the 'epenthetic' vowel gets incorporated into the lexical representation of the adapted loanword. However, the observed loanword adaptation pattern is more categorical than would be expected if it directly reflected speech perception, and crucially fails to fully reflect the pattern of perceptual similarity judgments documented in Experiment II. The most interesting cases are the adaptation of [−noisy] stop-nasal clusters, and [+noisy] stop-stop clusters. In the former case, Koreans judge forms like *pakna* as more similar to the repaired form *paŋna* (native) than to the epenthetic repair *pak^{h}ɨna* that they actually apply in loanwords. In the latter case, Experiments I and II strongly suggest that the release burst should trigger the illusory vowel effect, and yet stop-stop loanwords are adapted categorically without vowel epenthesis.

We stepped through a Bayesian account of speech perception, in which phonotactic knowledge serves as a 'prior,' and the listener's goal is to recover the lexical representation that generated the acoustic observations (cf. Wilson and Davidson 2013; Durvasula and Kahng 2015). The Bayesian account recapitulates many key aspects of the analysis of Boersma and Hamann (2009), and therefore inherits many of its salutory predictions. However, the Bayesian account offered here differs in two ways.

First, since it is a probabilistic theory, it offers the potential to explain stimulus (and listener?) dependent variability—although in principle, Boersma and Hamann's account could be extended in this way. Second, it offers an explanation for why Korean speakers exhibit good discrimination of *pakna* from both *pa*[kʰɨn]*a* and *pa*[ŋn]*a*, and choose the former for loanword adaptation even though they think the latter is a better perceptual match when the *k* is unreleased. Boersma and Hamann's account is unable to handle this fact, because it explicitly identifies non-native speech perception with native speech perception. Their account has nothing to say about discrimination, and incorrectly predicts that the identification pattern of Experiment II must match the loanword adaptation pattern.

It remains a mystery for every account why medial stop-initial clusters are adapted near-categorically in Korean. This is especially puzzling for stop-stop clusters, since word-final stops are adapted variably with and without epenthesis. There is reasonably compelling evidence that the variability in final stop adaptation is conditioned by the presence/strength of the release burst. We know from Davidson (2011) that medial stops also exhibit variability in whether they are released. Why then does vowel epenthesis occur so rarely in this position? Having exhausted the facts available to us, we turn now to speculation.

As shown in Davidson et al. (2015), listeners exposed to phonetic variability in different tokens of the same non-native type exhibit more consistent productions of it. According to the authors, this kind of higher-variability exposure helps to "stabilize" the listener's representation of the non-native item. In other words, the influence of (irrelevant) phonetic detail is minimized when the listener receives multi-talker and multi-token exposure.[9] The observation that (Korean) loanword adaptation recapitulates speech perception patterns, but more categorically, is also made in the work of de Jong and colleagues (Park and de Jong 2008; de Jong and Cho 2012). Stabilization could explain how Korean adapters form a coherent lexical representation of the initial stop in stop-nasal and stop-stop clusters. Some additional supporting evidence for this hypothesis comes from Kwon (2017), who showed that Koreans with low English proficiency were much more likely to adapt final stops with vowel epenthesis (in English nonwords). In effect, we are suggesting that stabilization provides a mechanism for the Preservation Principle: repeated exposures to source tokens enables Korean listeners to notice and abstract over the (irrelevant) variability in the burst release, forming a stable (phonological) representation.

The generalization that loanword adaptation becomes more regular over time is not new (Haugen 1950; Kang 2010). Nonetheless, the present state of affairs is that we simply don't know enough about what occurs between speakers and hearers dur-

[9]Regarding the relationship between the phonetic variability and categorical adaptation, a reviewer suggests the following. The word-final context admits variable adaptation because it is subject to more variation since it can be followed by a pause or any other sound. The word-medial context is more stable and allows categorical adaptation since it is surrounded by the 'same' sounds across different tokens. While this possibility seems intuitively appealing, we do not think it is fully adequate to the facts. As noted in the text, the initial stops in English clusters like [kn] and [pt] are variably released. Thus, the phonetic variable of interest here—stop burst releasednes—exhibits variability in the word-medial environment, despite the stability of identity of neighboring segments.

ing loanword adaptation. Is it really the case, as proponents of the phonological account claim, that all loanword adaptation is done by high-fluency bilinguals? Some degree of bilingualism must be present for loanword adaptation to occur in the first place, but we could imagine that some technical terms with a restricted distribution are adapted by speakers with low proficiency, namely those which are disproportionately likely to occur only in writing. How much exposure does the first adapter (or adapters) have to the source word before they adapt it? For that matter, when the first adapter borrows a word, who are they speaking to? Are they more likely to use English-derived forms with fellow bilinguals (even though the opportunity for code-switching is available), and if so, what effects does this have on the adaptation process? Finally, what are the social processes involved in the uptake and spread of a loanword, such that it eventually becomes known to the wider populace (including speakers with limited or no proficiency in English)? It seems likely the repeated use and re-use of a loanword within a speech community might stabilize that particular form's lexical representation (so that even if Korean speakers perceive any one source token variably, they tend to converge on one or another adaptation at a community level). Future research might concentrate on determining how these factors contribute to the regularities we observe in loanword adaptation but not in raw speech perception.

Acknowledgements We wish to acknowledge Sharon Peperkamp, Michael Kenstowicz, and an anonymous reviewer for constructive suggestions. This work has benefitted from the comments of various audiences, including the UCLA Phonology Seminar, Northwestern University Linguistics Department, Chonnam National University English Department, University of Arizona Linguistics Department, and the Keio Reading Circle (including Shigeto Kawahara, Junko Ito, and Armin Mester). We also wish to acknowledge Quinton Maynard and IMC for the use of office space during winter and summer breaks, and Syejeong Kim for facilitating the experiments in Korea. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A2A01024340).

Appendix A: Measurements of the burst release and yoked non-release in 'source' stimuli

[+noisy]	Energy (Pa^2ś)	Power (Pa^2)	Intensity (dB)	Illusory vowel rate
pt	1.32e−5	2.35e−4	57.69	0.18
pk	1.55e−4	2.17e−3	67.34	0.31
tp	9.73e−6	1.31e−4	55.16	0.62
tk	4.47e−5	6.72e−4	62.25	0.41
kp	4.26e−5	4.27e−4	60.29	0.38
kt	4.08e−5	6.02e−4	61.77	0.66
pn	1.00e−4	9.79e−4	63.89	0.62
tn	2.72e−5	3.24e−4	59.08	0.73
kn	3.10e−5	3.54e−4	59.47	0.89

[−noisy]	Energy (Pa^2ś)	Power (Pa^2)	Intensity (dB)
pt	3.17e−7	4.75e−6	40.75
pk	1.25e−7	1.91e−6	36.78
tp	1.38e−7	1.87e−6	36.69
tk	1.03e−7	1.63e−6	36.10
kp	3.06e−7	2.85e−6	38.53
kt	9.62e−8	1.42e−6	35.51
pn	4.01e−7	4.22e−6	40.23
tn	1.29e−7	1.53e−6	35.82
kn	1.05e−5	1.34e−5	55.24

Appendix B: Regression output

(B1) *Stop-stop*

```
Formula: isRespDiff ~ language * repair * noisy + (language | stim) +
     (repair * noisy | listener)

      AIC      BIC   logLik deviance df.resid
   1055.6   1166.0   -506.8   1013.6     1393

Fixed effects:

                                       Estimate Std. Error z value Pr(>|z|)

(Intercept)                               8.142      2.016   4.039 5.38e-05 ***
languageKorean                           -2.528      1.776  -1.423 0.15473
repairassim                              -8.583      2.067  -4.153 3.29e-05 ***
noisynoisy                               -4.091      2.054  -1.992 0.04638 *
languageKorean:repairassim                4.303      1.891   2.275 0.02290 *
languageKorean:noisynoisy                -1.250      1.799  -0.695 0.48736
repairassim:noisynoisy                    7.033      2.142   3.284 0.00102 **
languageKorean:repairassim:noisynoisy    -0.818      2.000  -0.409 0.68249
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

(B2) *Stop-nasal*

```
Formula: isRespDiff ~ language * repair * noisy + (language | stim) +
      (repair * noisy | listener)

       AIC      BIC   logLik deviance   df.resid
     345.3    441.5   -151.7    303.3        699

Fixed effects:

                                       Estimate Std. Error z value Pr(>|z|)

(Intercept)                             4.8586     1.4330   3.391 0.000697 ***
languageKorean                         -1.1461     1.3849  -0.828 0.407904
repairassim                             2.8902     3.4896   0.828 0.407551
noisynoisy                             -1.2919     1.9037  -0.679 0.497361
languageKorean:repairassim             -0.9175     2.5475  -0.360 0.718749
languageKorean:noisynoisy              -5.0256     2.8329  -1.774 0.076061.
repairassim:noisynoisy                 -0.9478     4.0985  -0.231 0.817116
languageKorean:repairassim:noisynoisy   7.6806     3.7973   2.023 0.043110 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

(B3) *Fricative-stop*

```
Formula: isRespDiff ~ language * repair + (language | stim) + (repair |
listener)

    AIC       BIC   logLik deviance df.resid
  385.7     429.7   -182.9    365.7      590

Fixed effects:

                            Estimate Std. Error z value Pr(>|z|)

(Intercept)                   3.1138     0.8161   3.815 0.000136 ***
languageKorean               -3.6105     0.9142  -3.949 7.83e-05 ***
repairassim                   1.2556     1.3656   0.919 0.357845
languageKorean:repairassim    4.7512     1.4342   3.313 0.000924 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

References

Abramson, Arthur S., and Kalaya Tingsabadh. 1999. Thai final stops: Cross-language perception. *Phonetica* 56: 111–122.

Ahn, Sang-Cheol. 1998. *An introduction to Korean phonology*. Seoul: Hanshin Publishing Company.

Best, Catherine T. 1991. The emergence of native-language phonological influences in infants: A perceptual assimilation model. *Haskins Laboratories Status Report on Speech Research* SR-107/108: 1–30.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, 171–204.

Best, Catherine T., and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Second language speech learning: The role of language experience in speech perception and production*, eds. Murray J. Munro and Ocke-Schwen Bohn, 13–34. Amsterdam: Benjamins.

Boersma, Paul, and Silke Hamann. 2009. Loanword adaptation as first-language phonological perception. In *Loanword phonology*, eds. Andrea Calabrese and W. Leo Wetzels, 11–58. Amsterdam: Benjamins.

Bohn, Ocke-Schwen, and Catherine T. Best. 2012. Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics* 40: 109–128.

Bradlow, Ann, David Pisoni, Reiko Akahane-Yamada, and Yoichi Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/, IV: Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America* 101(4): 2299–2310. https://doi.org/10.1121/1.418276.

Byrd, Dani. 1992. Perception of assimilation in consonant clusters: A gestural model. *Phonetica* 49: 1–24.

Chang, Charles B., and Alan Mishler. 2012. Evidence for language transfer leading to a perceptual advantage for non-native listeners. *The Journal of the Acoustical Society of America* 132(4): 2700–2710. https://doi.org/10.1121/1.4747615.

Chen, Larissa. 2003. Evidence for the role of gestural overlap in consonant place assimilation. In *International Congress of Phonetic Sciences (ICPhS) 15*, eds. Maria-Josep Solé, Daniel Recasens, and Joaquin Romero. Barcelona: Universitat Autónoma de Barcelona.

Cho, Taehong, and James M. McQueen. 2006. Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants. *The Journal of the Acoustical Society of America* 119: 3085–3096.

Daland, Robert, and Mira Oh. 2011. Stops and phrasing in Korean and English monolinguals and bilinguals. In *International Conference of the Phonetic Sciences (ICPhS) 17*. IPA: Hong Kong.

Davidson, Lisa. 2011. Characteristics of stop releases in American English spontaneous speech. *Speech Communication* 53(8): 1042–1058.

Davidson, Lisa, Sean Martin, and Colin Wilson. 2015. Stabilizing the production of nonnative consonant clusters with acoustic variability. *The Journal of the Acoustical Society of America* 137: 856–872. https://doi.org/10.1121/1.4906264.

de Jong, Kenneth, and Mi-Hui Cho. 2012. Loanword phonology and perceptual mapping: Comparing two corpora of Korean contact with English. *Language* 88(2): 341–368.

de Jong, Kenneth, and Hanyong Park. 2012. Vowel epenthesis and segment identity in Korean learners of English. *Studies in Second Language Acquisition* 34: 127–155.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jaques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology. Human Perception and Performance* 25(6): 1568–1578.

Dupoux, Emmanuel, Erika Parlato, Sonia Frota, Yuki Hirose, and Sharon Peperkamp. 2011. Where do illusory vowels come from? *Journal of Memory and Language* 64(3): 199–210.

Durvasula, K., and Jimin Kahng. 2015. Illusory vowels in perceptual epenthesis: The role of phonological alternations. *Phonology* 32(3): 385–416.

Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In *Speech perception and linguistic experience: Theoretical and methodological issues*, ed. Winifred Strange, 233–277. Baltimore: York Press.

Fleischhacker, Heidi. 2005. Similarity in phonology: Evidence from reduplication and loan adaptation. PhD diss., UCLA.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds /l/ and /r/. *Neuropsychologia* 9(3): 317–323. https://doi.org/10.1016/0028-3932(71)90027-3.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder, and Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology. Human Perception and Performance* 24(2): 592–608.

Hallé, Pierre A., Catherine T. Best, and Andrea Levitt. 1999. Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics* 27: 281–306.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26: 210–231.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440.

Hwang, J. 2011. Non-native perception and production of foreign sequences. PhD diss., Stony Brook University. Available at https://linguistics.stonybrook.edu/sites/default/files/uploads/2011_dissertation_Jiwon_Hwang.pdf. Accessed 26 July 2018.

Jun, Jongho. 1996. Place assimilation is not the result of gestural overlap: Evidence from Korean and English. *Phonology* 13: 377–407.

Kabak, Bariş, and William J. Idsardi. 2007. Speech perception is not isomorphic to phonology: The case of perceptual epenthesis. *Language and Speech* 50: 23–52.

Kang, Hyunsook. 1996. English loanwords in Korean. *Studies in Phonetics, Phonology and Morphology* 12: 21–47.

Kang, Yoonjung. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20: 219–273. https://doi.org/10.1017/S0952675703004524.

Kang, Yoonjung. 2010. The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology* 27: 225–253.

KBS Hangugeo Yeonkwuhwe [KBS Korean Language Research Team]. 1987. Reproduction of Weraeeo Phyogi Yongryejip [Examples of loanword transcription]. Seoul: Kwukeo yeonkwuso [Institute of the Korean Language].

Kim, Soohee, and Emily Curtis. 2002. Phonetic duration of English /s/ and its borrowing into Korean. *Japanese/Korean Linguistics* 10: 406–419.

Kuhl, Patricia K. 2004. Early language acquisition: Cracking the speech code. *Nature Reviews. Neuroscience* 5: 831–843.

Kwon, Harim. 2017. Language experience, speech perception and loanword adaptation: Variable adaptation of English word-final plosives into Korean. *Journal of Phonetics* 60: 1–19.

Kwulipkwukeyenkwuwen [The National Academy for the Korean Language]. 1991. *Oylaye sayong siltay cosa: 1990 nyendo [Survey of the state of loanword usage: 1990]*. Seoul: NAKL.

Maye, Jessica, and LouAnn Gerken. 2000. Learning phonemes without minimal pairs. In *Boston University Conference on Language Development (BUCLD)* 24, eds. Catharine Howell et al.. 522–533. Somerville: Cascadilla Press.

McClelland, James L., and Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18(1): 1–86. https://doi.org/10.1016/0010-0285(86)90015-0.

Miyawaki, Kuniko, James Jenkins, Winifred Strange, Alvin M. Liberman, Robert Verbrugge, and Osama Fujimura. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18(5): 331–340. https://doi.org/10.3758/BF03211209.

Mo, Yoonsook. 2007. Temporal, spectral evidence of devoiced vowels in Korea. In *International Conference of the Phonetic Sciences (ICPhS)* 2007, ed. Ingmar Steiner, 445–448. Saarbrücken: Universität des Saarlandes.

Oh, Mira. 1996. Linguistic input to loanword phonology. *Studies in Phonetics, Phonology and Morphology* 2: 117–126.

Oh, Mira. 2012. Adaptation of English complex words into Korean. *Journal of East Asian Linguistics* 21: 267–304.

Paradis, Carole, and Darlene LaCharité. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33(2): 379–430.

Paradis, Carole, and Darlene LaCharité. 2002. Addressing and disconfirming some predictions of phonetic approximation for loanword adaptation. *Langues et Linguistique* 28: 71–91.

Park, Hanyong, and Kenneth de Jong. 2008. Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics* 6: 704–723.

Peperkamp, Sharon, Inge Vendelin, and Kimihiro Nakamura. 2008. On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology* 25: 129–164.

Pitt, Mark A., and James M. McQueen. 1998. Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39(3): 347–370.

R Core Development Team. 2014. *R [3.02]: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at http://www.R-project.org/. Accessed 26 July 2018. [Software package for statistics and graphics].

Shin, Ji-young. 2011. *Hankukeoiy Malsori [Korean sounds]*. Korea: Gisikkwa Kyoyang.

Shinohara, Shigeko, Seong-rim Ji, Tomohiko Ooigawa, and Takahito Shinya. 2011. The limited role of perception in Korean loanword adaptation: The Korean three-way laryngeal categorization of Japanese, French, English, and Chinese plosives. *Lingua* 121: 1461–1484.

Silverman, Daniel. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9: 289–328.

Sundara, Megha, and Linda Polka. 2008. Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition* 106(1): 234–258.

Wilson, Colin, and Lisa Davidson. 2013. Bayesian analysis of non-native cluster production. In *North East Linguistic Society (NELS) 40*, Cambridge: MIT.

Wilson, Colin, Lisa Davidson, and Sean Martin. 2014. Effects of acoustic-phonetic detail on cross-language speech production. *Journal of Memory and Language* 77: 1–24.

Wright, Richard A. 2004. A review of perceptual cues and cue robustness. In *Phonetically based phonology*, eds. Bruce Hayes, Robert Kirchner, and Donca Steriade, 34–57. Cambridge: Cambridge University Press.

Yip, Moira. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116: 950–975.

Yun, Suyeon. 2016. A theory of consonant cluster perception and vowel epenthesis. PhD diss., MIT.