Evidence for language transfer leading to a perceptual advantage for non-native listeners

Charles B. Chang[a] and Alan Mishler

University of Maryland, College Park Center for Advanced Study of Language, 7005 52nd Avenue, College Park, Maryland 20742

Phonological transfer from the native language is a common problem for non-native speakers that has repeatedly been shown to result in perceptual deficits vis-à-vis native speakers. It was hypothesized, however, that transfer could help, rather than hurt, if it resulted in a beneficial bias. Due to differences in pronunciation norms between Korean and English, Koreans in the U.S. were predicted to be better than Americans at perceiving unreleased stops—not only in their native language (Korean) but also in their non-native language (English). In three experiments, Koreans were found to be significantly more accurate than Americans at identifying unreleased stops in Korean, at identifying unreleased stops in English, and at discriminating between the presence and absence of an unreleased stop in English. Taken together, these results suggest that cross-linguistic transfer is capable of boosting speech perception by non-natives beyond native levels.
© 2012 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4747615]

I. INTRODUCTION

A. Transfer in non-native speech perception

Anyone who has learned a second language as an adult will know how difficult it can be to understand a foreign tongue. Compared to native speakers, non-native speakers who learn a given language after childhood tend to be substantially worse in terms of listening skills across a range of conditions (Nábělek and Donahue, 1984; Bradlow and Pisoni, 1999; Garcia Lecumberri and Cooke, 2006; Cutler *et al.*, 2008). Comprehension of a non-native language can be challenging for a variety of reasons, including cognitive constraints and incomplete knowledge of the vocabulary, grammar, and culture (Goh, 2000; Bloomfield *et al.*, 2011). At the level of speech sounds specifically, perhaps no factor is as widely recognized as cause for non-native difficulty as phonological interference, or *transfer,* from the native language. The difficulties that non-native speakers encounter with accurately perceiving phonemes (i.e., sounds that can serve to distinguish words) in the non-native language are often explicable in terms of influence from sound patterns in their native language (Lado, 1957; Best, 1994, 1995). For instance, native speakers of Japanese, in which "l" and "r" sounds are variants of one Japanese phoneme, tend to have trouble perceiving the difference between the English /l/ and /r/ phonemes, and this perceptual deficit persists even after they have learned to pronounce the distinction reliably (Goto, 1971; Sheldon and Strange, 1982; Yamada and Tohkura, 1992; Yamada, 1995). Conversely, native speakers of English, which does not utilize consonant or vowel length contrastively, tend to have trouble perceiving the difference between Japanese short and long consonants and vowels (Han, 1992; Tajima *et al.*, 2008).

What these cases of phonological transfer have in common is a bias from perception of the native language that is detrimental to perception of the non-native language. Native perceptual habits that predispose individuals toward processing certain sounds as belonging to the same phoneme category in their mother tongue become a hindrance when transferred to a new language that requires the sounds to be processed as different phonemes (Flege, 1995; Cutler, 2001; Best and Tyler, 2007). More generally, transfer of native-language phonological patterns by non-native speakers has been shown to result in non-native performance that is either significantly worse than native performance ("negative" transfer) or, at best, not significantly different from native performance ("positive"—or, perhaps more aptly, neutral—transfer) (Odlin, 1989). Neutral transfer generally involves a parallelism between the native and non-native sound inventories, such that perception of the non-native sounds does not suffer significantly in comparison to native perception. Thus in the case of English /l/ and /r/, while Japanese learners of English tend to experience negative transfer from Japanese (which contains only one similar phoneme), French and German learners of English tend to experience more neutral transfer from their respective native languages (which, like English, contain separate /l/ and /r/ phonemes) (Hallé *et al.* 1999; Iverson *et al.*, 2003). This pattern of negative or null effects of cross-linguistic transfer is not surprising because it is not readily apparent how multilingual experience—that is, experience in more than one language—could be better for perceiving a given language than monolingual native experience in that language.

While the preponderance of evidence from second language acquisition studies suggests that cross-linguistic transfer has at best a null effect on non-native speech perception vis-à-vis native speech perception, a few studies have

produced findings suggesting that experience from the native language can be helpful for perception of a non-native language. Research examining the perception of non-native speech has found an "interlanguage speech intelligibility benefit" wherein non-native listeners of the same native language background as a non-native talker can be better at comprehending that talker than native listeners are (Bent and Bradlow, 2003; Bent et al., 2008). Meanwhile, work on the perception of approximants in a non-native language has shown that French, German, and Danish listeners are more sensitive to small steps on a speech continuum between English /w/ and /j/ than English listeners are (Hallé et al. 1999; Bohn and Best, 2012). In addition, work on the perception of final stops in a non-native language has shown that different groups of non-native listeners show distinct patterns of performance following from the utility of their native language experience: Korean listeners, who have extensive experience processing unreleased stops in their native language, outperform Dutch listeners in detection of unreleased stops in English (Cho and McQueen, 2006).

Although these findings suggest that cross-linguistic transfer can be helpful for perception of a non-native language, they leave open the question of whether transfer can actually result in better-than-native perceptual performance. After all, the interlanguage speech intelligibility benefit applies to a specific kind of speech—namely, foreign-accented speech produced by a non-native talker whose language background is shared by the non-native listener—not to the more general case of unaccented speech. Moreover, the differences found between native and non-native perception of approximants ostensibly favoring non-native listeners occurred in the discrimination of resynthesized speech continua. Therefore these data show only that French listeners, for example, are more sensitive than English listeners to minimal variations in the relevant acoustic parameters in the relevant ranges and cannot be interpreted as evidence that French listeners perceive English approximants more accurately than English listeners.[1] Finally, the disparities between different groups of non-native listeners in the perception of final stops do not provide evidence of better-than-native perceptual performance for either group because the native benchmark was never established.

B. The present study

Given the limited conclusions about cross-linguistic transfer that can be drawn from previous findings, the current study endeavored to address the question of whether transfer can in fact result in non-native listeners perceiving a given language more accurately than native listeners. It was hypothesized that transfer could in principle boost non-native speech perception beyond the standard of native speech perception due to the multiple possible intersections of two properties of language: *Frequency* and *canonicity*. In regard to speech sounds, frequency refers to the rate of occurrence of a phoneme's variants, which usually differ in terms of how often they occur in the language. Canonicity refers to the relative status of a phoneme's variants, of which one is usually regarded as the ideal or canonical form. Frequency and

canonicity often coincide, such that the most frequent variant of a phoneme is also its canonical form; however, this is not always the case. In American English, for example, the stop consonant /t/ at the end of a word may be variably pronounced as released (with an audible burst, $[t^{(h)}]$) or unreleased (with no audible burst, $[t^\urcorner]$). The unreleased variant is more frequent than the released variant; nevertheless, the released variant is the form that American English speakers seem to recognize as the basic one and store as their mental representation of /t/ (Sumner and Samuel, 2005).

The difference between released and unreleased stop variants is illustrated in Fig. 1 for the words *pup* and *putt*.[2] Here it can be seen, first, that the release burst associated with a final released stop differs markedly in energy profile between different places of articulation: Bilabial [p] has a relatively soft burst (evident in the light shading) with diffuse energy, while alveolar [t] has a relatively loud burst (evident in the dark shading) with concentrations of energy in the higher frequencies. Such burst properties thus provide a rich source of information about the identity of a final stop. In the case of an unreleased stop, however, these cues are not available to the listener because there is no burst. Instead the listener must rely on regular patterns of perturbation in the preceding vowel, such as in the second formant (F_2) resonance of the vocal tract, to recover the identity of the final stop. In Fig. 1, for example, it can be seen that a final /t/, whether released or unreleased, induces a rise in the F_2 of the given vowel, whereas final /p/ does not. These sorts of vocalic transition cues may be used in conjunction with burst cues in the perception of a final released stop but must serve as the primary source of information in the perception of a final unreleased stop.

As mentioned in the preceding text, final stop variants in American English demonstrate a disparity between frequency and canonicity; this provides an interesting test case for an investigation of perceptual biases from the native language because it allows for a comparison between native and non-native speakers that can actually favor the perceptual biases of the non-natives. Given native and non-native speakers of a language that are both experienced with hearing a particular form, the question posed in the current study was whether a difference in the form's canonicity for native and non-native speakers could boost non-native speech perception beyond the standard of native speech perception. To answer this question, this study examined whether the unreleased form of word-final voiceless stops in American English would be better perceived by native speakers of American English (where the unreleased form is frequent, but non-canonical) or by non-native speakers of English whose native language is Korean (where the unreleased form is obligatory and, therefore, both frequent and canonical) (Byrd, 1993; Davidson, 2011; Sohn, 1999).

As the object of perception was American English unreleased stops, two factors stood to influence listeners' performance: Experience with American English and experience with unreleased stops. Experience with unreleased stops favored the Koreans because their native language experience would bias them to attend to cues to a stop's place of articulation that are contained in the preceding vowel rather
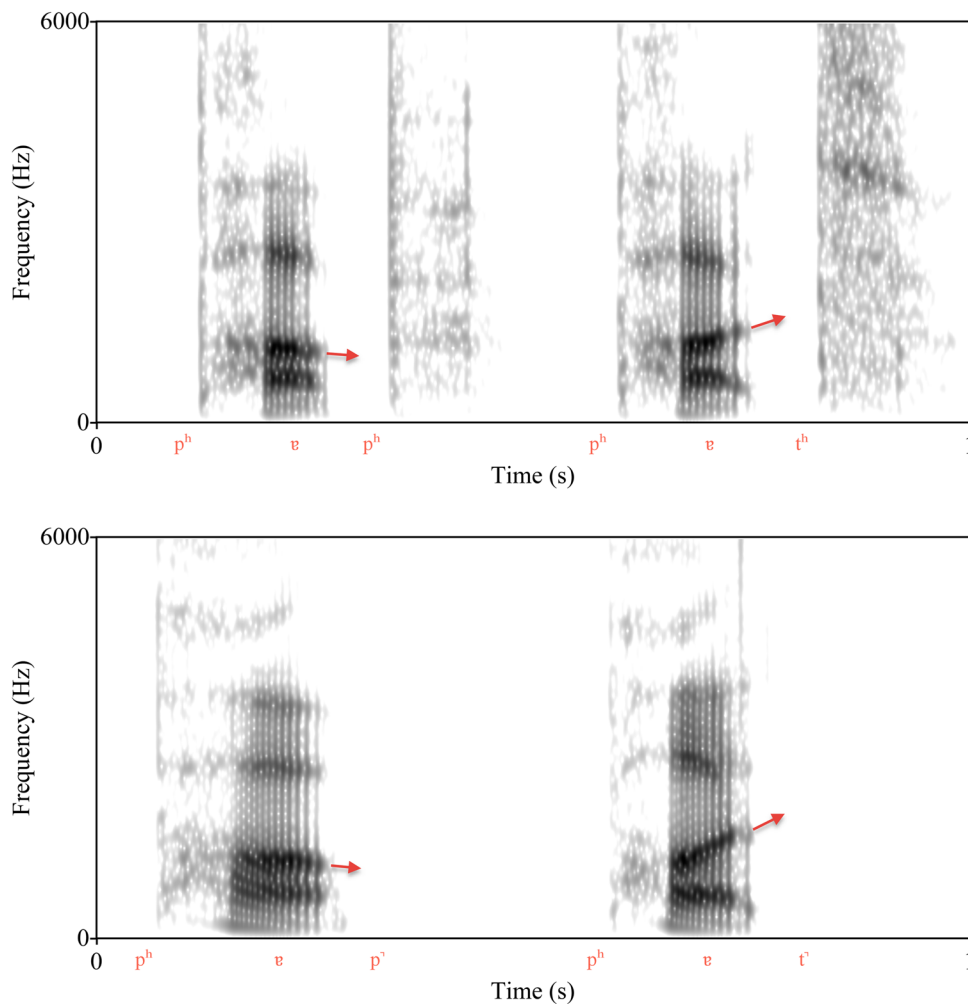
FIG. 1. (Color online) Spectrograms of a native English speaker uttering the words *pup* and *putt* with released final stops (top panel) and unreleased final stops (bottom panel). The different trajectories of the second formant preceding word-final /p/ and word-final /t/ are marked with arrows.

than in the stop's release burst (cf. Abramson and Tingsabadh, 1999; Cho and McQueen, 2006). On the other hand, experience with American English favored the Americans because this experience would better equip them to process English speech, including any properties of unreleased stops that might be particular to American English (cf. Tsukada, 2006; Tsukada *et al.*, 2007; Tsukada and Roengpitya, 2008). It was predicted that the Koreans' extensive experience processing unreleased stops in their native language would prove to be a decisive advantage, such that they would be significantly better than the Americans at perceiving unreleased stops across languages—that is, not only in their native language (Korean) but also in their non-native language (English). To test this prediction, three speech perception experiments were conducted with a group of native English speakers and a group of native Korean learners of English recruited from the University of Maryland community.

Experiments 1 and 2 tested whether native Korean learners of English would be better than native English speakers at identifying unreleased stops in Korean and English, respectively. These two experiments were focused specifically on how well the two groups would make use of the acoustic cues available to identify unreleased stops, so the influence of lexical frequencies was minimized by using nonce word materials. Using nonce words also made the task in Experiment 2 especially difficult for the Korean group as word

unfamiliarity has been shown to significantly worsen the perceptual performance of non-native listeners (Yamada *et al.*, 1996; Mora, 2005). Listeners in these experiments heard either Korean nonce words in isolation (Experiment 1) or English nonce words at the end of a carrier sentence (Experiment 2) in a speeded identification task. Each nonce word ended in either an unreleased voiceless stop (e.g., *ruzzepe* [ˈɹɛzip̚]) or a sonorant sound such as a vowel (e.g., *ruzzy* [ˈɹɛzi]), and listeners had to identify the final sound as /p/, /t/, /k/, or none of these as quickly and accurately as possible.

Experiment 3 tested whether native Korean learners of English would show a perceptual advantage over native English speakers even when native English speakers were allowed the additional benefit of lexical knowledge. Listeners in this experiment heard pairs of real English words uttered by different talkers in a speeded discrimination task. Each pair of words differed with respect to the identity of a final unreleased stop (e.g., *weep* vs *wheat*), with respect to the presence of a final unreleased stop (e.g., *beet* vs *bee*), or not at all, and listeners had to judge whether the two talkers had said the same word or different words. Experiment 3 thus constituted the strongest test of the prediction of a perceptual advantage for native Korean learners of English because in this experiment native English speakers had the benefit of both native English speech processing and native English lexical knowledge.

C. B. Chang and A. Mishler: The non-native advantage in speech perception

II. METHODS

A. Participants

Participants were recruited from the University of Maryland, College Park. None reported any history of hearing, speech, or language impairments, and all gave informed consent and were paid for their participation.

The talker who recorded the stimuli for Experiment 1 (Korean nonce words) was a male native speaker of Korean (age, 32 yr) who had been born and raised in Seoul. The talkers who recorded the stimuli for Experiment 2 (English nonce words) and Experiment 3 (English real words) were two male native speakers of American English (ages, 19 and 25 yr) who had grown up in Maryland since early childhood and had no experience with any language containing obligatorily unreleased stops.

Listeners who participated in Experiments 1-3 comprised 25 native speakers of American English (11 male, 14 female; mean age, 21.1 yr, SD 5.4) and 25 native speakers of Korean (11 male, 14 female; mean age, 27.6 yr, SD 7.5). American listeners were born and raised in the U.S. (mostly in Maryland, Virginia, or Washington, DC) and had no experience with a language containing obligatorily unreleased stops. Korean listeners were born and raised primarily in South Korea with a mean age of arrival to the U.S. of 21.1 yr (SD 9.0). Consistent with the compulsory nature of English education in modern South Korea, they generally reported extensive study of English (mean length, 12.3 yr, SD 7.0) but no experience with other languages containing variably or obligatorily unreleased stops.

B. Stimuli

The stimuli for Experiment 1 comprised 28 bisyllabic Korean nonce words that varied in terms of final consonant and final vowel. All items were of the shape $C_1V_1C_2V_2(C_3)$, where C = consonant and V = vowel. To provide a strong test of the hypothesis that native Korean learners of English would outperform native English speakers at identification of Korean unreleased stops, the Korean nonce words were made to be similar to English by filling the first two consonant slots with Korean consonants that also occur in English: The voiced bilabial nasal [m] (C_1) and the voiced alveolar flap [ɾ] (C_2). The first vowel slot (V_1) was filled with a high back rounded vowel with a palatal on-glide ([ju]), while the second vowel slot (V_2) ranged over the seven vowels in the vowel inventory of modern Korean: /i, u, a, ɛ, o, ʌ, ɯ/ (Lee, 1993; Ingram and Park, 1997; Ko, 2009). The final consonant slot (C_3) varied between /p/, /t/, /k/, and zero (i.e., absence of a final stop). This resulted in a set of 28 different nonce words (7 possible final vowels × 4 possible final consonants) shown in Table I.

The stimuli for Experiment 2 comprised 56 bisyllabic English nonce words that varied in terms of final consonant, final vowel, and stress pattern. As in Experiment 1, all items were of the shape $C_1V_1C_2V_2(C_3)$. To provide a strong test of the hypothesis that native Korean learners of English would outperform native English speakers at identification of English unreleased stops, the English nonce words were made to

be identifiably English-like and dissimilar from Korean by filling the first two consonant slots with English consonants absent from the Korean inventory: The voiced alveolar approximant [ɹ] (C_1) and the voiced alveolar fricative [z] (C_2). The first vowel slot (V_1) was filled with a mid central vowel (stressed [ɐ] or unstressed [ə]), while the second vowel slot (V_2) ranged over a set of seven English syllable nuclei: /i, u, ɑ, eɪ, oʊ, ɑɪ, ɑɹ/. The first three were chosen because they each have a parallel in a similar Korean vowel phoneme, while the latter four were chosen because they each lack a parallel in a Korean vowel phoneme due to their dynamic quality. To avoid confusion regarding which sound to identify in the identification task, only the final consonant slot (C_3) in any item was allowed to be a stop, and C_3 again varied among /p/, /t/, /k/, and zero. Finally, primary stress was varied between the initial and final syllables to create 56 different nonce words (7 possible final syllable nuclei × 4 possible final consonants × 2 possible stress patterns) shown in Table II.

The stimuli for Experiment 3 comprised 48 minimal pairs of monosyllabic English words differing in the presence and place of articulation of a final voiceless stop. The word pairs are shown in Table III. There were 12 stop/zero pairs differing with respect to the presence of a final stop (e.g., *peek, pee*) and 36 stop/stop pairs differing with respect to the place of a final stop (e.g., *lip, lick*). The word pairs represented most of the English vowels and were balanced for spoken frequency to guard against systematic disparities among /p/, /t/, /k/, and sonorants in terms of their default

TABLE I. Korean nonce word stimuli used in Experiment 1.

Final nucleus	Nonce words (in IPA)
/i/	mjuɾip̚, mjuɾit̚, mjuɾik̚, mjuɾi
/u/	mjuɾup̚, mjuɾut̚, mjuɾuk̚, mjuɾu
/a/	mjuɾap̚, mjuɾat̚, mjuɾak̚, mjuɾa
/ɛ/	mjuɾɛp̚, mjuɾɛt̚, mjuɾɛk̚, mjuɾɛ
/o/	mjuɾop̚, mjuɾot̚, mjuɾok̚, mjuɾo
/ʌ/	mjuɾʌp̚, mjuɾʌt̚, mjuɾʌk̚, mjuɾʌ
/ɯ/	mjuɾɯp̚, mjuɾɯt̚, mjuɾɯk̚, mjuɾɯ

TABLE II. English nonce word stimuli used in Experiment 2.

Stress	Final nucleus	Nonce words (in IPA)
Initial	/i/	ˈɹɛzip̚, ˈɹɛzit̚, ˈɹɛzik̚, ˈɹɛzi
	/u/	ˈɹɛzup̚, ˈɹɛzut̚, ˈɹɛzuk̚, ˈɹɛzu
	/ɑ/	ˈɹɛzɑp̚, ˈɹɛzɑt̚, ˈɹɛzɑk̚, ˈɹɛzɑ
	/eɪ/	ˈɹɛzeɪp̚, ˈɹɛzeɪt̚, ˈɹɛzeɪk̚, ˈɹɛzeɪ
	/oʊ/	ˈɹɛzoʊp̚, ˈɹɛzoʊt̚, ˈɹɛzoʊk̚, ˈɹɛzoʊ
	/ɑɪ/	ˈɹɛzɑɪp̚, ˈɹɛzɑɪt̚, ˈɹɛzɑɪk̚, ˈɹɛzɑɪ
	/ɑɹ/	ˈɹɛzɑɹp̚, ˈɹɛzɑɹt̚, ˈɹɛzɑɹk̚, ˈɹɛzɑɹ
Final	/i/	ɹəˈzip̚, ɹəˈzit̚, ɹəˈzik̚, ɹəˈzi
	/u/	ɹəˈzup̚, ɹəˈzut̚, ɹəˈzuk̚, ɹəˈzu
	/ɑ/	ɹəˈzɑp̚, ɹəˈzɑt̚, ɹəˈzɑk̚, ɹəˈzɑ
	/eɪ/	ɹəˈzeɪp̚, ɹəˈzeɪt̚, ɹəˈzeɪk̚, ɹəˈzeɪ
	/oʊ/	ɹəˈzoʊp̚, ɹəˈzoʊt̚, ɹəˈzoʊk̚, ɹəˈzoʊ
	/ɑɪ/	ɹəˈzɑɪp̚, ɹəˈzɑɪt̚, ɹəˈzɑɪk̚, ɹəˈzɑɪ
	/ɑɹ/	ɹəˈzɑɹp̚, ɹəˈzɑɹt̚, ɹəˈzɑɹk̚, ɹəˈzɑɹ

TABLE III. English real word stimuli used in Experiment 3.

Contrast	Words
/p/ vs /t/	weep-wheat, whip-wit, rape-rate, cap-cat, hoop-hoot, taupe-tote, pop-pot, pup-putt, tripe-trite, tarp-tart, warp-wart, kelp-Celt
/t/ vs /k/	seat-seek, sit-sick, bait-bake, net-neck, rat-rack, loot-Luke, oat-oak, cot-cock, mutt-muck, bite-bike, Bart-bark, port-pork
/k/ vs /p/	chic-sheep, lick-lip, peck-pep, wreck-rep, tack-tap, slack-slap, coke-cope, soak-soap, shock-shop, pike-pipe, hike-hype, hark-harp
/p/ vs zero	keep-key, type-tie, ripe-rye, gulp-gull
/t/ vs zero	beet-bee, suit-sue, mart-mar, silt-sill
/k/ vs zero	peek-pee, make-may, lake-lay, spike-spy
No change	ape, dupe, hop, cup, quit, great, tot, curt, cheek, slick, lock, cork, new, row, four, hell

likelihood to be perceived as the final sound of a word. Spoken frequencies were obtained using the Corpus of Contemporary American English (Davies, 2008) and took into account the spoken frequencies of all words with the same phonological form (i.e., all homophones); for instance, the spoken frequency of the item *peek* /pik/ was calculated as the sum total of the spoken frequencies of *peek, peak,* and *pique*. In this manner, the frequency balance across the two members of a minimal pair was controlled, such that all pairs had phonological forms differing in spoken frequency by less than an order of magnitude. The 48 word pairs were divided roughly equally among low-frequency (<1 word per million), mid-frequency (1–10 words per million), and high-frequency (>10 words per million) items. The average log (base 10) spoken frequencies of the /p/-, /t/-, /k/-, and sonorant-final items contrasted in the experiment were, respectively, 0.61, 0.65, 0.82, and 1.03.

C. Procedure

1. Stimulus recording

The stimuli for Experiments 1–3 were recorded in a sound-attenuated booth at the University of Maryland, College Park, using a Zoom H4N mobile audio recorder and an Audix HT5 head-mounted condenser microphone positioned approximately 2 cm to the left of the talker's mouth. Audio was recorded with 44.1-kHz sampling and 24-bit resolution. Items for Experiment 1 were presented via Korean spelling, and items for Experiments 2 and 3, via English spelling (with the stressed syllable underlined for the bisyllabic nonce items), on randomized individual index cards three times. Clarification of the desired pronunciation of a nonce item was provided in the few cases where this was not clear to the talker from the spelling of the item. To help maintain a steady rate of presentation, a Qwik Time QT-3 metronome set at 60 beats/min was used to present items at a rate of approximately one every 2 s.

The blocks of tokens that ultimately provided the English stimuli were ones in which the talker had been instructed to "fully pronounce" (i.e., release) the final consonants. These tokens were edited in PRAAT (Boersma and Weenink,

2011) to remove the releases and normalize the peak intensity to 0.99. Tokens that were produced as released and then "dereleased" in this way were used instead of tokens that were naturally produced as unreleased because the presence of a release made it clear that the oral closure of the final stop consonant was realized (rather than replaced with a glottal stop).[3] Although naturally unreleased stops may encode place of articulation information differently from (de)released stops (e.g., with more distinct formant transitions in the preceding vowel), previous research comparing the perception of dereleased and unreleased stops found the two to be not significantly different from each other (Lisker, 1999), suggesting that results found with dereleased stops are likely to generalize to unreleased stops.

2. Perception experiments

All three experiments took place in the same sound-attenuated booth in which the stimuli were recorded. Listeners were provided with oral instructions by the experimenter and written instructions on screen, both in their native language. They first had the tasks explained to them and were then told to listen carefully to the stimuli and respond as quickly and accurately as possible. Stimuli were presented in E-PRIME (Psychology Software Tools, Inc., 2002) on a Dell Latitude D430 laptop computer through Audio-Technica QuietPoint ATH-ANC7 binaural headphones, and listeners entered their responses on a Psychology Software Tools Model 200A serial response box connected to the computer. The experiments were completed in the following order with intervening breaks: Experiment 3, Experiment 2, Experiment 1. Experiment 1 was completed last to avoid any potential influence of the processing of Korean stimuli on the processing of English stimuli in Experiments 2 and 3.

Experiment 1 consisted of a speeded four-alternative forced choice (4AFC) identification task with Korean nonce word stimuli. These stimuli were presented in isolation to minimize the difficulty of the task for native English speakers unfamiliar with Korean. Each trial consisted of the presentation of a trial counter on screen for 1 s and then the playing of one of the 28 nonce words, which listeners had to identify as ending in /p/, /t/, /k/, or something else ("other") as quickly and accurately as possible. Listeners heard eight practice trials and then three randomized blocks of 28 test trials.

Experiment 2 consisted of a similar 4AFC identification task with English nonce word stimuli. These stimuli were uttered by two different talkers, as well as presented at the end of an English sentence, to maximize the difficulty of the task for non-native English speakers. Each trial consisted of the presentation of a trial counter on screen for 1 s and then the playing of a randomly selected precursor (*This word is…, Now the word is…,* or *The next word is…*) and one of the 56 nonce words, which listeners again had to identify as ending in /p/, /t/, /k/, or something else ("other") as quickly and accurately as possible. Listeners heard eight practice trials and then three randomized blocks of 56 test trials. The precursor and nonce word in a given trial were always spoken by the same talker. Trials in the first block were spoken

by the first talker; trials in the second block, by the second talker; and trials in the third block, by either talker.

Experiment 3 consisted of a speeded AX discrimination task with English word stimuli. Word pairs were uttered by two different talkers to maximize the difficulty of the task for non-native English speakers and to encourage discrimination of the words at an abstract level (i.e., not at an acoustic level; see, e.g., Flege, 2003). Each trial consisted of the presentation of a trial counter on screen for 1 s, the playing of the first word (A), an inter-stimulus interval (ISI) of 1 s, and then the playing of the second word (X), which listeners had to identify as either the same word or a different word as quickly and accurately as possible. Like the use of different talkers for the two stimuli in a trial, the use of a long ISI was meant to discourage listeners from discriminating the stimuli at an acoustic level; by imposing a memory demand on processing of the first stimulus, the long ISI instead encouraged higher-level encoding of the stimuli using long-term phonological representations associated with lexical items. Listeners heard 12 practice trials and then a total of 192 test trials (96 "same" trials and 96 "different" trials) divided into two randomized blocks. The "same" and "different" trials were distributed evenly across the two blocks and across both possible talker orders.

III. RESULTS

A. Experiment 1: Korean nonce words

As expected, the Koreans were much better than the Americans at identifying final sounds in Korean (Fig. 2). To analyze the identification data, a logistic mixed-effects model of identification accuracy[4] was built, in stepwise fashion, starting with random-effect terms for Participant and Item and then adding fixed-effect terms for Final (unreleased

stop or sonorant), Group (Americans or Koreans), and a Final × Group interaction. Adding Final significantly improved the model's predictions [$\chi^2(1) = 14.198$, $P = 0.0002$], as did adding Group [$\chi^2(1) = 31.13$, $P < 0.0001$]; however, adding the Final × Group interaction did not [$\chi^2(1) = 0.825$, $P = 0.364$]. Thus the final model [$N = 4200$, log-likelihood $= -1539$] included fixed effects for Final and Group with no interaction term. The results of this model showed that the odds of American listeners correctly identifying a final unreleased stop were better than 50-50 [$\beta = 0.819$, $z = 2.907$, $P = 0.004$] and that American listeners were better at identifying final sonorants (i.e., vowels) as "other" sounds than at identifying final stops as such [$\beta = 2.021$, $z = 4.216$, $P < 0.0001$]. Final sounds, however, were overall more likely to be identified accurately by Korean listeners [$\beta = 1.624$, $z = 6.527$, $P < 0.0001$], and mixed-effects models built to examine the effect of Group on identification accuracy for each final type separately confirmed that the perceptual advantage for Korean listeners held true of both final stops [$\beta = 1.660$, $z = 6.387$, $P < 0.0001$] and final sonorants [$\beta = 1.657$, $z = 2.783$, $P = 0.005$]. Moreover, an examination of accuracy on stop-final items by place of articulation of the final stop revealed that Korean listeners showed higher accuracy for all three places: /p/ (82% to the Americans' 57%), /t/ (91% to the Americans' 80%), and /k/ (90% to the Americans' 32%). American listeners, meanwhile, showed a bias toward identifying final stops as /t/, which was by far the more frequent incorrect response (of the two incorrect stop options) for both final /p/ and final /k/.

To examine whether the Koreans' higher accuracy in identification of Korean finals could be attributed to systematic differences in speed of identification (i.e., responding accurately more often due to taking longer to enter responses), response times for correct identification judgments in the two groups were log-transformed (Newell and Rosenbloom, 1981; Johnson, 2008) and analyzed in a linear mixed-effects model with random-effect terms for Participant and Item (Baayen et al., 2008) and a fixed-effect term for Group. The model showed no significant effect of Group on response times for entering correct identification judgments [$\beta = -0.024$, $t = -1.240$, $P = 0.214$], suggesting that the Koreans' superior identification performance in Experiment 1 was not an artifact of a speed-accuracy tradeoff.

In short, the results of Experiment 1 provided evidence that, in a Korean context, native Korean learners of English were significantly better than native speakers of American English at both perceiving final unreleased stops and perceiving the occurrence of other final sounds.[5]

B. Experiment 2: English nonce words

In accordance with the prediction of a cross-linguistic perceptual advantage for the Koreans, the Koreans were better than the Americans at identifying unreleased stops in English, too; however, the two groups were similarly proficient at identifying the occurrence of other final sounds in English (Fig. 3). As in Experiment 1, the identification data were analyzed by building a logistic mixed-effects model of identification accuracy with random-effect terms for
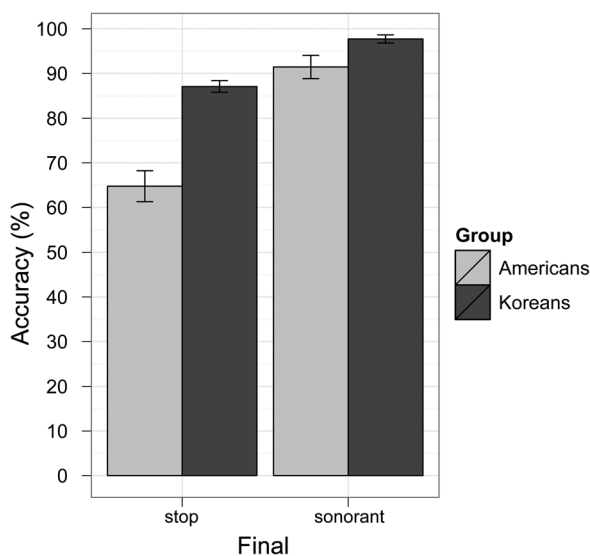

FIG. 2. Identification accuracy in Experiment 1 (Korean nonce words) by final type and group. The leftmost bars plot mean accuracy for identification of Korean unreleased stop finals; the rightmost bars, mean accuracy for identification of Korean sonorant finals as "other" sounds (i.e., not /p/, /t/, or /k/). Light gray and dark gray bars represent the American and Korean groups, respectively. Each error bar marks ±1 standard error of the mean (over 25 participants).
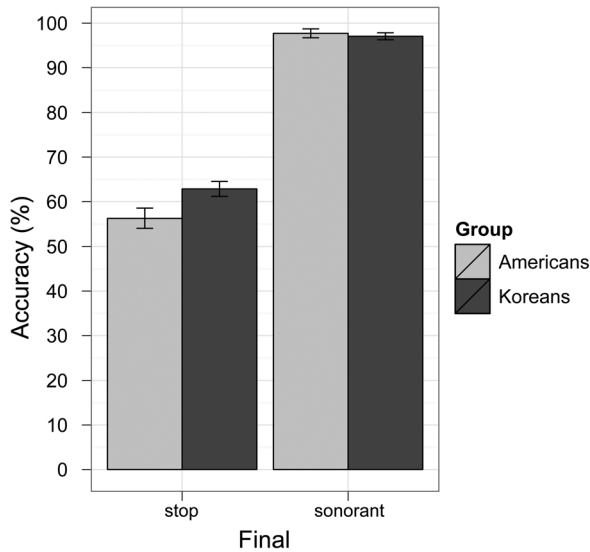
FIG. 3. Identification accuracy in Experiment 2 (English nonce words) by final type and group. The leftmost bars plot mean accuracy for identification of English unreleased stop finals; the rightmost bars, mean accuracy for identification of English sonorant finals as other sounds (i.e., not /p/, /t/, or /k/). Light gray and dark gray bars represent the American and Korean groups, respectively. Each error bar marks ±1 standard error of the mean (over 25 participants).

Participant and Item and fixed-effect terms for Final (unreleased stop or sonorant), Group (Americans or Koreans), and a Final × Group interaction. Adding Final significantly improved the model's predictions [$\chi^2(1) = 106.71$, $P < 0.0001$] and so did adding Group [$\chi^2(1) = 4.633$, $P = 0.031$] and the Final × Group interaction [$\chi^2(1) = 5.534$, $P = 0.019$]. Consequently, the final model [$N = 8400$, log-likelihood $= -3776$] included all three fixed effects. The results of this model showed that the odds of American listeners correctly identifying a final unreleased stop were not significantly better than 50-50 [$\beta = 0.296$, $z = 1.794$, $P = 0.073$] but that American listeners were far better at identifying the occurrence of other final sounds [$\beta = 4.014$, $z = 11.574$, $P < 0.0001$]. As in Experiment 1, an unreleased stop was more likely to be identified accurately by Korean listeners than American listeners [$\beta = 0.355$, $z = 2.387$, $P = 0.017$]. However, consistent with the informativeness of the interaction term, a second mixed-effects model just for identification of the sonorant-final stimuli indicated that, unlike a final unreleased stop, the occurrence of other final sounds was not significantly more likely to be identified accurately by Korean listeners than American listeners [$\beta = -0.648$, $z = -1.111$, $P = 0.267$] as both showed near-perfect identification on sonorant-final items. An examination of accuracy on stop-final items by place of articulation of the final stop revealed that Korean listeners showed higher accuracy for /p/ (76% to the Americans' 57%) and /k/ (55% to the Americans' 32%) but not for /t/ (59% to the Americans' 80%). Again, however, American listeners showed a systematic bias toward identifying unreleased stops as /t/, which was even stronger here than in Experiment 1, making it unclear whether the Americans' high accuracy on /t/ was more than the result of their general tendency to label an unreleased stop as /t/.

To check that the Koreans' superior stop identification performance in Experiment 2 was not isolated to syllables containing English vowels that are similar to Korean vowels (i.e., /i, u, ɑ/), a third mixed-effects model was built on a subset of the data comprising identification judgments on items containing a final unreleased stop and one of the diphthongal English syllable nuclei that are absent from the Korean vowel inventory (i.e., /eɪ, oʊ, ɑɪ, ɑɹ/). Consistent with the results of the first model, the results of this model showed that a final unreleased stop was still more likely to be identified accurately by Korean listeners than American listeners when the preceding vowel was dissimilar from any Korean vowel [$\beta = 0.492$, $z = 3.142$, $P = 0.002$].

To examine whether the Koreans' higher accuracy in identification of English unreleased stops could be attributed to systematic differences in speed of identification, log-transformed response times for correct stop identification judgments in the two groups were analyzed as for Experiment 1 in a linear mixed-effects model with random-effect terms for Participant and Item and a fixed-effect term for Group. The model showed no significant effect of Group on response times for entering correct stop identification judgments [$\beta = 0.020$, $t = 0.960$, $P = 0.339$], suggesting that the Koreans' superior stop identification performance in Experiment 2 was not an artifact of a speed-accuracy tradeoff.

In short, the results of Experiment 2 provided evidence that, in an English context, native Korean learners of English were significantly better than native speakers of American English at perceiving a final unreleased stop, regardless of whether the vowel preceding was similar to a Korean vowel. Furthermore, both groups were excellent at perceiving the occurrence of other final segments.

C. Experiment 3: English real words

The Koreans were no better than the Americans at discriminating minimal pairs of English words differing in the identity of a final unreleased stop (e.g., *weep, wheat*); however, they were significantly better at discriminating minimal pairs differing in terms of the presence of a final unreleased stop (e.g., *beet, bee*), as shown in Fig. 4. Participants' perceptual sensitivity to stimulus changes was calculated in terms of d', a measure of signal detection that takes into account both discrimination accuracy and response bias (Macmillan and Creelman, 2005). Participants' d' scores for discriminating the two contrast types were analyzed in a linear mixed-effects model, starting with a random-effect term for Participant and adding fixed-effect terms for Contrast (stop/stop or stop/zero), Group (Americans or Koreans), and a Contrast × Group interaction. Adding Contrast significantly improved the model's predictions [$\chi^2(1) = 63.862$, $P < 0.0001$]. Adding Group did not further improve the model's predictions [$\chi^2(1) = 1.733$, $P = 0.188$] but adding the Contrast × group interaction did [$\chi^2(1) = 8.274$, $P = 0.004$]. The results of the final model [$N = 100$, log-likelihood $= -55.873$] showed that sensitivity to a change between two stop-final words was significantly greater than zero for American listeners [$\beta = 1.191$, $t = 13.841$, $P < 0.0001$] and, moreover, not significantly different for Korean listeners
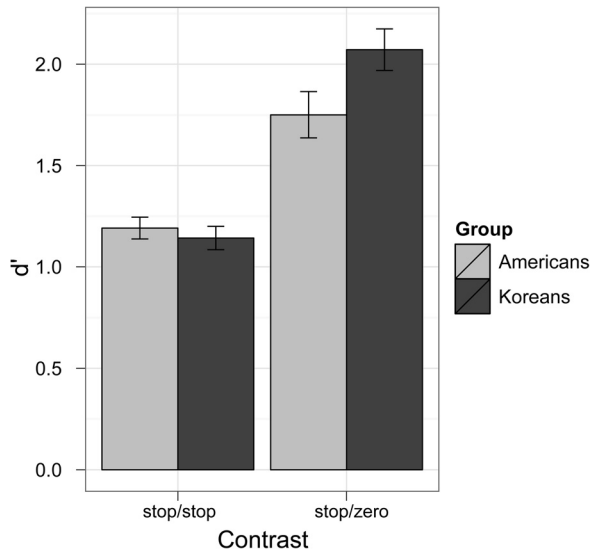
FIG. 4. Perceptual sensitivity in Experiment 3 (English real words) by contrast type and group. The leftmost bars plot d' for discrimination of English minimal pairs differing in terms of final stop (e.g., *weep, wheat*); the rightmost bars, d' for discrimination of English minimal pairs differing in terms of the presence of a final stop (e.g., *beet, bee*). Light gray and dark gray bars represent the American and Korean groups, respectively. Each error bar marks ±1 standard error of the mean (over 25 participants).

compared to American listeners [$\beta = -0.049$, $t = -0.404$, $P = 0.688$].[6] American listeners were significantly more sensitive to stop/zero contrasts than stop/stop contrasts [$\beta = 0.559$, $t = 6.284$, $P < 0.0001$]. However, consistent with the informativeness of the Contrast × Group interaction, a second mixed-effects model just for discrimination of stop/zero pairs indicated that sensitivity to a stop/zero contrast was significantly greater for Korean listeners than for American listeners [$\beta = 0.321$, $t = 2.095$, $P = 0.041$].

To examine whether the Koreans' greater displayed sensitivity to stop/zero contrasts in English minimal pairs could be attributed to systematic differences in speed of discrimination, log-transformed response times for correct stop/zero discrimination judgments in the two groups were analyzed in a linear mixed-effects model with random-effect terms for Participant, Item 1, and Item 2 and a fixed-effect term for Group. The model showed no significant effect of Group on response times for entering correct stop/zero discrimination judgments [$\beta = 0.019$, $t = 0.790$, $P = 0.432$], suggesting that the Koreans' superior discrimination of stop/zero contrasts in Experiment 3 was not an artifact of a speed-accuracy tradeoff.

Thus the results of Experiment 3 provided evidence that in an English lexical context, native Korean learners of English were significantly more sensitive than native speakers of American English to the presence of a final unreleased stop. On the other hand, the two groups were not significantly different in terms of sensitivity to a change in a final unreleased stop.

IV. DISCUSSION

To summarize, it was found that, as a group, native Korean learners of English were better than native speakers of American English at perceiving unreleased stops across languages. In Experiment 1, Koreans outperformed Americans

in identification of unreleased stops in Korean nonce words, and in Experiment 2, they did so in English nonce words as well. Finally, in Experiment 3, Koreans outperformed Americans in discrimination between the presence and absence of an unreleased stop in real English words. It remains to be seen whether the pattern found in Experiments 1 and 2 with Korean speakers who are familiar with English also obtains with Korean speakers who are not familiar with English. Nevertheless, the findings of all three experiments are consistent in suggesting that experience with Korean results in more accurate perception of unreleased stops in American English.[7]

These results are noteworthy because they provide evidence that instead of unilaterally putting non-native speakers at a disadvantage, cross-linguistic transfer from the native language can lead to a non-native advantage over native speakers. In light of the current literature on native versus non-native speech perception, this is a remarkable finding, which can be attributed to the role of canonicity in shaping perceptual biases in speech processing. To be specific, the canonicity of unreleased stops in Korean seems to hone the ability to utilize vocalic transition cues to a stop's place of articulation to a greater extent for Korean speakers than the high frequency of unreleased stops in American English seems to do so for American English speakers. Instead, for American English speakers the canonicity of released stops seems to result in some degree of reliance on consonantal burst cues to place of articulation in final stops, consistent with previous findings showing that English speakers tend to follow burst cues over transition cues when the two conflict with each other (Wang, 1959). This reliance on burst cues then makes the perception of unreleased stops relatively challenging because burst cues are not available for unreleased stops.

Crucially, the results of Experiment 2 showed that the Koreans were better than the Americans at identifying English unreleased stops in vowel contexts that are absent from Korean, suggesting that the cross-linguistic transfer resulting in their superior performance was general rather than specific to Korean-like sequences of speech segments. To put it another way, the Koreans' superior performance could not have been a convenient accident; they could not have performed better, for instance, by simply imagining they were listening to Korean because certain components of the stimuli such as the onset consonant [z] and the syllabus nucleus [ɑɹ] do not have phonetically close parallels in Korean, and, moreover, the stimuli were all embedded in an English sentential context. Instead the Koreans seem to have abstracted from their native language experience a general capacity for extracting information about a final stop from coarticulatory cues in a preceding vowel, which they are then able to apply to new vowels in a different language.

The results of Experiment 3 showed, furthermore, that the Koreans' perceptual advantage extended to an English lexical context, where they were clearly handicapped relative to the Americans. Native English lexical knowledge stood to provide considerable benefits to the Americans here as it would have allowed for better delimitation of the range of possible parses of a stimulus and made it more likely for the stimulus to be encoded lexically using long-term

phonological representations. For example, a stimulus [meɪ…] with a final stop ambiguous between /p/ and /k/ would have been quickly recognized as the word *make* by American listeners, who would have known that "mape" is not a real word of English (cf. the Ganong effect; Ganong 1980). By contrast, Korean listeners, given their non-native knowledge of the English lexicon, would have been less inclined to eliminate "mape" as a candidate parse because "mape" could, after all, simply be an English word unknown to them; incorrectly parsing an initial stimulus as such a non-word, moreover, would have made it more difficult for the stimulus to be held in memory to be compared with a second stimulus because memory for non-words is generally worse than memory for words (e.g., Hulme *et al.*, 1991). However, even when the Americans were provided with these benefits of their native lexical knowledge, they were not significantly better than the Koreans at detecting a change between different unreleased stops, and in fact, the Koreans were significantly better at detecting a change between the presence and absence of an unreleased stop.

Taken together, the results of Experiments 1–3 demonstrate that under the right circumstances cross-linguistic transfer from the native language can have a positive effect on non-native speech perception, boosting non-native perception beyond native levels. The implication is that the prevailing conceptualization of transfer in terms of either negative or neutral effects on non-native speech perception is an inadequate characterization of its perceptual consequences. Rather than exerting a specific kind of influence by nature, transfer seems to impact non-native perception in a manner that is highly dependent upon the alignment of properties in the native and non-native language. It is argued that positive effects of transfer have failed to be found as a result of the kinds of alignments that tend to be examined in the literature—namely, those that do not favor the native perceptual biases of non-native listeners. Here it has been demonstrated that when properties of the native and non-native language do align to favor native perceptual biases, transfer is beneficial to non-native perception.

The sort of alignment required for beneficial effects of transfer, however, is likely to involve not only the phonological level (e.g., the occurrence of unreleased final stops) but also other levels of linguistic patterning, such as the acoustic phonetic level (e.g., the specific coarticulatory patterns distinguishing final stops of different places of articulation). For instance, as suggested by the findings of Tsukada and colleagues (Tsukada *et al.*, 2007; Tsukada and Roengpitya, 2008), which show that Cantonese, Korean, Thai, and Vietnamese speakers differ in their perception of final stops, it is possible that other language groups that also have extensive native-language experience with unreleased final stops might not show the same amount of native-language transfer benefit for perception of unreleased English stops that Korean speakers show. If this turned out to be the case, such cross-linguistic differences in the native-language transfer benefit could be the product of cross-linguistic differences in the degree to which the native phonetic realization of unreleased stops resembles the non-native (English) realization of unreleased stops. In Vietnamese, for example, so-called

unreleased stops are orally unreleased but nasally released (Michaud *et al.*, 2006); moreover, coarticulatory patterns associated with final stops in Vietnamese may differ significantly from those in English. This could lead to different expectations of vowel-consonant formant transitions and, thus, different patterns of perception of these transitions in the processing of unreleased English stops. These factors may help to explain why—at least in the perception of unreleased Thai stops and released English stops—native Vietnamese speakers are consistently worse than native Korean speakers (Tsukada *et al.*, 2007), a disparity that might also be found in the perception of unreleased English stops.

Nevertheless, the current findings suggest that to the extent that alignments favoring native perceptual biases are common, native-language transfer has positive effects more often than research on second language acquisition would suggest. Thus in other cases of non-coincidence between frequency and canonicity, the expectation is that non-native speakers will display a similar perceptual advantage over native speakers. For example, nasalized vowels are a frequent, but non-canonical realization of vowels before final nasal consonants in American English, whereas they are perfectly fine phonemes in languages like French and Portuguese. Under the current view, therefore, it would be reasonable to predict that—all other things being equal—vowel nasalization in American English would be better perceived by French learners of English than by native speakers of English. This is the sort of empirical question that suggests fruitful avenues of further research into the full range of transfer effects, which appear to be much more varied than previously imagined.

ACKNOWLEDGMENTS

[1]On the contrary, depending on the actual distribution of English approximants with respect to the relevant acoustic parameters (in this case, F_2 and F_3), these data could mean that French listeners perceive English approximants less accurately than English listeners do. For instance, French listeners might be more inclined to incorrectly perceive a slightly /j/-like token of English /w/ as /j/, whereas an English listener might be more inclined to disregard the small deviance and correctly perceive the token as /w/.

[2]In Fig. 1 and elsewhere, the mid central vowel of American English is transcribed as the near-open central vowel [ɐ], following Roca and Johnson (1999). In particular, it is not transcribed as the open-mid back unrounded vowel [ʌ] because that symbol is reserved for transcribing the mid back unrounded vowel of Korean, which is significantly more back than the American English mid central vowel.

[3]In addition to replacement with a glottal stop, a final oral stop may also be produced with glottalization, which is left in the preceding vowel after audio editing. It is important to note, however, that any such glottalization trace of a final dereleased stop probably gave American listeners even more of an advantage over Korean listeners in Experiments 2 and 3

because as native listeners, they had more experience with glottalization as a cue to final stops in English.

[4]See Jaeger (2008) and Dixon (2008) for reasons to prefer mixed-effects logistic regression to analysis of variance for the analysis of accuracy data.

[5]Note that correct selection of the "other" response option for sonorant-final items in Experiments 1 and 2 is more precisely described as accurate identification of the occurrence of other final segments than as accurate identification of final sonorants. Recall from Sec. II C 2 that in Experiments 1 and 2, listeners were given three response options for specifically identifying the final stops in stop-final items (/p, t, k/), but only one "other" response option for identifying final sonorants as something other than /p, t, k/. Consequently, listeners' "other" responses for sonorant-final items did not, strictly speaking, indicate accurate identification of final sonorants because the other response option was also consistent with other possible percepts (e.g., nasals, voiced plosives) that would have been incorrect for the vowels and approximants that terminated the sonorant-final items.

[6]Individual place contrasts differed with respect to which group had the higher d'. On the /p/-/t/ contrast, the Koreans (mean $d' = 1.09$) were slightly worse than the Americans (mean $d' = 1.22$); on the /t/-/k/ contrast, the Koreans (mean $d' = 1.19$) were slightly better than the Americans (mean $d' = 1.09$); and on the /p/-/k/ contrast, the Koreans (mean $d' = 1.31$) and the Americans (mean $d' = 1.32$) were nearly identical. However, consistent with the overall pattern, the effect of Group on d' was not significant for any of the individual place contrasts: /p/-/t/ [$\beta = -0.124$, $t = -1.017$, $P = 0.314$], /t/-/k/ [$\beta = 0.098$, $t = 1.024$, $P = 0.311$], or /p/-/k/ [$\beta = -0.015$, $t = -0.149$, $P = 0.882$].

[7]Thanks to Bruce Hayes for pointing out a possible alternative interpretation of these data attributing the Americans' relatively poor performance to the type of speech they were hearing (laboratory speech) in which listeners might reasonably expect final stops to be produced with release. There are two reasons why this is not a convincing explanation of the findings. First, as found by Byrd (1993), even in read speech utterance-final stops in American English are frequently produced as unreleased (e.g., 51% of the time for final /p/), so strong expectations of release do not follow from the statistics of final stop realization in relatively careful speech. Second, if one was to argue that the statistics of final stop realization in careful speech varieties such as laboratory speech are significantly different from those documented in studies such as Byrd (1993), the fact remains that non-native listeners also have access to these statistics, albeit perhaps not to the same degree as native listeners. That is to say, there is no reason to expect only native English speakers to have been misled by the statistics of careful English to anticipate release bursts in Experiments 2 and 3; non-native English speakers should have been misled in a similar way. In addition, it would have been fairly clear after the practice session in each experiment that final stops were consistently being produced without release, so if anything this should have led all listeners to expect final stops to be unreleased. These facts suggest that the difference found between Americans and Koreans in perception of English unreleased stops is unlikely to be the product of Americans' expectations of a release burst that faithfully reflect the statistics of the type of speech they were hearing.

Abramson, A. S., and Tingsabadh, K. (**1999**). "Thai final stops: Cross-language perception," Phonetica **56**, 111–122.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (**2008**). "Mixed-effects modeling with crossed random effects for subjects and items," J. Mem. Lang. **59**, 390–412.

Bent, T., and Bradlow, A. R. (**2003**). "The interlanguage speech intelligibility benefit," J. Acoust. Soc. Am. **114**, 1600–1610.

Bent, T., Bradlow, A. R., and Smith, B. L. (**2008**). "Production and perception of temporal patterns in native and non-native speech," Phonetica **65**, 131–147.

Best, C. T. (**1994**). "The emergence of native-language phonological influences in infants: A perceptual assimilation model," in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by J. C. Goodman and H. C. Nusbaum (MIT Press, Cambridge, MA), pp. 167–224.

Best, C. T. (**1995**). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 171–204.

Best, C. T., and Tyler, M. D. (**2007**). "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, edited by O.-S. Bohn and M. J. Munro (John Benjamins Publishing, Amsterdam, The Netherlands), pp. 13–34.

Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., and Ross, S. (**2011**). "What makes listening difficult? Factors affecting second language listening comprehension," Technical Report TTO 81434 E. 3.1 (University of Maryland Center for Advanced Study of Language, College Park, MD).

Boersma, P., and Weenink, D. (**2011**). "PRAAT: Doing phonetics by computer (version 5.3) [computer program]," http://www.praat.org (Last viewed 2/28/2012).

Bohn, O.-S., and Best, C. T. (**2012**). "Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners," J. Phonetics **40**, 109–128.

Bradlow, A. R., and Pisoni, D. B. (**1999**). "Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors," J. Acoust. Soc. Am. **106**, 2074–2085.

Byrd, D. (**1993**). "54,000 American stops," UCLA Work Pap. Phonetics **83**, 97–116.

Cho, T., and McQueen, J. M. (**2006**). "Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants," J. Acoust. Soc. Am. **119**, 3085–3096.

Cutler, A. (**2001**). "Listening to a second language through the ears of a first," Interpreting **5**, 1–23.

Cutler, A., Garcia Lecumberri, M. L., and Cooke, M. (**2008**). "Consonant identification in noise by native and non-native listeners: Effects of local context," J. Acoust. Soc. Am. **124**, 1264–1268.

Davidson, L. (**2011**). "Characteristics of stop releases in American English spontaneous speech," Speech Commun. **53**, 1042–1058.

Davies, M. (**2008**). "The corpus of contemporary American English: 425 million words, 1990-present," http://corpus.byu.edu/coca/ (Last viewed 2/28/2012).

Dixon, P. (**2008**). "Models of accuracy in repeated-measures designs," J. Mem. Lang. **59**, 447–456.

Flege, J. E. (**1995**). "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 233–272.

Flege, J. E. (**2003**). "A method for assessing the perception of vowels in a second language," in *Issues in Clinical Linguistics*, edited by E. Fava and A. Mioni (Unipress, Padova, Italy), pp. 19–43.

Ganong, W. F. (**1980**). "Phonetic categorization in auditory perception," J. Exp. Psychol. Hum. Percept. Perform. **6**, 110–125.

Garcia Lecumberri, M. L., and Cooke, M. (**2006**). "Effect of masker type on native and non-native consonant perception in noise," J. Acoust. Soc. Am. **119**, 2445–2454.

Goh, C. C. M. (**2000**). "A cognitive perspective on language learners' listening comprehension problems," System **28**, 55–75.

Goto, H. (**1971**). "Auditory perception by normal Japanese adults of the sounds 'L' and 'R,'l" Neuropsychologia **9**, 317–323.

Hallé, P. A., Best, C. T., and Levitt, A. (**1999**). "Phonetic vs. phonological influences on French listeners' perception of American English approximants," J. Phonetics **27**, 281–306.

Han, M. (**1992**). "The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers," Phonetica **49**, 102–127.

Hulme, C., Maughan, S., and Brown, G. D. A. (**1991**). "Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span," J. Mem. Lang. **30**, 685–701.

Ingram, J. C. L., and Park, S.-G. (**1997**). "Cross-language vowel perception and production by Japanese and Korean learners of English," J. Phonetics **25**, 343–370.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., and Siebert, C. (**2003**). "A perceptual interference account of acquisition difficulties for non-native phonemes," Cognition **87**, B47–B57.

Jaeger, T. F. (**2008**). "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models," J. Mem. Lang. **59**, 434–446.

Johnson, K. (**2008**). *Quantitative Methods in Linguistics* (Blackwell Publishing, Malden, MA).

Ko, I. (**2009**). "The merger of ey /e/ and ay /ɛ/ of Seoul Korean," Eoneo Jeongbo [Lang. Inform.] **10**, 73–89.

Lado, R. (**1957**). *Linguistics Across Cultures: Applied Linguistics for Language Teachers* (University of Michigan Press, Ann Arbor, MI).

Lee, H. B. (**1993**). "Korean," J. Int. Phonetic Assoc. **23**, 28–31.

Lisker, L. (**1999**). "Perceiving final voiceless stops without release: Effects of preceding monophthongs versus nonmonophthongs," Phonetica **56**, 44–55.

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

Michaud, A., Ngọc, T. V., Amelot, A., and Roubeau, B. (**2006**). "Nasal release, nasal finals and tonal contrasts in Hanoi Vietnamese: An aerodynamic experiment," Mon-Khmer Stud. **36**, 121–137.

Mora, J. C. (**2005**). "Lexical knowledge effects on the discrimination of non-native phonemic contrasts in words and non-words by Spanish/Catalan bilingual learners of English," in *Proceedings of the ISCA Workshop on Plasticity in Speech Perception (PSP2005)*, edited by V. Hazan and P. Iverson (International Speech Communication Association, London, UK), pp. 43–46.

Nábělek, A. K., and Donahue, A. M. (**1984**). "Perception of consonants in reverberation by native and non-native listeners," J. Acoust. Soc. Am. **75**, 632–634.

Newell, A., and Rosenbloom, P. S. (**1981**). "Mechanisms of skill acquisition and the law of practice," in *Cognitive Skills and Their Acquisition*, edited by J. R. Anderson (Erlbaum Associates, Hillsdale, NJ), Chap. 1, pp. 1–55.

Odlin, T. (**1989**). *Language Transfer: Cross-Linguistic Influence in Language Learning* (Cambridge University Press, Cambridge, UK).

Psychology Software Tools, Inc. (**2002**). "E-PRIME, version 2.0," http://www.pstnet.com (Last viewed 2/28/2012).

Roca, I., and Johnson, W. (**1999**). *A Course in Phonology* (Blackwell Publishing, Malden, MA).

Sheldon, A., and Strange, W. (**1982**). "The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception," Appl. Psycholing. **3**, 243–261.

Sohn, H.-M. (**1999**). *The Korean Language, Cambridge Language Surveys* (Cambridge University Press, Cambridge, UK).

Sumner, M., and Samuel, A. G. (**2005**). "Perception and representation of regular variation: The case of final /t/," J. Mem. Lang. **52**, 322–338.

Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R., and Munhall, K. G. (**2008**). "Training English listeners to perceive phonemic length contrasts in Japanese," J. Acoust. Soc. Am. **123**, 397–413.

Tsukada, K. (**2006**). "Cross-language perception of word-final stops in Thai and English," Bilingualism: Lang. Cognit. **9**, 309–318.

Tsukada, K., Nguyen, T. T. A., Roengpitya, R., and Ishihara, S. (**2007**). "Cross-language perception of word-final stops: Comparison of Cantonese, Japanese, Korean and Vietnamese listeners," in *Proceedings of the 16th International Congress of Phonetic Sciences*, edited by J. Trouvain and W. J. Barry (Pirrot, Dudweiler, Germany), pp. 1781–1784.

Tsukada, K., and Roengpitya, R. (**2008**). "Discrimination of English and Thai words ending with voiceless stops by native Thai listeners differing in English experience," J. Int. Phonetic Assoc. **38**, 325–347.

Wang, W. S.-Y. (**1959**). "Transition and release as perceptual cues for final plosives," J. Speech Hear. Res. **2**, 66–73.

Yamada, R. A. (**1995**). "Age and acquisition of second language speech sounds: Perception of American English /r/ and /l/ by native speakers of Japanese," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 305–320.

Yamada, R. A., and Tohkura, Y. (**1992**). "Perception of American English /r/ and /l/ by native speakers of Japanese," in *Speech Perception, Production and Linguistic Structure*, edited by Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Ohmsha, Tokyo, Japan), pp. 155–174.

Yamada, R. A., Tohkura, Y., and Kobayashi, N. (**1996**). "Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese," in *Second-Language Speech: Structure and Process*, edited by A. James and J. Leather (Mouton de Gruyter, Berlin, Germany), pp. 103–117.