Journal of Memory and Language 81 (2015) 72–90

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design

Jiwon Hwang [a,*], Susan E. Brennan [b], Marie K. Huffman [a]

[a] Department of Linguistics, Stony Brook University, United States
[b] Department of Psychology, Stony Brook University, United States

ARTICLE INFO

Article history:
Received 15 June 2013
revision received 29 December 2014
Available online 14 February 2015

Keywords:
Phonetic adaptation
Audience design
Priming
Pragmatic disambiguation
Korean
Non-native spoken dialogue

ABSTRACT

To be understood, non-native speakers must adapt their speech in order to produce contrasts in their second language (L2) that are not present in their first language (L1). Here we examine mechanisms hypothesized to facilitate such adaptation within spoken dialogue: *priming*, *affiliation*, and *audience design*. In two experiments, Korean non-native speakers of English interacted in a referential communication task with a Korean English-speaking confederate (Experiment 1) and a monolingual American English-speaking confederate (Experiments 1 and 2). The task required them to spontaneously produce labels containing segments from English that do not exist in Korean (/æ/ and coda /b/), which, when spoken with a Korean accent, can result in ambiguous homophones (e.g., *pat* pronounced like *pet*, or *mob* pronounced like *mop*). The Koreans produced more English-like phonetic segments not only immediately after hearing similar segments primed by the American partner, but also when the task required the partner to distinguish two potentially ambiguous items. The first time the Koreans referred to potentially ambiguous objects, utterances took longer to initiate; once they were aware of the potential for ambiguity, initiating contrasting labels took no more time than initiating labels primed by the partner. Findings suggest that priming effects in dialogue are not obligatory but may be motivated, and that phonetic adaptation is shaped by awareness of a partner's pragmatic needs.

Introduction

One of the major challenges to understanding how people process and represent speech comes in the form of the enormous variability an individual encounters on a daily basis. Sources of this variability include speaker physiology, dialect, and the speaker's language background. Variability is particularly challenging when native and non-native speakers come into contact; when a non-native speaker's first language (L1) lacks a phonetic segment or segmental contrast included in the second, target language (L2), the result can range from mildly to heavily foreign-accented speech. Virtually everyone encounters foreign-accented speech; given the increasing interconnectedness and mobility of the worlds' populations, this phenomenon can only be increasing.

One way in which speakers respond to variability is by adapting the way in which they speak, often producing forms that they have just heard from another speaker. For example, after hearing a double object construction (e.g., *I gave you the book*), speakers tend to reproduce the same syntactic structure (e.g., *He threw her the ball*) rather than a prepositional phrase (e.g., *He threw the ball to her*) (Bock, 1996; Branigan, Pickering, & Cleland, 2000). At the

* Corresponding author at: Department of Linguistics, Stony Brook University, Stony Brook, NY 11794-4376, United States.
E-mail address: jiwon.hwang@stonybrook.edu (J. Hwang).

http://dx.doi.org/10.1016/j.jml.2015.01.001
0749-596X/

lexical level, two partners in conversation typically come to use the same or similar referring expressions, providing evidence that they share a perspective and are referring to the same thing (Bortfeld & Brennan, 1997; Brennan & Clark, 1996; Garrod & Anderson, 1987). And at the phonetic level, speakers often (albeit not always) adapt their pronunciation toward that of a partner or even to that of passively heard speech (e.g., Babel, 2010; Giles & Powesland, 1975; Kim, 2012; Kim, Horton, & Bradlow, 2011; Pardo, Cajori, & Krauss, 2010; Pardo, Gibbons, Suppes, & Krauss, 2012; Willemyns, Gallois, Callan, & Pittam, 1997). Here, we examine the forces underlying adaptation in interactive spoken dialogue, through the lens of accented speech produced by non-native speakers.

Theoretical accounts of adaptation

Of theoretical interest is why and how adaptation happens; several accounts have been proposed. A *priming* account says that adaptation occurs automatically and passively, after a speaker is primed with a similar form. This sort of explanation underlies the "output–input coordination" account of Garrod and Anderson (1987), as well as the view that adaptations are "generic" or driven by what is easiest for speakers, even when helpful to addressees (Brown & Dell, 1987; Dell & Brown, 1991). More recently, priming was incorporated as a cornerstone of the "interactive alignment" account of Pickering and Garrod (2004). To the extent that adaptation is driven by priming, this suggests that it is inflexible and possibly encapsulated from slower, higher-level pragmatic influences (e.g., see proposals by Brown & Dell, 1987, Barr & Keysar, 2002, and Pickering & Garrod, 2004). Although the interactive alignment account was modified somewhat in Costa, Pickering, and Sorace (2008) to account for lower rates of adaptation on the part of non-native speakers, these proposals still posit priming to be a default process, with pragmatic adaptation taking additional time. The interactive alignment account has been offered as a general explanation at multiple linguistic levels of adaptation in dialogue, including at the lexical, syntactic, and phonological levels (see also Pickering & Garrod, 2013, for an update that focuses on automatic co-activation of neural systems for perception and action rather than on a "priming" explanation per se).

Another account, which we will call *affiliation*, focuses on sociolinguistic forces (such as those addressed by accommodation theory, Giles & Powesland, 1975), including the status and group identity of a partner, and the speaker's relationship with that partner. On this approach, adaptation in speaking is considered to be due to factors such as the desire to show solidarity, or to affiliate with, or to mark one's membership in, a social group or category (Beebe, 1981; Giles & Powesland, 1975; Gumperz, 1982). This approach accounts for not only how speakers tend to become more similar in their speech patterns, but also how they may diverge in situations in which they disagree or do not wish to affiliate with one another (see, e.g., Babel, 2010, 2012; Bly, 1993; Bourhis & Giles, 1977; Kim, 2012; Kim et al., 2011; Willemyns et al., 1997). The affiliation account predicts that adaptation is used to actively define and acknowledge speakers' identities and relationships.

On the third account, which we will call audience design,[1] adaptation is shaped in a way that is partner-specific, in that it is responsive to the perceived needs of a partner at a particular point in the conversation. For example, partners in a conversation typically come to *entrain* on the same terms in referring to an object, a way of marking that they believe they share a conceptual perspective on it (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Metzing & Brennan, 2003). This adaptation is flexible; that is, a previously entrained-upon referring expression is revised when a change of context requires more information to uniquely identify the referent or (sometimes) when a new addressee enters the conversation. When context changes such that a *less* informative expression could suffice to identify the referent, speakers tend to continue using the over-informative entrained-upon expression with the same partner (rather than break the conceptual pact they have established), but less so with a new partner (Brennan & Clark, 1996). This effect has also been shown in comprehension: Addressees experience interference or delay when a speaker appears to abandon a previously entrained-upon expression and uses a new expression for no apparent reason, but not when a *new* speaker uses the same new referring expression (Metzing & Brennan, 2003; replicated by Matthews, Lieven, & Tomasello, 2010).

The theoretical mechanisms underlying the three accounts outlined above—priming, affiliation, and audience design—need not be mutually exclusive of course, but could work in concert to shape adaptation in spoken dialogue. However, the priming account differs from the other two in its predictions about the *timing* with which such adaptation occurs. Most accounts that appeal to priming as an explanation argue specifically for a modular, two-stage architecture. These include the interactive alignment proposal with its immediate priming and delayed use of "full common ground" (Pickering & Garrod, 2004); the *dual process hypothesis* (Bard et al., 2000; Bard & Aylett, 2000), in which automatic processes (such as articulation) are considered to be obligatory and therefore not influenced by audience design; and the *perspective adjustment* account (including Horton & Keysar's, 1996 *monitoring-and-adjustment* model for production and Barr & Keysar's, 2002 *anchoring and adjustment* heuristic for comprehension) in which a fast-acting, inflexible, "egocentric" stage precedes a slow-acting, inferential, partner-specific stage (see also Brown & Dell, 1987). This implies that adaptation due to priming should take place rapidly, whereas an adaptive response that is not preceded by a prime should take place more slowly.

Explanations that attribute adaptation to sociolinguistic factors (such as affiliation) do not necessarily make a commitment a priori to any particular psychological model, so do not lead to specific predictions about the timing with which inferences are made (although some, e.g., Babel, 2010 and Kim, 2012, examine both issues in tandem).

[1] This term was coined by Bell (1984) to cover a wide variety of partner-specific influences on speaking, including "all a person's attributes, psychological and social, permanent and temporary" (p. 169). Here we use it in the sense of being responsive to a partner's needs at a particular point in the conversation.

Audience design accounts that do make claims about cognitive architecture tend to predict that partner-specific information (e.g., knowledge about a partner's needs, or common ground with a partner) can act like any other source of information, probabilistically constraining processing from the earliest moments (e.g., Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003; Kraljic & Brennan, 2005; Metzing & Brennan, 2003). This does not mean that the system will always succeed in taking the partner into account—cognitive systems are subject to processing limitations, interference, and errors—but if partner-specific information is *available* (known, salient, already computed) early enough in processing, there is no architectural barrier that prevents the system from using it.

Adaptation in conversations between native and non-native speakers

When conversing in a second language, non-native speakers produce phonetic, lexical and syntactic forms that differ systematically from those of native speakers, setting the stage for adaptation. Relatively few studies document how non-native speakers adapt their pronunciation in conversation with native speakers (see reviews of theoretical and methodological issues in Beebe & Giles, 1984; Costa et al., 2008). Beebe (1981) looked at interviews with 61 Thai children bilingual in Thai and Chinese, conducted by an ethnic Thai interviewer and an ethnic Chinese interviewer who spoke Thai without an accent; Beebe concluded that the children used Thai vowels less often with the Chinese interviewer than with the Thai interviewer and attributed this to ethnic accommodation (related to what we call *affiliation*), claiming to have ruled out imitation (related to what we call *priming*). Kim et al. (2011) examined adaptation using perceptual judgments of words excised from native–native and native–non-native conversation, finding alignment toward a partner's accent in some cases and divergence in others. These results add to the evidence suggesting that adaptation need not be an obligatory, automatic process (see also Babel, 2011, 2012; Kim, 2012). If priming does not automatically drive phonetic adaptation, then it is all the more important to understand which factors interact with or interfere with priming.

The current project

In this paper, we examine phonetic adaptation by non-native speakers as a means of testing and differentiating different accounts of adaptation. Many second language learners produce L2 pronunciations that are notably unlike those of native speakers, often showing a heavy influence of L1 sound inventory and sound patterning (e.g., Best, McRoberts, & Goodell, 2001; Brannen, 2002; Clements, 2001; Eckman & Iverson, 2013; Flege & Eefting, 1987; Hancin-Bhatt, 1994; Sirsa & Redford, 2013; White & Mattys, 2007). Our experimental design took advantage of the fact that ambiguities arise when non-native speakers fail to make necessary contrasts in the target language. For example, Korean lacks the voicing contrast that distinguishes English sounds such as /p/ and /b/ at the end of a syllable, as well as the vowel contrast between /æ/ and /ɛ/ (Sohn, 1999). This leads many Korean speakers of English to produce their L2 /b/ much like their L1 /p/, pronouncing *mo**b*** as something that sounds like *mo**p*** (Major & Faudree, 1996), and producing their L2 /æ/ much like their L1 /ɛ/, pronouncing *p**a**t* as something that sounds like *p**e**t* (Flege, Bohn, & Jang, 1997; Yang, 1996). In Korean-accented English, the failure to make such contrasts can lead to ambiguity in the form of unintended (by speakers) homophones (as perceived by addressees), making the accented speech especially difficult to understand. We are particularly interested in discovering the conditions under which adaptation may result in more native-like pronunciation by non-native speakers.

Because priming, affiliation, and audience design have all been associated with adaptation, we attempted to tease these explanations apart in two experiments investigating phonetic adaptation by naive non-native speakers of English whose L1 was Korean. We took advantage of phonological differences between English and Korean and quantified phonetic adaptation with acoustic measures that could reveal the effects of one or more of the factors outlined above. In both experiments, subjects participated in a spontaneous conversational task with a confederate in which they matched cards labeled with single-syllable forms conforming to English phonology, the critical items of which contained segments missing phonemically from Korean, /æ/ or final /b/. Experiment 1 aimed to test for effects of priming of these phonetic segments along with effects of affiliation based on whether language background was shared with the confederate partner (who was either a Korean speaker of English like the subjects, or else an American monolingual native speaker of English). The Korean confederate was a fluent bilingual who, during the experimental sessions, spoke English with a strong Korean accent (where final /b/ is produced like /p/ and /æ/ is produced like /ɛ/). The monolingual American English speaking confederate exhibited a clear contrast between /b/ and /p/ and between /æ/ and /ɛ/. Each subject interacted with both confederates in two subsequent sessions (counterbalanced for order). Experiment 2 aimed to replicate the priming effect of Experiment 1, along with manipulating the partner's pragmatic need to uniquely identify a referent in a particular context (an audience design effect). The second experiment had only one confederate, the monolingual American English speaker.

General predictions

To the extent that phonetic adaptation is driven by low-level priming, non-native speakers may better approximate the target segments in L2 that are lacking in their L1 immediately after they hear their American English-speaking partner produce the segment, but not when the previous utterance lacked this segment (either because the previous utterance by the American partner contained no similar segment to prime the target segment, or because the previous utterance was produced by the Korean partner using L1 phonology). To the extent that adaptation is driven by the partner's identity as a member of the same or a different language community (as in Beebe, 1981), Korean speakers of English might simply use more Korean phonology to

their Korean partner and more English-like phonology to their American partner (regardless of what they hear immediately before producing the target item). If both affiliation and priming are at work (Experiment 1), the non-native speakers should adapt toward the American partner's targets not only right after being primed, but also in the absence of priming with similar segments (which would lead to a difference between the Korean vs. American confederate baseline conditions, albeit less of a difference than between the two priming conditions). To the extent that adaptation is driven by dynamic factors having to do with audience design—the pragmatic needs of a partner at a particular point in the dialogue (Experiment 2)—speakers may use more English-like phonology when it is needed to make a contrast that enables the addressee to uniquely identify a referent. Again, such a pragmatic effect need not be mutually exclusive with a priming effect. If both emerge in Experiment 2, then comparing the time to initiate speaking in critical turns (pragmatic vs. priming conditions) may shed light on whether audience design can be a fast-acting process (as opposed to a late repair).

Experiment 1

Experiment 1 tested for effects of priming and affiliation by measuring acoustic parameters of target items (containing /æ/ or final /b/) spontaneously produced by the subjects and comparing these (within-speakers) across conditions. In two back-to-back referential communication sessions, the Korean subjects conversed in English during a matching task done with the Korean and American confederate partners. Just before critical conversational turns, subjects were prompted by the confederate partner who asked what was next to a landmark labeled with a prime that rhymed with a target item the subject was to say (Priming trials). Other turns were prompted by the confederate asking about a landmark labeled with an unrelated form (Baseline trials). Specifically, if the language background of the partner is what primarily drives Korean speakers' phonetic adaptation toward making L2 contrasts, speakers should produce more English-like /b/ and /æ/ with the American partner than with the Korean partner, regardless of what landmark the partner has just produced (e.g., for both Baseline and Priming trials). If adaptation is *not* driven by affiliation, then Korean speakers' baseline forms should be no more English-like when addressed to the native-English-speaking partner than to the native-Korean-speaking partner. If priming drives phonetic adaptation, Korean speakers should produce more English-like /b/ and /æ/ when they have just heard phonetically relevant landmark labels produced by the native-English-speaking confederate (Priming condition) and possibly less English-like forms when they have just heard phonetically relevant landmarks produced by the native-Korean-speaking partner.

Method

The subject/director was seated before a board with cards in a prearranged pattern, and the confederate/matcher was seated behind a barrier before a similar board with half of the cards already in place (serving as potential landmarks). The matcher asked questions in English to locate target cards, referring to the landmark cards (e.g., What is below *Hob*?), and the director answered by pronouncing the appropriate label (e.g., *Gob*.).

Materials and design

Pairs in each round used two 5-cell × 6-cell boards like those in Figs. 1a and 1b, along with two identical sets of five cards (one for the matcher and one for the director) on which English strings were written. Two of these were critical target cards displaying labels that included the L2 target sounds, and three were filler cards. Each subject completed a total of eight rounds with each of the two confederates; each round used a different board and set of cards. Cards were distributed to two lists (A or B) that contained the same types of labels, as shown in Table 1; half of the subjects used List A with the Korean partner and List B with the native English-speaking partner, and the other half of the subjects used these lists with the other partner.

Experimental items were one-syllable labels consisting of words and pronounceable English non-words that contained sounds requiring a phonemic contrast that does not exist in Korean and is therefore likely to be ambiguous

		Eye		
		Moon		
Cap	Bap			
		By		Hob
Beat		Zin		Gob
Yon				

Fig. 1a. A sample board for directors (subjects) in Experiment 1: Labels highlighted in beige were initially present on both the director's and (confederate) matcher's boards and were mentioned as landmarks by the matcher in her questions to prompt the director's responses; the rest are targets.

		Eye		
		?		
Cap	?			
		?		Hob
?		Zin		?
Yon				

Fig. 1b. A sample board for matchers (confederates) in Experiment 1: The confederates referred to the beige landmark cards in order to ask the subjects about which card to place in each cell marked with "?" (During the experiment, the beige highlighting was present on both boards.)

Table 1
/b/ and /æ/ experimental items (with landmark listed directly above each item) for each Round, distributed to Lists A and B in Experiment 1. Items in white appeared in the baseline condition, and those in gray were in the priming condition. Each subject did one list with the Korean confederate and one with the English confederate, counterbalanced for order.

	Round	1	2	3	4	5	6	7	8
List A	Landmark	Pix	Joy	Vox	Bey	Hob	Kib	Kob	Tib
	/b/ items	Nib	Dob	Sib	Job	Gob	Jib	Bob	Mib
	Landmark	Fox	Ton	Hum	Kin	Cap	Pack	Cat	Zap
	/æ/ items	Pat	Hat	Nap	Dack	Bap	Sack	Mat	Tap
List B	Landmark	Cue	Box	Coz	Fix	Kob	Pib	Tob	Lib
	/b/ items	Dib	Mob	Zib	Sob	Nob	Gib	Zob	Bib
	Landmark	Bun	Win	Mot	Hut	Lap	Tack	Gap	Zat
	/æ/ items	Hack	Sat	Nack	Dap	Sap	Gack	Map	Bat

in Korean accented English speech (L2). Items were chosen based on their phonetic segments; they were intended to be treated as abstract labels or names that made it possible to refer to the cards, rather than as lemmas (in fact, subjects were unfamiliar with the lexical status of many of the labels, as quantified in Experiment 2). These critical labels contained /b/ in coda position or else the vowel /æ/. /b/ is likely to be produced as [p] because stop voicing contrasts are neutralized to voiceless unaspirated stops in coda position (although Korean does have a three-way contrast on stops in onset position (e.g., aspirated /p^h, t^h, k^h/, unaspirated /p, t, k/ and fortis /p^*, t^*, k^*/)). The other voiced stops, /d/ and /g/, were excluded because post-vocalic /b/ is more likely to be devoiced in Korean-accented English (Hwang, 2011; Major & Faudree, 1996). The /b/ experimental items were generated to have onsets and vowels that are likely to be easy for Korean speakers to pronounce: onsets were consonants that exist in Korean and the vowel was either [i] or [ɑ], which have correspondents in Korean (Yang, 1996). The /æ/ items were generated to have a voiceless stop in the coda. Fillers were all one-syllable words and ended with a vowel, a voiceless stop, or a nasal. A complete list of the experimental items, fillers and landmarks is in Appendix A.

The *Baseline* condition was designed to measure the extent to which the naive subjects naturally produced the L2 contrast missing from their L1; for this reason, all subjects experienced that condition first, in Rounds 1–4. In the Baseline condition, all of the landmarks were phonetically irrelevant to the critical targets; none of the landmarks in the first four rounds included any of the key phonemes, coda /b/ and the vowel /æ/ (examples are shown in Fig. 2). In each round of the *Priming* condition, which each subject experienced from Rounds 5–8, two of the landmarks that were pronounced by the confederate in the conversational turn immediately before the subject produced the target labels had forms that were phonetically similar to these target labels, to test for immediate phonetic priming to the L2 sounds. Neither the 3 filler items on each board nor their landmarks contained the key phonetic segments /b/ or /æ/; however, to break any pattern that subjects might otherwise perceive in the stimuli, two of the three filler items in each of the first four rounds had phonetically similar landmarks. In each of the last four rounds, three other landmarks were phonetically irrelevant to their filler targets.

To set up the test for effects of partner affiliation, in one session, the Korean-speaking subjects directed a Korean-speaking confederate who used a heavily Korean-accented English phonology, and in the other session, they directed a native English-speaking confederate. The order of these sessions was counterbalanced (half of subjects interacted with the Korean partner first and half with the American, with one session immediately following the other). Thus, this partner order factor was the only between-subject factor while all the others were within-subject factors.

Procedures

Upon arriving in the lab, the subject was introduced to the partner, instructed about the communication task by the experimenter, and assigned to the role of director. Both partners knew that they had copies of the same card set. The subject/director was instructed to answer the questions of the matcher so that the matcher could accurately arrange her cards in the same order as the subject's. Subjects were informed that they would be conducting the card arrangement task multiple times with the same partner, and would then do the same task again in a session with a different partner. At the beginning of each session, the pair spent about 2 or 3 min getting to know one another by talking about everyday topics (weather, majors, housing, etc.). Before doing the experimental rounds, the director and first matcher took part in a short practice round where a set of 3 cards had to be matched. Once subjects were familiar with the task and the role of director, they did 8 rounds of the matching task with each matcher. After each round, the director and matcher together compared their boards to check that the cards were matched correctly; no errors were made.

Participants

Subjects

Twenty graduate and undergraduate students (7 male and 13 female) from Stony Brook University participated as naive subjects in the director role and received $10 or research credit in a psychology course for their participation. All were 18 years of age or older and native speakers of Korean, and all gave informed consent before participating. Subjects' accent in English was rated by the English-speaking confederate at the end of the experiment on a scale of 0 (*None*/*Weak*) to 5 (*Strong Accent*), with an average accentedness rating of 3.88. Fig. 3 shows the distribution of

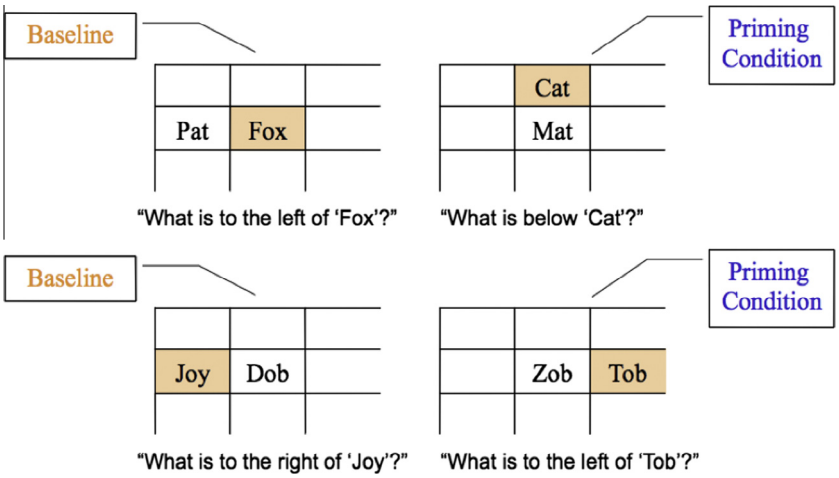

Fig. 2. Examples from directors' (subjects') boards for Baseline and Priming conditions in Experiment 1.

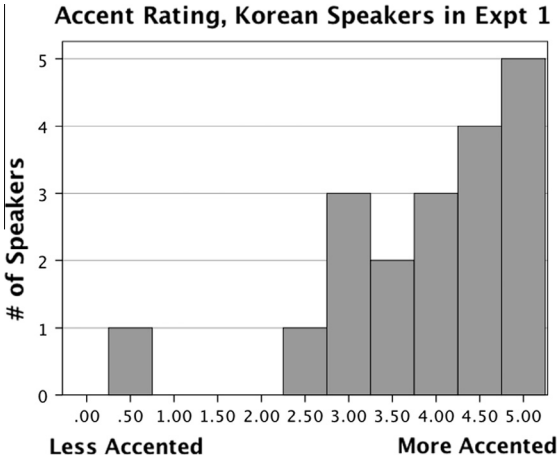

Fig. 3. Accent ratings for Korean speakers of English (naive subjects) in Experiment 1.

subjective ratings for 19 of the 20 Korean speakers of English who participated in Experiment 1 (note the strong skew, as most subjects were perceived to have heavy accents; one subject was not rated by mistake). The one subject who spoke lightly accented English showed the same patterns of effects as the rest of the speakers who had heavier accents.

Confederates

Two confederates (a monolingual speaker of English and a native speaker of Korean) participated as matchers. The monolingual English speaker was a 24 year-old female speaker of American English from the New York area. The Korean-speaking confederate was a 27 year-old female living in New York, whose first language was Korean and second language was English. The native English speaker was not previously acquainted with any of the subjects; the Korean-speaking confederate was previously acquainted with only one. The confederates' status was not disguised; subjects were told that their partner worked in the lab. The confederate partners actually performed the matching task and so had an authentic need to communicate with the subjects (following Kuhlen & Brennan, 2013's recommendations about using confederates in dialogue experiments).

Measures and coding

All of the dialogues from the two sessions were recorded using a Marantz PMD 660 digital recorder at a 44,100 Hz sampling rate. Directors and matchers wore Shure SM10ACN head-mounted dynamic microphones for recording speech. The Praat analysis program (Boersma & Weenink, 2007) was used for acoustic measures, which were made by the first author and other trained phoneticians in the lab who were blind to the experimental conditions.

Consonants: /b/ experimental items. Two acoustic characteristics of coda voicing were measured in order to compare experimental items in different conditions. Vowel duration is the primary cue that English native speakers use in distinguishing voicing of coda stops (Hogan & Rozsypal, 1980; Peterson & Lehiste, 1960; Raphael, 1972); vowels before voiced stops typically have longer duration than before voiceless stops. We used F2 onset and offset from a spectrogram with a 5 ms window length as the basis for measuring vowel onset and offset. In cases where F2 was not a clear cue for vowel onset or offset, we took the place where two or more formants appeared or dropped out together as the onset and offset, respectively. A second acoustic cue that English native speakers use for the coda voicing distinction is closure voicing duration (Jones, 2003; Lisker, 1957; Nittrouer, 2004); voiced stops have longer closure voicing than voiceless stops. Closure voicing duration was measured as the period from vowel offset to where the periodicity ended in the waveform.

Vowels: /æ/ experimental items. Three acoustic measures were made for the /æ/ experimental items: duration, F1 and F2. For native speakers, /æ/ has a longer vowel duration than /ɛ/ (e.g., Bohn & Flege, 1990; Peterson &

Lehiste, 1960). To measure vowel duration, the same general procedure described above was used to identify vowel onset and offset for the cases in which the consonant before a vowel was voiced (e.g., *Bat*). After voiceless fricatives and aspirated stops (e.g., *Sack* or *Tap*), which often showed formant structure within the noise, an alternative method was used; the first point at which modal voicing began was taken as the vowel onset based on both waveforms and spectrograms. Another important acoustic difference between /æ/and /ɛ/ is formant values: higher F1 and lower F2 are expected for /æ/ than for /ɛ/. The formant measurements were done by a Praat script that uses linear predictive coding to identify formants and then automatically measures F1 and F2 with a 2.5 ms step size and 25 ms window size at the midpoint of the vowel (since the vowels in the experimental items were all monophthongs). To rule out error in these automated measures, any tokens beyond 2 standard deviations from the mean in each of the conditions were re-measured by hand. This was done separately for F1 and F2 in four different groups by the kind of vowel and by the gender (female or male). No data were excluded. After this, formants were normalized to the Bark scale (a nonlinear scale for frequency that correlates with perceived pitch) using the formula in Traunmüller (1990).

Confederates' speech

The American monolingual confederate produced clear final /b/ and /æ/; the Korean confederate intentionally used her Korean phonology for these segments in order to maximize the phonetic difference in the priming landmarks. Both confederates aimed to remain constant in their pronunciation of the landmark items. As a manipulation check on whether they provided the expected acoustic cues in the priming condition, we conducted a one-way ANOVA (by-items) to compare the two confederates' vowel duration and closure voicing for /b/ landmarks and on the vowel duration of the first and second formants for the /æ/ landmarks during Rounds 5–8. As expected, the American and Korean confederates presented different English phonetic profiles in their landmarks (see Tables 2a and 2b). For the /b/ landmarks, the American confederate spontaneously produced significantly longer vowels (128 ms longer on average) and closure voicing (48 ms longer on average) than the Korean confederate. For the /æ/ landmarks, the American produced significantly longer vowels (77 ms longer) with much higher F1, by 1.7 Bark (260 Hz) and lower F2 by .7 Bark (214 Hz). In other words, her /æ/ vowels were longer, lower, and backer than the vowel /æ/ produced by the Korean partner. These differences between the two confederates persisted throughout the 4 rounds that each confederate did with the subjects, with no reliable Round × Partner interactions. (Because labels differed between rounds, there was variability caused by the vowel environment of the label for /b/ items, but this variability was present for both confederates, with no other systematic changes from one round to the next.)

Results

To look for effects of partner and priming, we computed 2 × 2 ANOVAs for Partner (American vs. Korean) × Priming (non-primed Baseline vs. Primed) for five measures: vowel duration and the first and second formants for the /æ/ items, and vowel duration and closure voicing duration for the /b/ items; comparisons of interest were all within-subjects. If Korean speakers considered their partner's language background (as a native speaker of English or of Korean) and adapted accordingly, then all else being equal, they should produce more English-like /b/ and /æ/ segments with the American partner than with the Korean partner (not only in the Priming condition, but also in Baseline). If such adaptation is driven by priming, Korean speakers should produce more English-like /b/ and /æ/ segments only when they have just heard those segments produced with English phonology (predicting a Priming × Partner interaction).

Priming

Evidence for priming was found from the vowel duration measurements of both /b/ and /æ/ items. Korean speakers produced significantly longer vowels immediately after they were primed by hearing the native-English-speaking partner produce the similar landmark terms for the /b/ items, Partner × Priming interaction, $F(1,19) = 21.18$, $p < .001$, $n_p^2 = .527$ and for the /æ/ items, Partner × Priming interaction, $F(1,19) = 6.74$, $p = .018$, $n_p^2 = .262$. For the /æ/ items, the first formant also showed a significant Partner × Priming interaction, $F(1,19) = 4.80$, $p = .041$, $n_p^2 = .202$. The other measures showed no reliable Partner × Priming interactions (for the /b/ items, closure voicing duration, $F(1,19) = 2.9$, n.s., and for the /æ/ items,

Table 2a
Acoustic measures (SDs) of the /b/ priming landmarks in ms from the two confederate partners' spontaneous speech, Experiment 1.

	Korean speaker of English as L2	American speaker of English as L1	Comparisons of confederates' speech
Vowel duration (*SD*)	85.1 (23.8)	212.6 (59.3)	$F(1,6) = 47.11$, $p < .001$, $n_p^2 = .887$
Closure voicing duration (*SD*)	12.0 (19.3)	59.6 (27.6)	$F(1,6) = 63.10$, $p < .001$, $n_p^2 = .913$

Table 2b
Confederates' spontaneous speech: acoustic measures (SDs) of the /æ/ priming landmarks (vowel duration in ms; formants in the Bark scale), Experiment 1.

	Korean speaker of English as L2	American speaker of English as L1	Comparisons of confederates' speech
Vowel duration (*SD*)	95.8 (19.7)	172.8 (38.3)	$F(1,7) = 271.70$, $p < .001$, $n_p^2 = .975$
F1 (*SD*)	6.7 (.4)	8.5 (.4)	$F(1,7) = 938.14$, $p < .001$, $n_p^2 = .993$
F2 (*SD*)	13.2 (.9)	12.5 (.6)	$F(1,7) = 30.43$, $p = .001$, $n_p^2 = .813$

second formants ($F(1,19) = .23$, n.s.). The means of vowel duration and other acoustic measures are presented in Tables 3a and 3b.

Partner's L1

Native Korean speakers produced more English-like forms of English /b/ and /æ/ words (more distinct from /p/ and /ɛ/ words in terms of vowel duration) after they had just heard a phonetically similar landmark term produced by the native-English-speaking partner (see the starred cells in Tables 3a and 3b). Although there was a main effect of partner's language background on vowel duration for /æ/ targets, $F(1,19) = 6.44$, $p = .020$, $n_p^2 = .253$), the means in Table 3b show that these are entirely due to the interaction. That is, those vowel durations addressed to the English-speaking confederate but not primed by that partner (Baseline) were just as short as those addressed to the native-Korean-speaking. The fact that Korean native speakers' Baseline productions addressed to the English-speaking confederate were not significantly different from the ones addressed to the Korean-speaking confederate suggests that phonetic adaptation was shaped *not* by the partner's status as a native or non-native speaker of English, but by the coda or the vowel just heard (primed). The only main effect of partner's language background that was not due to the interaction was found on closure voicing in /b/ items ($F(1,19) = 7.3$, $p = .014$, $n_p^2 = .278$). However, the difference went in the wrong direction and was due to a difference between the two Baseline conditions ($F(1,19) = 7.0$, $p = .016$, $n_p^2 = .269$), with closure voicing significantly shorter to the English-speaking confederate than to the Korean-speaking confederate. There was no reliable adaptation of F2 in /æ/ items, raising the question of why some but not all aspects of the L2 speaker's speech were shaped by the native English-speaking confederate's pronunciation of the landmark primes.

Table 3a
Mean vowel durations and closure voicing durations in ms (*SD*) for the subjects' /b/ targets in Experiment 1.

	Partner's L1	Baseline (not primed)	Primed
Vowel duration	Korean	177.0 (56.9)	170.4 (60.4)
	English	178.5 (63.0)	209.8 (65.4)[a]
Closure voicing duration	Korean	13.7 (25.0)	11.7 (17.9)
	English	7.6 (11.1)	10.7 (13.9)

[a] The reliably different cell in the Partner × Priming interaction for vowel duration.

Table 3b
Mean (*SD*) for vowel durations in ms and formants in normalized to the Bark scale for the subjects' /æ/ targets in Experiment 1.

	Partner's L1	Baseline (not primed)	Primed
Vowel duration	Korean	153.1 (38.2)	158.4 (40.5)
	English	155.6 (43.5)	177.0 (43.7)[a]
F1 (in Bark)	Korean	6.94 (.80)	6.93 (.82)
	English	6.96 (.98)	7.11 (.83)[a]
F2 (in Bark)	Korean	12.89 (.93)	12.79 (.82)
	English	12.97 (.84)	12.83 (1.03)

[a] The reliably different cells in the Partner × Priming interaction for vowel duration and F1.

Discussion

Korean speakers of English adapted by producing more English-like segments immediately after priming by similar segments from their native-English-speaking partner. They did *not* produce more English-like segments in either of the Baseline conditions (to either the American or the Korean partner). The fact that Koreans produced more English-like /b/s or /æ/s to the American partner *only* when they had just heard such segments suggests that simply knowing whether or not the addressee belongs to the same language community is not sufficient to drive phonetic adaptation (at least with the current measures). These results differ from those of Kim et al. (2011), who found that Korean and English speakers were more likely to converge with a partner they had more in common with (in language and/or dialect). The difference may be due in part to the strong accent of our Korean confederate, which may have reinforced subjects' desire either to speak less Korean-accented English or else to speak more like the native English-speaking confederate.

Although hearing standard English segments from the American English-speaking partner primed the Korean subjects to produce segments that were more English-like than they did in the absence of such priming, these were only approximations; vowel duration increased to native-like levels for both /b/ and /æ/, but closure voicing for /b/ and F2 for /æ/ did not (although F1 did increase in the direction of the English target). This pattern may have occurred because, for the /b/ items, vowel duration is more easily processed and manipulated by the foreign language speaker than are the other acoustic cues such as closure voicing. It has been shown that Mandarin learners of English, who have no voicing contrast in their language, produce final [b] with less closure voicing than English native speakers, having no significant difference from final [p] (Flege, McCutcheon, & Smith, 1987; Hayes-Harb, Smith, Bent, & Bradlow, 2008). In addition, Flege (1988) found that Mandarin speakers did produce longer vowel duration before final [b] than before [p]. These results taken together suggest that sustaining voicing during stop closure is generally more difficult than lengthening vowels for L2 learners. Thus it appears that the Korean subjects' target items, though primed by all available native cues, were produced as more English-like only in the cues that they were better at producing.

Note that Korean subjects produced a much longer vowel for both baseline /b/ and /æ/ (170.0 and 153.1 ms longer, respectively, Table 3) than did the confederate Korean partner (who intentionally produced Korean /p/ and /ɛ/-like vowel durations for English /b/ and /æ/, 85.1 and 95.8 ms respectively, Table 2). In other words, Korean subjects' baseline productions of the English target words were not nearly as Korean-accented as the productions of the Korean confederate. Although subjects *could* have produced segments that were more Korean-like after hearing the Korean confederate's heavily accented speech (indeed,

that is what an inflexible priming account would predict), they did not. So in this conversational setting, adaptation was not automatic in the sense of being obligatory, but flexible; the fact that it occurred *only* toward native and not toward Korean accented English indicates that the direction of priming was modulated by another factor. Given that the subjects recruited for the experiment were Korean L2 learners of English, they were likely aware that English has vowel and consonant contrasts that Korean does not, and probably had the goal of trying to become more proficient in the L2. The one-directional adaptation to the native rather than the non-native English speaker in the priming condition may have arisen from a motivation to produce more native-like English, with the desirable but difficult distinctions becoming easier to produce right after priming.

Experiment 1's design was constrained by the need to elicit an uncontaminated baseline in order to test a global version of the audience design hypothesis with respect to the language community membership of the partner. Although the results of Experiment 1 did not show partner-specific adaptation based simply on the partner's language background, they are consistent with (motivated) priming, particularly for vowel duration. However, the fact that the Baseline condition (Rounds 1–4) preceded the Priming condition (Rounds 5–8) for each confederate partner presents a potential confound that leaves open the possibility that over time, speakers may have become more sensitive to the contrast and therefore more likely to produce it (although that would not explain the difference between Baseline and Priming conditions in speech addressed to the American confederate). In Experiment 2 we replicated the priming effect without this condition-order confound by unblocking Baseline from Primed items and systematically counterbalancing their order. In addition Experiment 2 tested another kind of critical partner-specific effect: one based on the pragmatic needs of the partner rather than on the partner's identity as a member of a language community.

Experiment 2

Experiment 1 did not show an effect of affiliation based on whether the addressee was from the same or different language community. In Experiment 2 we tested a finer-grained partner-specific factor, that of the addressee's pragmatic needs. Note that Experiment 1's task did not require speakers and addressees to distinguish /b/ and /æ/ from /p/ and /ɛ/; that is, even though these sounds are not contrastive in Korean phonology, the task contained no referential ambiguity that needed to be resolved. There were no cards that contained the coda /p/ and the vowel /ɛ/ on the board, so even if the native-Korean speaker said '*Hop*' for '*Hob*', the partner could still determine that the intended referent was '*Hob*', and there were no cases where forms like '*Bat*' had to be distinguished from a minimally different form like '*Bet*'. So it is unclear whether the Korean speakers simply did not notice an opportunity for making the /b/–/p/ and /æ/–/ɛ/ contrasts (particularly in the Baseline condition), or whether they failed to produce these contrasts because they were not pragmatically necessary to succeed in the matching task.

In Experiment 2, native-Korean speakers of English as an L2 did a modified version of the card-matching task that required them to make distinctions between /b/–/p/ and /æ/–/ɛ/ in order to communicate unambiguously with the native-English-speaking partner. This provided a strong test of an audience design hypothesis based on the pragmatic needs of a communicative partner.

In addition, to detect whether partner-specific adaptation always necessitates extra planning or whether it can be done flexibly and swiftly, we measured speakers' time to initiate speaking. Note that proposals such as Horton and Keysar's (1996) monitoring-and-adjustment model and Pickering and Garrod's (2004) alignment model posit a modular architecture for speech planning in which initial planning is inflexible and egocentric. On this view, any partner-specific adaptations that take place would do so as a kind of repair or afterthought, requiring inferences about the partner that take additional time. Alternatively, audience design could guide planning early in speaking, as long as speakers are aware of addressees' pragmatic needs.

Predictions

As in Experiment 1, we predicted that items would elicit more English-like forms when primed by the English native speaker than when not primed. Moreover, if phonetic adaptation is flexible enough to be shaped by the needs of the addressee, then Korean speakers of English should use more English-like forms when needed to contrast two similar items ('*Sob*' vs. '*Sop*' or '*Bat*' vs. '*Bet*') that were adjacent on the board. To the extent that audience design requires additional inferences, it should take more time for speakers/directors to initiate their answers in the pragmatic contrast condition than in the Priming or Baseline conditions. If initiating a pragmatic contrast *always* takes longer, that would be evidence for egocentric or inflexible processing. However, if, once the Korean speakers of English became aware of the potential for phonetic ambiguity, they adapt to their matchers' needs by initiating pragmatically needed phonetic contrasts without delay, that would be evidence for flexible, "smart" audience design.

Method and design

Experiment 2 used the same matching task as in Experiment 1, but had non-native speakers interact with a single native-English-speaking confederate. To replicate Experiment 1's priming effect, some coda /b/ and vowel /æ/ landmarks were phonetically similar to the targets as in Experiment 1, with the modification that the priming condition and baseline condition blocks were counterbalanced to determine whether evidence for priming would be found in early rounds of the experiment as well. In addition, to determine whether non-native speakers are able to flexibly adapt their utterances in ways consistent with their partners' communicative needs, some targets were situated in a context in which a minimal pair of cards

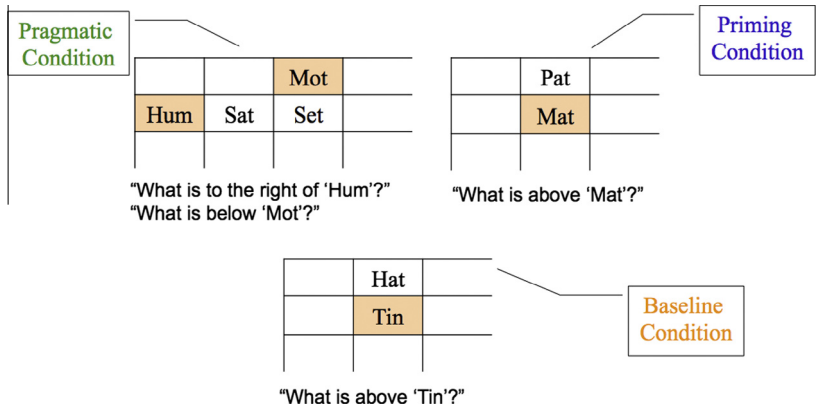

Fig. 4. Examples from boards illustrating the three conditions and the confederate matcher's utterances in Experiment 2.

Table 4
Target items from Lists A and B in Experiment 2. Baseline items are in white, primed items are in gray, and pairs of pragmatically contrasting items are in light gray. The pragmatically contrasting items within a pair were referred to one right after another, counterbalanced for whether the one with the phonetic segment missing from L1 was referred to first or second.

		Block 1 (Rounds 1–4)				Block 2 (Rounds 5–8)			
Round		1	2	3	4	5	6	7	8
List A	/b/–/p/	Nib	Dob	Nob	Sib	Mob Mop	Dib Dip	Bib Bip	Sob Sop
	/æ/–/ɛ/	Pat Pet	Hat Het	Tap Tep	Sack Seck	Hack	Bat	Sat	Sap
List B	/b/–/p/	Nib Nip	Dob Dop	Nob Nop	Sib Sip	Mob	Dib	Bib	Sob
	/æ/–/ɛ/	Pat	Hat	Tap	Sack	Hack Heck	Bat Bet	Sat Set	Sap Sep

containing both members of the L2 contrast (e.g., *Ni**b*** and *Ni**p***) were adjacent on the board, and where the need for a contrast was prompted by the difference in English orthography in the cards' labels (spelled with 'b' vs. 'p' or 'a' vs. 'e'). We reasoned that when a segment that does not exist in L1 was necessary for a pragmatic distinction in L2, the Korean speakers of English would try to make this distinction for their partner, even if it required approximating the distinction in a non-native-like fashion. Finally, as in Experiment 1, some landmarks were phonetically different from the targets in order to provide a baseline.

This design enabled us to compare the same types of labels (e.g., *Nib*, *Sib*, *Dib*, etc.) in three contexts; one where there is a priming landmark providing phonetic cues to coda voicing or to the vowel /æ/ (priming condition), one where the situation encourages subjects to make the contrast of '*Nib*' and '*Nip*' (pragmatic contrast condition) and one where there is a phonetically irrelevant priming landmark, as illustrated in Fig. 4 (baseline condition). As shown in Fig. 4, the pair '*Sa**t***' and '*Se**t***' in the pragmatic contrast condition were adjacent to each other on the board, indirectly cueing subjects about the communicative need for contrasts. The word '*Pat*' illustrates the priming condition, primed by the question 'what is above *Mat*?' spoken by the confederate. '*Hat*' is a baseline word, as it neither occurred with a contrasting form, nor was it phonetically primed.

Materials and procedures

Each 8-by-8-cell board held one set of 10 target cards (3 critical cards and 7 filler cards), as well as 10 landmarks, for a total of 24 critical labels. Table 4 shows the critical labels (e.g., the three in Round 1 were *Nib*, *Pat*, and *Pet*). The critical labels were the same kind (containing coda /b/ and the vowel /æ/) as in Experiment 1. Eight different sets of critical and filler targets and landmarks were generated, for a total of 8 rounds. There were two lists of stimuli, with each subject matching the cards from one list (see Table 4). For landmarks and fillers, see Appendix B. As Table 4 shows, each block contained baseline, priming and pragmatic items; List was the only between-subject factor and all the others were within-subject.

As in Experiment 1, Korean-speaking subjects were assigned to the director role; the same English native speaker as in Experiment 1 played the matcher. The matcher and director took part in a short practice round in which they matched a set of 6 cards, with the matcher asking questions about target cards vis-a-vis landmark cards (as in Experiment 1). The matcher could ask about the target card locations in any order, except that she

always asked about the cards with pragmatically contrasting labels one right after the other, in a predetermined counterbalanced order with /æ/ or /b/ first half the time and with /ɛ/ or /p/ first the rest of the time. After each round (one practice round and eight experimental rounds), the pair checked to see whether their cards were correctly matched.

At the end of the experiment we had the Korean subjects classify the target labels as words or nonwords of English (without having to give the meanings of the words, and with encouragement to guess). Of the labels that were words, 87.8% were correctly classified as words; of the nonwords, only 44% were correctly classified as nonwords.

Participants

Subjects

Twenty-two graduate and undergraduate students (7 male and 15 female) from Stony Brook University volunteered for this study and received $10 or research credit in a psychology course for their participation. All were 18 years of age or older and native speakers of Korean; all provided informed consent. None had participated in Experiment 1.

After the experimental session, the confederate rated the accentedness of each subject's speech from 0 (no accent) to 5 (heaviest accent; see Fig. 5) and the subject filled out a language background questionnaire. Accentedness was correlated (r (19) = .49, p = .030) with the age at which subjects began acquiring English (the only subject who began learning at home rather than at school had the lowest accentedness rating, 0.5). The questionnaire and subjects' responses are available online at http://www.psychology.sunysb.edu/sbrennan-/korean_subjects_JML_e2/ (see Supplementary material).

Confederate

The same native speaker of English who served as the American confederate in Experiment 2 served as the lone confederate in Experiment 1. She was not acquainted with any of the subjects.

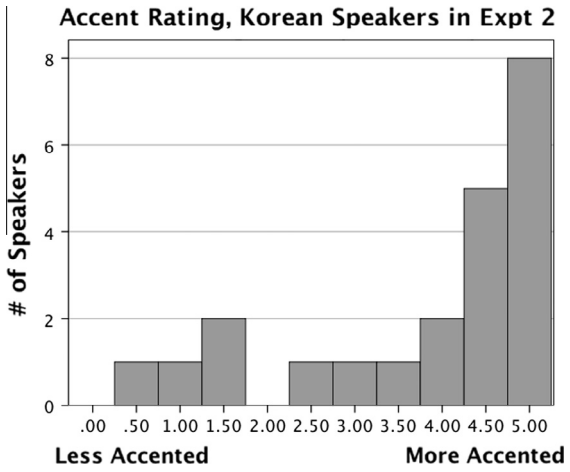

Fig. 5. Accent ratings for Korean speakers of English (naive subjects) in Experiment 2.

Coding

Acoustic measures of adaptation

The same acoustic measures were made as in Experiment 1; vowel duration, and closure voicing for the /b/ and /p/ items, and vowel duration, F1, and F2 for the /æ/ and /ɛ/ items. In addition to these, stop release duration was measured for the /b/ and /p/ items on the expectation that Koreans would make use of release characteristics to signal stop contrasts (note that stop release duration was not measured in Experiment 1 as there were no /p/ segments in the critical stop targets). The majority of cues to Korean stop (onset) contrasts are phonetically manifested near the moment of release, and it has been shown that Koreans use English releases in stop detection (Cho & McQueen, 2006), so we would expect that Koreans would be sensitive to the role of stop bursts in disambiguating stops; Hayes-Harb et al. (2008) also report that Mandarin speakers (who also have no stop voicing contrast in codas) use stop release duration to differentiate voiced and voiceless English stops in production. Stop release duration was computed as the period from the beginning of the burst transient to the point where the noise abruptly decreased in intensity in most frequency ranges in spectrogram; we found that 107 out of 264 items were released (6 of the subjects did not release /p/, /b/ or both). For the vowel targets /æ/ and /ɛ/, measures of formants were obtained automatically by a Praat script and were double-checked, with outliers re-coded by hand, following the same method as in Experiment 1.

Timing of speech onsets

For the analysis of latency to speaking, we measured the interval duration from the end of the confederate/matcher's question to the onset of the Korean speaker/director's answer to that question for items in all three conditions (baseline, priming, and pragmatic contrast).

Results

When the Korean speakers conversed in English, both priming and the need for pragmatic contrast shaped how they produced the targets missing from Korean (see Figs. 6a and 6b).

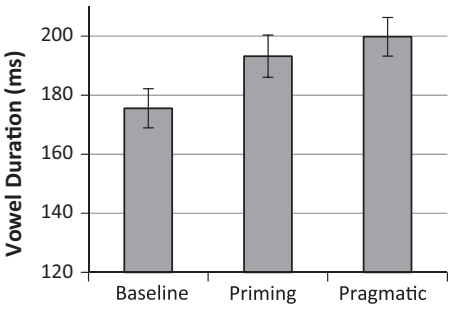

Fig. 6a. /b/ items, Experiment 2. Vowel duration (with *SE* bars) when phonetic segments are not primed (Baseline), primed (Priming condition), or in need of disambiguation (Pragmatic condition).

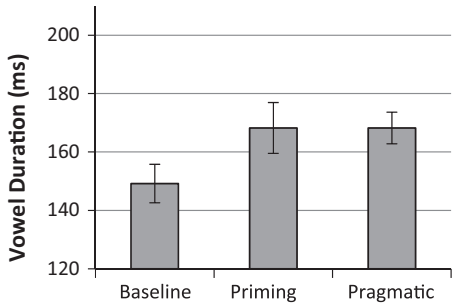

Fig. 6b. /æ/ items, Experiment 2. Vowel duration (with *SE* bars) when phonetic segments are not primed (Baseline), primed (Priming condition), or in need of disambiguation (Pragmatic condition).

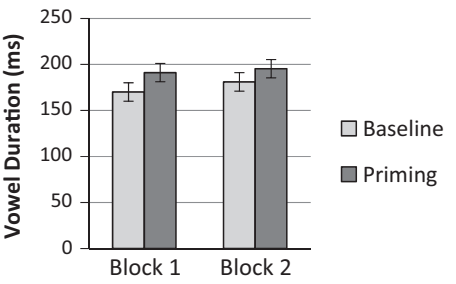

Fig. 7a. /b/ items. Within-speaker priming as measured by vowel duration, Experiment 2.

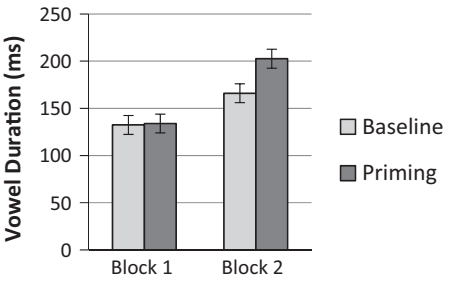

Fig. 7b. /æ/ items. Within-speaker priming as measured by vowel duration, Experiment 2.

Priming

To look for effects of priming, we computed ANOVAs (by-subjects) on vowel duration and closure voicing duration for the /b/ items, as well as vowel duration, F1, and F2 for the /æ/ items, with priming as a factor (comparing Baseline to Primed conditions). Recall that in Experiment 1, these conditions were confounded with the order in which subjects experienced them (baselines were collected in Rounds 1–4, with primed items collected in Rounds 5–8). There, it was possible that the priming effect may have been due in part to speakers' learning to produce the phonetic segments /æ/ and coda /b/ from hearing the native-English-speaking confederate over time, rather than due (only) to local priming from the confederate's similar landmarks (heard immediately before producing the target items). Experiment 2 removed this confound by unblocking Primed from Baseline items (so that both appeared in on the same board, within the same block), by counterbalancing the order of the blocks (as shown in Table 4, with each speaker experiencing the /æ/ items in one block and the /b/ items in the other block), and by including order (priming block first vs. priming block second) as a between-subjects factor in the ANOVAs.

The priming effect found in Experiment 1 was replicated; immediately after hearing phonetically similar landmarks, the Korean-speaking subjects produced /b/s and /æ/s that were more English-like than after hearing phonetically irrelevant landmarks. The vowel durations associated with both sounds were significantly longer when primed than when not (for /b/, 19 ms longer, $F(1,20) = 9.50$, $p = .006$, $n_p^2 = .322$; for /æ/, nearly 18 ms longer, $F(1,20) = 12.29$, $p = .002$, $n_p^2 = .381$; see Figs. 6a and 6b. F1 for /æ/ tokens was significantly higher when primed than when not ($M = 7.18$ for Baseline, $M = 7.35$ for Primed, $F(1,20) = 4.42$, $p = .048$ $n_p^2 = .181$), replicating the result in Experiment 1. Consistent with Experiment 1, closure voicing in /b/ items showed no reliable priming ($M = 23.3$ for Baseline, $M = 26.2$ for Primed, $F(1,20) = .51$, $p =$ n.s.), nor did F2 ($M = 1932.1$ for Baseline, $M = 1913.9$ for Primed, $F(1,20) = 3.89$, $p = .063$) in /æ/ items).

Order effects for priming. Next we looked for order effects (which may have been in play in Experiment 1); see Figs. 7a and 7b. There was no effect of block order in the vowel duration effect for the /b/ items (Priming × Order interaction: $F(1,20) = 2.22$, n.s.), nor for F1 for the /æ/ items (Priming × Order interaction: $F(1,20) = 3.96$, $p = .060$). This is consistent with Experiment 1's findings of a priming effect. However, there *was* an order effect for /æ/ items for vowel duration (Priming × Order interaction: $F(1,20) = 10.43$, $p = .004$, $n_p^2 = .343$); Fig. 7b shows that there was no priming effect for the vowel duration measure in Block 1 (Rounds 1–4), but a strong one in Block 2 (Rounds 5–8), where the Primed condition shows a longer vowel than the Baseline condition. That the Baseline condition also shows a longer vowel in the second block than in the first block suggests learning for vowel contrasts, as /æ/ vowels become longer and thus more native-English-like in the later rounds. And that Block 2's Primed condition shows a longer vowel than Block 2's Baseline condition suggests that local priming leads to an even more native-English-like vowel (amplifying the effect of learning).

Pragmatic contrast

The key question for Experiment 2 was whether the Korean-speaking subjects would produce the critical consonant (/b/) and vowel (/æ/) differently from /p/ and /ɛ/ respectively in the Pragmatic Contrast condition, as well as from /b/ and /æ/ in the baseline condition. We conducted ANOVAs comparing /b/ and /æ/ in the Pragmatic condition to Baseline for the same kinds of acoustic measures used in Experiment 1, and in addition, compared those measures for /b/ vs. /p/ and /æ/ vs. /ɛ/, as these were directly comparable within paired (temporally and spatially adjacent) items. We also added release duration with voicing (voiced or voiceless) as a factor to the comparison of /b/ vs. /p/ paired items.

Table 5a
Means (*SD*) of acoustic measures of /b/ and /p/ items (in ms), Pragmatic condition, Experiment 2. (Note that release duration can be compared here because the Pragmatic condition included both /b/ items and /p/ items.)

	/b/	/p/	Within-speaker /b/ vs. /p/ comparisons
Vowel duration	199.8 (61.4)	128.4 (38.4)	$F(1,21) = 121.09$, $p < .001$, $n_p^2 = .852$
Closure voicing duration	24.6 (17.6)	10.7 (7.0)	$F(1,21) = 25.56$, $p < .001$, $n_p^2 = .549$
Release duration	63.7 (33.2)	96.5 (50.0)	$F(1,15) = 14.12$, $p = .002$, $n_p^2 = .485$

Table 5b
Means (*SD*) of acoustic measures of /æ/ and /ɛ/ items (vowel duration in ms; formants in the Bark scale), Pragmatic condition, Experiment 2.

	/æ/	/ɛ/	Within-speaker /æ/ vs. /ɛ/ comparisons
Vowel duration	168.2 (50.8)	127.2 (43.7)	$F(1,21) = 28.76$, $p < .001$, $n_p^2 = .578$
F1 (in Bark)	7.30 (1.11)	6.75 (.98)	$F(1,21) = 23.26$, $p < .001$, $n_p^2 = .526$
F2 (in Bark)	12.88 (.88)	12.91 (.80)	$F(1,21) = .78$, n.s.

Pragmatic contrast within paired items

When two contrastively labeled items were adjacent to each other on the board, Korean-speaking subjects indeed produced them differently, even though their first language lacked the target contrasts. The vowel and closure voicing of the /b/ items were significantly longer than for the /p/ items and the release duration of the /p/ items was significantly longer than the /b/ items (Table 5a); vowel duration was significantly longer for the /æ/ items than /ɛ/ items (Table 5b). As for spectral cues, only the first formant differed between the vowel pairs, showing the expected higher first formant for /æ/. These clear differences suggest that the adjacency of phonetically similar labels that needed to be distinguished may have made Korean-speaking subjects particularly aware of the potential for ambiguity.

Pragmatic contrast compared to baseline

Not only were /b/s and /æ/s pronounced differently from /p/s and /ɛ/s, respectively, but they were pronounced in a more English-like manner when they needed to be contrasted with sounds from L2 than when they did not, as predicted by the audience design hypothesis. Korean speakers of English produced significantly longer vowels for both /b/ and /æ/ items in contexts with potentially ambiguous competitors than in the Baseline condition (with no similar competitors). That is, Korean speakers produced more English-like /b/s and /æ/s when the English-speaking partner needed unambiguous input in order to do the collaborative task. That speakers make such distinctions when needed by addressees (even when the distinctions are not facilitated by the speakers' L1) supports the audience design hypothesis.

Matching order

Speakers can engage in audience design only when their interlocutors' needs are known (Horton & Gerrig, 2002; Kraljic & Brennan, 2005). This raises the interesting question of what happens when one turn in the matching task implicitly removes the need to disambiguate a referring expression in the next turn. In other words, the order of matching competing items in the pragmatic condition could matter, because once the matcher has correctly resolved the referent of the first expression, the referent of the second expression is no longer ambiguous (it no longer has a competitor in the matching task). Therefore, we included two contrasts in the ANOVA: the first compared both orders (combined) to the Baseline condition, and the second compared (within the Pragmatic condition) the order in which the /b/ and /æ/ items were produced. Because this involved two comparisons rather than one, we adopted a Bonferroni correction of $p < .025$ (see Table 6).

We found mixed evidence about the degree to which audience design is precise. The means for vowel duration and closure voicing duration of the stop consonant /b/ items were not reliably different, whether they were matched before the /p/ items or afterward (although as Table 6a shows, the means were in the right direction, with greater ambiguity avoidance with competitors present than without). However, the /æ/ items did show an order difference, with longer vowel duration for /æ/ items when the /ɛ/ items were still competitors than when they were not (Table 6b).

Latency to onset of speaking

To assess whether the processes underlying any of these effects required additional planning that could delay speaking, we compared the latencies to speech onset for utterances in the Pragmatic and Priming conditions to those in the baseline condition, using one-way ANOVAs. Critically, in the pragmatic contrast condition, Korean subjects needed to answer the matcher's two subsequent questions with two contrasting words in successive turns (this means that although they were not aware that the matcher's next question would be about the second item, they always saw the contrasting pair of labels side-by-side before they answered the first question and so were aware of the potential for ambiguity). We predicted that it would take them longer to initiate speaking when answering the first question (presumably noticing the ambiguity and planning to avoid it), than to initiate speaking in the second one (when they were already aware of the ambiguity). In the latter case, as long as speakers are aware of an addressee's needs, designing an unambiguous utterance should not take longer unless audience design is

Table 6a
Pragmatic contrast effects for /b/ items, Experiment 2. Means (*SD*) of the acoustic measures are in ms. For each measure, the overall comparison to Baseline is followed by the comparison of whether the /b/ item is matched before or after the /p/ item (with Bonferroni correction of $p < .025$).

	Baseline	Pragmatic condition		Comparisons
Vowel duration	175.6 (44.1)	199.8 (61.4)		$F(1,21) = 7.82$, $p = .011$, $n_p^2 = .271$
		Matched first	Matched second	
		202.3 (66.4)	197.4 (56.5)	$F(1,21) = .20$, n.s.
Closure voicing duration	23.3 (22.1)	24.6 (17.6)		$F(1,21) = .25$, n.s.
		Matched First	Matched Second	
		24.4 (16.6)	24.8 (18.7)	$F(1,21) = .02$, n.s.

Table 6b
Pragmatic contrast effects for /æ/ items, Experiment 2. Means (*SD*) of the acoustic measures include vowel duration in ms and formants in the Bark scale. For each measure, the overall comparison to Baseline is followed by the comparison of whether the /æ/ item is matched before or after the /ɛ/ item (with Bonferroni correction of $p < .025$).

	Baseline	Pragmatic condition		
Vowel duration	149.2 (43.8)	168.2 (50.8)		$F(1,21) = 7.69$, $p = .01$, $n_p^2 = .268$
		Matched first	Matched second	
		179.9 (50.4)	156.6 (49.1)	$F(1,21) = 7.03$, $p = .015$, $n_p^2 = .251$
F1 (in Bark)	7.18 (1.01)	7.31 (1.11)		$F(1,21) = 2.75$, $p < .20$, n.s.
		Matched first	Matched second	
		7.21 (1.09)	7.40 (1.14)	$F(1,21) = 5.03$, $p = .036$, n.s.
F2 (in Bark)	12.93 (.78)	12.88 (.88)		$F(1,21) = .27$, n.s.
		Matched first	Matched second	
		12.97 (.78)	12.78 (.97)	$F(1,21) = .21$, n.s.

Table 7
Mean latency to speaking, with Priming, First Pragmatic, and Second Pragmatic conditions compared to Baseline condition, Experiment 2. Data have been log-transformed for comparisons; a Bonferroni correction of $p < .017$ is used.

Baseline	Priming	Pragmatic contrast	
		First item	Second item
1006.8 ms (994.2)	1026.8 ms (801.3)	1438.2 ms (1185.3)	966.5 ms (1287.8)
Comparisons to baseline	$F(1,21) = .328$, n.s.	$F(1,21) = 14.34$, $p = .001$, $n_p^2 = .406$	$F(1,21) = 4.87$, $p = .039$, n.s.

constrained to follow an egocentric, modular stage; therefore we expected a faster speech onset to the second contrasting item than the first one for two reasons: speakers would be more likely to have noticed the potential for ambiguity before pronouncing the second one, and they would not have to locate its landmark since it would be close to the previous one.

Because the response times were positively skewed, we normalized them using a log transformation; because the logic of this measure involved making three comparisons to the Baseline condition rather than one, we used a Bonferroni adjustment of $p < .017$ for the significance values. Results are in Table 7.

The latency to speaking was indeed longer than baseline (by 431 ms) when speakers referred to the first item in the pragmatic contrast, compared to the second item (which was not reliably different from baseline). It appears, then, to take measurable time to notice the potential for ambiguity and plan an utterance tailored to a partner's needs, but not to tailor the utterance once this initial adaptation has been made. That the second contrasting item was so fast to produce is also affected in part by the fact that the subjects did not need to spend much time locating the landmark, since it was close to the landmark for the previous item; however, this does not present a confound, since this search time is also required in the Baseline (and Priming) conditions. There was no difference in speech latency between Priming and Baseline conditions, which is consistent with the idea that priming influences phonetic adaptation automatically, implicitly, and outside of speakers' awareness.

Discussion

Experiment 2 provides several types of evidence for audience design in speech production. When the Korean speakers of English needed to pronounce a segment contrastively in their L2 in order to be understood by their partners, they did so, even though this contrast was not present in their L1. This adaptation resulted in pronunciations that were more similar to English than were the baseline utterances. These adaptations were made locally (in that they were tailored to the potential for ambiguity in the specific referential context), rather than globally (they were not present in the comparisons of the Baseline condition for speech addressed to the two confederate partners in Experiment 1, where pragmatic ambiguity was absent).

When speakers were directly contrasting /b/ and /p/, they manipulated release duration as well as vowel duration. The release duration of /p/ in the pragmatic contrast condition was reliably longer than that of /b/ (by nearly 33 ms), even though post-vocalic /p/ in Korean is strictly unreleased. In other words, the Korean speakers of English not only approximated the /b/s missing from their L1 by lengthening preceding vowels, but also adopted a strategy not available in their L1, producing (normally unreleased) coda /p/ with an unusually long release burst (enhancing the L2 contrast between /p/ and /b/).

The fact that speakers made distinctions in the Pragmatic condition, but did so just as strongly for consonants in the first-mentioned competitor as in the second-mentioned competitor, suggests that dynamic, local adaptation to an audience's needs can be done efficiently, without having to first analyze the figure-ground relationship of a referent for ambiguity in the referential context. This finding is consistent with Brown & Dell (1987; Dell & Brown, 1991) and Kraljic and Brennan (2005), who found that speakers disambiguate when there is a likelihood (as opposed to an actuality) of ambiguity. That audience design seemed to be more precise for the vowel contrast is therefore surprising. In this specific case, it could be that implementing the goal of vowel disambiguation is particularly difficult for our subjects, so they make a special effort to do so only when it is absolutely necessary (the first mention).

Converging evidence for pragmatic influence on adaptation comes from the finding of a greater latency to speech onset before the first reference within a pair of potentially ambiguous references. This result suggests that adapting to an addressee's need for information takes time, but only when the potential for ambiguity has not yet been noticed by the speaker. When speakers were already aware of addressees' needs (by the second reference, where the need for a pragmatic contrast was evident), they were just as fast to initiate speaking as they were in the baseline (non-ambiguous) condition. This constitutes evidence that even rapid adaptation in pronunciation can be "smart"—that is, a flexible process that is sensitive to partner-specific information (as proposed in Galati & Brennan, 2010; Stent, Huffman, & Brennan, 2008); it need not unfold as a repair following a reflexive, encapsulated, "egocentric" process.

Could this "pragmatic effect" instead be a lexical effect, due to neighborhood density? Baese-Berk and Goldrick (2009) report that having more similar-sounding competitors in the lexicon can result in words being produced more distinctively. They found that there was greater voice onset time (VOT) for voiceless stops in items with a minimal competitor (e.g., *cot* with *got*, etc.) than with no competitor (*pep* with **bep*), and even greater VOT when the competitor was displayed on the screen. Baese-Berk and Goldrick argue that this is due to hyperarticulation caused by higher activation. There are several reasons why neighborhood density is not a satisfactory explanation for our pragmatic effect. First, work by Munson and Solomon (2004) has shown that while neighborhood density does affect vowel formants (greater density means more vowel space dispersion), it does not affect vowel duration; in our pragmatic-to-baseline comparison (Table 6b), speakers manipulated only vowel duration. Furthermore, it is not clear how neighborhood density would explain the fact that first mentions for /æ/ items are produced with greater vowel durations but not higher F1 (hyperarticulation would influence both phonetic properties, in the absence of a principled reason to predict otherwise). On a neighborhood density explanation, whether our first-mentioned items were produced more clearly should be determined by how many lexical neighbors those words had, and would presumably not be affected by the difficulty of conveying the particular contrast involved. Yet, we found a first mention effect for the /æ/–/ɛ/ contrast and not for the /b/–/p/ contrast. Finally, it should be noted that our subjects were learners of English, whose English lexicons were relatively small and idiosyncratic (they were not particularly accurate at distinguishing words from nonwords in the rating task). Without measuring these individuals' lexicons, it is impossible to assess any impact of neighborhood. That we found the pragmatic effect despite this potential source of noise suggests that considerations beyond lexical neighborhoods are involved. Consistent with Experiment 1, Experiment 2's findings suggest that priming is not an inflexible, obligatory process. We found priming effects for /b/ on vowel duration and /æ/ items on both vowel duration and first formant, replicating the results of Experiment 1. For items with coda /b/, priming effects were found in both early and late blocks of the experiment, while for items with the vowel /æ/, priming on vowel duration was found only in the later blocks, although priming on the first formant was consistently found in both blocks. If priming were inflexible and obligatory, this variability would not be expected. We suspect that this difference is due in part to the fact that Korean speakers are generally more facile with the vowel duration differences associated with /b/ than with the vowel differences that distinguish /ɛ/ and /æ/ (see, e.g., Chang & Idsardi, 2001; Flege et al., 1997; Ingram & Park, 1997; Oh, 2011). These properties may be differentially represented and thus differentially primed. Given Korean speakers' reportedly weak ability to produce and perceive temporal differences between these vowels, they may have needed more exposure to the native speaker producing the /æ/ vs. /ɛ/ vowel length differences before they would be able to produce them reliably themselves. This may, then, have been an effect similar to that reported by Babel (2010), who proposes an activation threshold for imitation effects, since her (female) subjects showed cumulative F1 imitation effects over multiple exposures to the same voice. However, in our study it was not just increased exposure to the partner that may have produced the order effect for priming of vowel duration for /æ/. The subjects who did the priming condition in Block 2 experienced the pragmatic condition in the earlier rounds, suggesting that the awareness of contrast produced by the pragmatic condition task helped set the stage for priming in the later rounds.

General discussion

We found that Korean non-native speakers adapted their pronunciation to become more native-English-like in two situations: (1) when there was potential for ambiguity from similar adjacent competitors and (2) when the target label was primed by a phonetically similar item in the native-English-speaking partner's immediately preceding utterance. Pronunciation was not shaped by affiliation (as represented by the partner's language background in Experiment 1), but by the American confederate's immediately preceding speech (Experiments 1 and 2) and by a pragmatic need for disambiguation (Experiment 2). Our data on adaptation in non-native speech have revealed a number of subtle aspects of how priming and audience design work.

Although we found the predicted priming effects in both experiments, note that priming was not observed across the board. Priming occurred *only* when speakers spoke with the native English speaking partner; vowel durations for the L2 sounds missing from L1 never adapted in the direction of the much shorter vowels of the heavily accented Korean-speaking partner in Experiment 1 (as shown by the patterns of vowel durations in Tables 3 and 4). This suggests that priming in spoken dialogue is not automatic in the sense of being obligatory, but flexible, likely fueled in part by a motivation to speak more native-like English (see also Costa et al., 2008).

Turning to pragmatic effects, when speakers were prompted by orthographically contrasting labels that were visually adjacent (and that warranted pronouncing /b/ and /æ/ contrastively with /p/ and /ɛ/, respectively), Korean speakers of English produced more English-like /b/ and /æ/ as needed to be unambiguous. We also found that Korean speakers of English took more time to initiate labeling the first card in a pragmatic contrast when they had to first notice and plan for their partners' need for disambiguation. This planning cost for audience design did not need to be paid repeatedly, however; once speakers were aware of ambiguity, they were just as fast (and in fact, numerically if not significantly even faster) to initiate speaking in the second card in the contrast as they were in the Baseline condition. This result is consistent with the logic that a speaker cannot adapt to an addressee's pragmatic needs unless the speaker is aware of those needs (e.g., Horton & Gerrig, 2002; Kraljic & Brennan, 2005), but not with the proposal that speakers default to being egocentric for reasons of modularity, or that partner-specific adaptations are essentially repairs. Our findings illustrate how processes such as priming and audience design can work together to shape speakers' adaptation to addressees.

Our data also support the notion that disambiguating cues are produced by speakers not only on an as-needed basis, but rather when the *potential* for ambiguity is salient, even if a particular utterance does not turn out to be ambiguous once it unfolds in a particular context. For example, in Kraljic and Brennan's (2005) Experiment 2, speakers began to speak before analyzing whether there was actually ambiguity between similar, potential competitors; however, they provided prosodically disambiguating cues regardless of whether the competitors made the situation truly ambiguous. Similarly, in Stent et al. (2008), speakers continued to produced hyperarticulated speech for several turns after a specific error had been repaired and communication with the partner was proceeding without problems. If audience design were done only on an as-needed basis, this would require monitoring and analyzing the referential context continuously for ambiguity, which would slow the initiation of speaking. Our speakers did not appear to monitor precisely, but made stronger /b/ contrasts when they perceived a potential for pragmatic ambiguity (even when it had just been resolved by eliminating the competitor in the previous turn). On the other hand, that does not mean that speakers *never* do such monitoring; they did make stronger /æ/ contrasts when the /æ/ item still had an /ɛ/ competitor than when it did not. To the extent that the vowel contrast may be more difficult than the consonant one, L2 speakers may monitor more precisely for a contrast they know they are more likely to "get wrong", and they may be more likely to manipulate cues in accordance with their mastery of those cues as learners. We propose that examining adaptation in conversations between second language learners and native speakers can yield insights into the state of a learner's knowledge of the language being acquired.

In closing, although partner-specific adaptation in spoken dialog is limited by speakers' linguistic repertoire, ultimately it can serve the needs of communication. When our Korean speakers of English as a second language produced a contrast in L2 that was not present in L1, this was facilitated by immediate priming *and* by the need for pragmatic disambiguation. It was *not* affected by affiliation with the addressee as a native speaker of Korean. These findings suggest nuanced roles for both priming and audience design as forces behind phonetic adaptation. That the priming we found occurred only toward more English-like speech (and not toward a stronger Korean accent for vowel durations) is evidence that priming can be flexible. And that pragmatic contrasts were produced more clearly and without delay at moments when the potential for ambiguity was perceptually evident demonstrates that a speaker can design utterances for a partner non-egocentrically, without having to elaborately model the partner's needs.

Acknowledgments

This material is based upon work supported by NSF under Grant ITR-0325188. We thank our colleagues from the Adaptive Spoken Dialogue Project, the Dialogue Matters Network (funded by Leverhulme Trust), the Gesture Focus Group, and Arthur Samuel and Nancy Franklin for helpful discussions. We are especially grateful to Elizabeth Cohen Fanning for her assistance in running the

experiments and to our English-speaking Korean volunteers.

Appendix A

Appendix A.1. Stimuli from List A in Experiment 1. Items in white appeared in the baseline condition, and those in gray were in the priming condition.

List A	Target		Fillers		
Landmark	pix	fox	buy	dry	doom
1	nib	pat	dot	try	boom
Landmark	joy	ton	day	rem	fit
2	dob	hat	seat	rep	fin
Landmark	vox	hum	but	clow	sate
3	sib	nap	cook	slow	bate
Landmark	bey	kin	wit	rick	zake
4	job	dack	name	rit	cake
Landmark	hob	cap	eye	zin	yon
5	gob	bap	moon	by	beat
Landmark	kib	pack	wis	age	tux
6	jib	sack	sun	neck	seak
Landmark	kob	cat	hin	nee	cot
7	bob	mat	toy	came	heak
Landmark	tib	zap	nut	in	seat
8	mib	tap	net	boy	mum

Appendix A.2. Stimuli from List B in Experiment 1. Items in white appeared in the baseline condition, and those in gray were in the priming condition.

List B	Target		Fillers		
1	dib	hack	sot	rip	sane
Landmark	box	win	pox	slot	pot
2	mob	sat	sum	plot	hot
Landmark	coz	mot	dow	ret	take
3	zib	nack	dook	reck	tape
Landmark	fix	hut	bog	tream	sing
4	sob	dap	dame	dream	king
Landmark	kob	lap	son	one	non
5	nob	sap	boon	fame	heat
Landmark	pib	tack	tea	con	foy
6	gib	gack	pun	nick	peat
Landmark	tob	gap	pea	wiz	cow
7	zob	map	bate	goy	seek
Landmark	lib	zat	pey	six	cox
8	bib	bat	nit	soy	teak

Appendix B

Appendix B.1. Stimuli from List A in Experiment 2. Baseline items are in white, primed items are in gray, and pairs of pragmatically contrasting items are in dark gray. The pragmatically contrasting items within a pair were referred to one right after another, counterbalanced for whether the one with the phonetic segment missing from L1 was referred to first or second. Those in bold face were referred to first.

Target (List A)			Fillers				
Round	**b - p**	**æ-ɛ**					
Landmark	kib	fox bun	tow	buy dun	son	nee one	doom
1	nib	pat **pet**	rick	dot sot	heat	came fame	boom
Landmark	nix	ton win	rem	day pox	foy	gut mon	fit
2	dob	**hat** het	rep	mum sum	peat	mip hip	fin
Landmark	hob	hum mot	mim	but mig	cow	age con	sate
3	nob	tap **tep**	ret	seep meep	seek	neck nick	bate
Landmark	ace	kin hut	ring	gum dow	cox	six wiz	zake
4	sib	**sack** seck	rip	cook dook	teap	soy goy	cake
Landmark	pix cue	pin	cop hun	wit bog	yon	soon	jane
5	mob **mop**	hack	ren len	name dame	beat	noon	sane
Landmark	joy box	mat	sox aim	eye nut	tux	zin	pot
6	**dib** dip	bat	rit lit	moon boon	seap	by	hot
Landmark	vox coz	tin	jot hey	wis tea	cot	vet	take
7	bib **bip**	sat	rim lim	sun pun	heak	net	tape
Landmark	bey fix	cap	bow kif	hin pea	mun	tin	sing
8	**sob** sop	sap	reck leck	toy hoy	seat	boy	king

Appendix B.2. Stimuli from List B in Experiment 2. Baseline items are in white, primed items are in gray, and pairs of pragmatically contrasting items are in dark gray. The pragmatically contrasting items within a pair were referred to one right after another, counterbalanced for whether the one with the phonetic segment missing from L1 was referred to first or second. Those in bold face were referred to first.

Target (List B)			Fillers				
Round	**b - p**	**æ-ɛ**					
Landmark	joy box	mat	sox aim	wit bog	yon	zin	pot
1	nib **nip**	pat	rick lick	name dame	beat	by	hot
Landmark	pix cue	tin	cop hun	eye nut	tux	soon	jane
2	**dob** dop	hat	rep lep	moon boon	seap	noon	sane
Landmark	fix bey	cap	bow kif	wis tea	cot	tin	sing
3	nob **nop**	tap	ret let	sun pun	heak	boy	king
Landmark	vox coz	pin	jot hey	hin pea	mun	vet	take
4	**sib** sip	sack	rip lip	toy hoy	seat	net	tape
Landmark	nix	fox bun	rem	buy dun	son	nee one	doom
5	mob	hack **heck**	ren	dot sot	heat	came fame	boom
Landmark	kib	ton win	tow	day pox	foy	six wiz	fit
6	dib	**bat** bet	rit	mum sum	peat	soy goy	fin
Landmark	ace	hum mot	ring	but mig	cow	age con	sate
7	bib	sat **set**	rim	seep meep	seek	neck nick	bate
Landmark	hob	kin hut	mim	gum dow	cox	gut mon	zake
8	sob	**sap** sep	reck	cook dook	teap	mip hip	cake

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jml.2015.01.001.

References

Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society, 39*, 437–456.
Babel, M. (2011). Imitation in speech. *Acoustics Today, 7*, 16–23.
Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*, 177–189.

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes, 24*, 527–554.

Bard, E. G., & Aylett, M. (2000). Accessibility, duration, and modeling the listener in spoken dialogue. In *Proceedings of Gotalog 2000, 4th workshop on the semantics and pragmatics of dialogue*. Gothenburg University, Gotalog, Sweden..

Bard, E. G., Anderson, A. H., Sotillo, C. F., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language, 42*, 1–22.

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language, 46*, 391–418.

Beebe, L. M. (1981). Social and situational factors affecting the communicative strategy of code-switching. *International Journal of the Sociology of Language, 32*, 139–149.

Beebe, L., & Giles, H. (1984). Speech-accommodation theories: A discussion in terms of second-language acquisition. *International Journal of the Sociology of Language, 46*, 5–32.

Bell, A. (1984). Language style as audience design. *Language in Society, 13*, 145–204.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775–794.

Bly, B. (1993). *Uncooperative language and the negotiation of meaning*. Unpublished doctoral dissertation, Stanford University, Stanford, CA..

Bock, J. K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review, 3*, 395–421.

Boersma, P., & Weenink, D. (2007). *Praat: Doing phonetics by computer* (Version 4.6.29)..

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*, 303–328.

Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes, 23*, 119–147.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 119–135). London: Academic Press.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition, 75*, B13–B25.

Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics, 47*, 1–46.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 6*, 1482–1493.

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology, 19*, 441–472.

Chang, W., & Idsardi, W. J. (2001). Speech perception and production of English final voicing contrast by Korean speakers. *Harvard Studies in Korean Linguistics, IX*, 117–129.

Cho, T., & McQueen, J. (2006). Phonological versus phonetic cues in native and nonnative listening: Korean and Dutch listeners' perception of Dutch and English consonants. *Journal of the Acoustical Society of America, 119*, 3085–3096.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as collaborative process. *Cognition, 22*, 1–39.

Clements, G. N. (2001). Representational economy in constraint-based phonology. In T. Alan Hall (Ed.), *Distinctive feature theory* (pp. 71–146). Berlin & New York: Mouton de Gruyter.

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes, 23*, 528–556.

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptations in language production: Limiting the role of the "Model of the Listener". In D. Napoli & J. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman* (pp. 105–129). Hillsdale, NJ: Erlbaum.

Eckman, F. R., & Iverson, G. K. (2013). The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition, 35*, 67–92.

Flege, J. (1988). The development of skill in producing English word-final stops: Kinematic parameters. *Journal of the Acoustical Society of America, 84*, 1639–1652.

Flege, J., Bohn, O.-S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics, 25*, 437–470.

Flege, J., & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics, 15*, 67–83.

Flege, J., McCutcheon, J., & Smith, S. (1987). The development of skill in producing word final English stops. *Journal of the Acoustical Society of America, 82*, 433–447.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee?. *Journal of Memory and Language, 62*, 35–51.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27*, 181–218.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. London: Academic Press.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.

Hancin-Bhatt, B. J. (1994). Segment transfer: A consequence of a dynamic system. *Second Language Research, 10*, 241–269.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science, 28*, 105–115.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*, 43–61.

Hayes-Harb, R., Smith, B., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*, 664–679.

Hogan, J. T., & Rozsypal, A. J. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America, 67*, 1764–1771.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language, 47*, 589–606.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground?. *Cognition, 59*, 91–117.

Hwang, J. (2011). *Non-native perception and production of foreign sequences*. Unpublished doctoral dissertation, Stony Brook University, Stony Brook, NY..

Ingram, J. C. L., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25*(3), 343–370.

Jones, C. (2003). *The development of phonological categories in children's perception of final voicing in dialects of English*. Unpublished doctoral dissertation, University of Massachusetts Amherst, Amherst, MA..

Kim, M. (2012). *Phonetic accommodation after auditory exposure to native and nonnative speech*. Unpublished doctoral dissertation, Northwestern University, Chicago, IL..

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Journal of Laboratory Phonology, 2*, 125–156.

Kraljic, T., & Brennan, S. E. (2005). Using prosody and optional words to disambiguate utterances: For the speaker or for the addressee?. *Cognitive Psychology, 50*, 194–231.

Kuhlen, A., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20*, 54–72.

Lisker, L. (1957). Closure duration and the intervocalic voiced–voiceless distinction in English. *Language and Cognitive Processes, 33*, 42–49.

Major, R. C., & Faudree, M. C. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition, 18*, 69–90.

Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology, 46*, 749–760.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research, 47*, 1048–1058.

Nittrouer, S. (2004). The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults. *The Journal of the Acoustical Society of America, 115*, 1777–1790.

Oh, G. (2011). *The effect of age of acquisition and second-language experience on segments and prosody: a cross-sectional study of Korean bilinguals' English and Korean production*. Ph.D. dissertation, University of Oregon..

Pardo, J. S., Cajori, Jay I., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, and Psychophysics, 72*, 2254–2264.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics, 40*, 190–197.

Peterson, G., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.
Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 167–226.
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329–392.
Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.
Sirsa, H., & Redford, M. A. (2013). The effects of native language on Indian English sounds and timing patterns. *Journal of Phonetics, 41*, 393–406.
Sohn, H.-M. (1999). *The Korean language*. Cambridge: Cambridge University Press.
Stent, A., Huffman, M., & Brennan, S. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication, 50*, 163–178.
Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of Acoustical Society of America, 88*, 97–100.
White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*, 501–522.
Willemyns, M., Gallois, C., Callan, V. J., & Pittam, J. (1997). Accent accommodation in the job interview: Impact of interviewer accent and gender. *Journal of Language and Social Psychology, 16*, 3–22.
Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*, 245–261.