

# On the assimilation-discrimination relationship in American English adults' French vowel learning<sup>a)</sup>

Erika S. Levy<sup>b)</sup>

*Program in Speech and Language Pathology, Department of Biobehavioral Sciences, Teachers College, Columbia University, 525 West 120th Street, Box 180, New York, New York 10027*

(Received 30 January 2009; revised 22 June 2009; accepted 15 August 2009)

A quantitative “cross-language assimilation overlap” method for testing predictions of the Perceptual Assimilation Model (PAM) was implemented to compare results of a discrimination experiment with the listeners' previously reported assimilation data. The experiment examined discrimination of Parisian French (PF) front rounded vowels /y/ and /œ/. Three groups of American English listeners differing in their French experience (no experience [NoExp], formal experience [ModExp], and extensive formal-plus-immersion experience [HiExp]) performed discrimination of PF /y-u/, /y-o/, /œ-o/, /œ-u/, /y-i/, /y-ɛ/, /œ-ɛ/, /œ-i/, /y-œ/, /u-i/, and /a-ɛ/. Vowels were in bilabial /rabVp/ and alveolar /radVt/ contexts. More errors were found for PF front vs back rounded vowel pairs (16%) than for PF front unrounded vs rounded pairs (2%). Overall, ModExp listeners did not perform more accurately (11% errors) than NoExp listeners (13% errors). Extensive immersion experience, however, was associated with fewer errors (3%) than formal experience alone, although discrimination of PF /y-u/ remained relatively poor (12% errors) for HiExp listeners. More errors occurred on pairs involving front vs back rounded vowels in alveolar context (20% errors) than in bilabial (11% errors). Significant correlations were revealed between listeners' assimilation overlap scores and their discrimination errors, suggesting that the PAM may be extended to second-language (L2) vowel learning. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3224715]

PACS number(s): 43.71.Hw, 43.71.An, 43.71.Es, 43.71.Gv [AJ]

Pages: 2670–2682

## I. INTRODUCTION

A predominant model of cross-language speech perception, the Perceptual Assimilation Model (PAM) (Best, 1995), posits that the perceived similarity of non-native segments to native categories, i.e., gestural constellations in native phonological space, predicts the difficulties naïve listeners will encounter in discriminating speech sounds in a non-native language. Exploring the extension of the PAM from the realm of naïve listeners to second-language (L2) learners, Best and Tyler (2007) proposed the PAM-L2 and called for research examining whether the principles involved in cross-language speech perception by naïve listeners also apply to L2 learning.

A limitation of the PAM (Best, 1995), the PAM-L2 (Best and Tyler, 2007), and of other speech perception and production models, such as Flege's (1995) Speech Learning Model (SLM), is that they are formulated qualitatively, with no objective measure of similarity between native and L2 speech sounds. The present study introduces the “cross-language assimilation overlap” method as a quantitative method for testing the claim by PAM and PAM L2 (Best, 1995; Best and Tyler, 2007) that perceived similarity of native and non-native (or L2) speech sounds predicts how accurately the non-native sounds will be discriminated. A study is reported on the discrimination of Parisian French (PF) front rounded

vowels by native American English (AE) L2 learners of French. The discrimination results are compared to Levy's (2009) perceptual assimilation results by means of the cross-language assimilation overlap method.

In its original form, the PAM (Best, 1995) posits that naïve listeners perceptually assimilate speech sounds of an unfamiliar language into native categories and that their assimilation patterns predict the relative accuracy with which they will discriminate the segments. Non-native segments assimilate as gradiently “good” to “poor” instances of native categories along a continuum. In single-category assimilation, for example, segments that contrast in a non-native language are both assimilated as equally good or poor exemplars of the same native language category, yielding the highest degree of discrimination difficulty. In category-goodness assimilation, two non-native speech sounds are assimilated into the same native category, but one of the segments is perceived as a better instance than the other. In proposing the PAM-L2, Best and Tyler (2007) posited that when a category-goodness assimilation pattern occurs, there is little incentive for a new category to be learned for the less deviant L2 phone. The authors suggest that the deviant phone may be initially learned as a variant of the native category and that with continued L2 exposure, the language learner becomes more attuned to the relevant contrasts between the phones and creates a new L2 category. A factor in determining the creation of new L2 categories is whether the L2 contains minimally contrasting words that occur frequently in dense phonological neighborhoods, increasing the communicative necessity of perceiving the contrast.

<sup>a)</sup> Portions of this work were presented at the Acoustical Society of America meeting held in Providence, RI, in June, 2006.

<sup>b)</sup> Author to whom correspondence should be addressed. Electronic mail: levy@tc.columbia.edu

In the PAM (Best, 1995) framework, two-category assimilation involves each non-native segment assimilating to a separate native category. An uncategorizable segment is assimilated within the native phonological space, but outside any native category. When the uncategorizable segment is paired with a segment that is similar to an AE category, an “uncategorized-categorized” assimilation pattern emerges. Both segments may also be uncategorizable in the native language. When two segments assimilate to separate native categories or as worse or better exemplars or as uncategorizable and categorizable exemplars, these patterns are expected to yield more accurate discrimination than pairs that fall into a single-category assimilation pattern. According to the PAM-L2 (Best and Tyler, 2007), if both L2 phones are assimilated in an uncategorizable pattern, learning depends, to some extent, on how similar the L2 phones are perceived to be to native phones that approximate them in phonological space.

Researchers have operationalized definitions of assimilation patterns referred to by the PAM (Best, 1995) in diverse ways. For example, Best *et al.* (2001) designated a non-native speech sound as “uncategorized” if a listener’s orthographic transcription of the sound suggested one that fell between two or more native English categories. Other researchers (e.g., Levy, 2009; Strange *et al.*, 2009) have used inter- and intra-subject consistencies of categorization as indications of whether sounds are uncategorized. Harnsberger (2001) determined a speech sound to be uncategorized when its top label represented less than 90% of a group’s responses. A limiting consequence of this classification method, which apports continuous ranges into categories, is evident in those patterns referred to by Harnsberger (2001) as “borderline” cases. For example, if a group assimilates a non-native sound to a native category on 89% of trials and the other to another on 91% of another native category, this pattern is considered uncategorized-categorized, as it just misses criterion for the “two-category” or uncategorized-uncategorized patterns. Harnsberger (2001) responded to this type of problem by including borderline scores in more than one assimilation type (e.g., uncategorized-categorized and uncategorized-uncategorized) in his analysis.

Category goodness-of-fit ratings have also been relied on diversely in the field. For example, Best (1995) and Kuhl and Iverson (1995) found goodness ratings to be a strong predictor of discrimination accuracy. Guion *et al.* (2000) combined identification and goodness ratings into one metric in order to examine the relationship between cross-linguistic mapping patterns and discrimination. In contrast, Levy (2009), Strange *et al.* (2005, 2009) found that in phrase- and sentence-level non-native vowel perception experiments, listeners made use of a small range of goodness-of-fit ratings; thus, these studies made limited use of ratings in their analyses.

The various operational definitions of “categorization” and of “category goodness” may yield more than one way to classify assimilation patterns, thus leading to different predictions of discrimination accuracy. The cross-language assimilation overlap method introduced in this study was developed as a quantitative technique for examining perceptual

assimilation and discrimination relationships. Rather than relying on the more typically used method of categorizing patterns according to type of perceptual assimilation (e.g., two-category, category-goodness, etc.) and then comparing expected performance based on perceptual assimilation type with actual performance (e.g., Best *et al.*, 1996, 1988; Harnsberger, 2001), this method ranks perceived similarity, which is quantified by an “overlap” score, and examines the correlation between listeners’ overlap and their discrimination accuracy. Overlap is defined as the smaller percentage of responses when two members of a pair of non-native (or L2) speech sounds are assimilated to the same native category. This method permits the rank ordering of vowel contrasts in terms of difficulty predicted from perceptual assimilation patterns, without making reference to goodness ratings. A goal of the present study was to determine whether such an analysis would find a relationship between perceptual assimilation overlap and discrimination errors in French vowel learning.

It should be noted that the studies using perceptual assimilation and discrimination tasks in the PAM (Best, 1995) tradition have focused mostly on naïve listeners’ performance (e.g., Best *et al.*, 1996, 1988; Best and Strange, 1992; Strange *et al.*, 2001). Few experiments thus far (e.g., Guion *et al.*’s [2000] study of consonant perception) have examined L2 learners’ discrimination patterns. To the author’s knowledge, none has been used to examine vowel perception by experienced learners, even though accurate vowel recognition has been found to be more important than consonant recognition for overall sentence intelligibility (Kewley-Port *et al.*, 2007).

## II. THE DISCRIMINATION EXPERIMENT

This section reports a study of the effects of formal and immersion language experience and consonantal context on AE listeners’ discrimination of PF contrasts involving front rounded vowels. French high front rounded /y/ and mid front rounded /œ/ are produced with the tongue forward and the lips protruded (Tranel, 1987). English, in contrast, has no canonical front rounded vowels, although in several AE dialects, /u/, /ʊ/, and /o/ have become more “fronted,” i.e., produced with the tongue farther forward in the oral cavity (Clopper *et al.*, 2005; Strange *et al.*, 2007). Findings are mixed regarding AE speakers’ discrimination of front rounded vowels from other French vowels. High accuracy is reported in Best *et al.*’s (1996) categorial<sup>2</sup> discrimination study, in which naïve AE listeners discriminated Bretagne French /sœ-sy/ syllables with fewer than 5% errors. Similarly, in a study involving L2 learners, Flege and Hillenbrand (1984) tested native English speakers proficient in French on paired /tu-ty/ tokens produced by seven native French speakers from France and Belgium. Listeners identified which member of the pair was /ty/ with an error rate of only 10%.

Greater problems in discrimination of front rounded vowels were found for even advanced AE learners of French in Gottfried’s (1984) categorial discrimination study. AE listeners with and without French experience and native French listeners heard productions of PF vowels /e-ɛ/, /a-ɛ/, /i-ɛ/,

/a-ɔ/, /y-u/, /a-a/, /y-ø/, and /œ-ø/ in /tVt/, /Vt/, /tV/, and /V/ syllabic contexts, uttered as if in sentences. Vowels in isolation were discriminated more accurately than vowels in /tVt/ context by all three groups.

In the reviewed studies, the vowel stimuli were presented either in alveolar context or in isolation. Production studies indicate that vowels vary depending on their consonantal contexts (Hillenbrand *et al.*, 2001) and that patterns of variation differ in different languages (Strange *et al.*, 2007), suggesting that learning coarticulatory patterns of variation may be part of the L2 speech learning process (Beddor *et al.*, 2002; Levy and Law II, 2008; Manuel, 1999; Oh, 2008). Phonetic context may affect vowel perception (Bohn and Steinlen, 2003), as well. Strange *et al.* (2009) found effects of consonantal context and speaking style (i.e., citation form disyllables vs sentences) on assimilation of French and German vowels in sentences by naïve AE listeners. For example, PF /y/ was more often assimilated to AE /u/ in alveolar (94%) than in bilabial (74%) context.

In an investigation of context effects in L2 learning, Levy and Strange (2008) extended Gottfried's (1984) study, examining AE listeners' discrimination of PF vowels /y/, /œ/, /u/, and /i/ in /rabVp/ and /radVt/ bisyllables in AXB triads of the phrase "neuf /raCVC/ à des amis," ("nine /raCVC/ to some friends"). (In the AXB paradigm, stimuli are presented in triads, with the second matching the first or the third.) Two groups of AE listeners participated: The "inexperienced group" consisted of AE listeners with no French experience. The "experienced group" was highly proficient in French, with extensive classroom and immersion French experience. Results showed effects of French language experience and consonantal context on AE listeners' discrimination of the French contrasts. The experienced group made fewer errors (5%) than the inexperienced group (24%) for three experimental pairs (PF /y-i/, /œ-u/, and /y-œ/). For PF /y-u/, no statistically significant difference (24% for Inexp vs 30% for Exp) was revealed as a function of language experience, pointing to this contrast as a particularly difficult one to master. The inexperienced group made more PF /y-u/ errors in alveolar context (8% in bilabial vs 39% in alveolar), but more PF /y-i/ errors in bilabial context (24% in bilabial vs 8% in alveolar). The experienced group confused PF /y-u/ in both contexts (24% in bilabial and 35% in alveolar context) with great between-subject variation. For all contrasts except PF /y-i/, where the opposite was the case, the inexperienced group made fewer errors in bilabial than in alveolar context. No significant context effect was found for the experienced group.

An explanation for the context-dependent performance by these L2 learners makes reference to the relationship between native and L2 vowel production. High back AE vowels /u/, /ʊ/, and /o/, to a lesser extent, are fronted in alveolar contexts (Hillenbrand *et al.*, 2001; Strange *et al.*, 2007). (See Fig. 1 in Levy, 2009, for a vowel plot of the PF vowel stimuli superimposed onto AE vowel space.) Thus, in AE, the phonological category /u/ has relatively back rounded [u] instantiations in most contexts (e.g., "cool" [kul]), but when the tongue is forward, as in alveolar context in "dude" [dyd], for example, AE /u/ approximates a front rounded vowel. Thus,

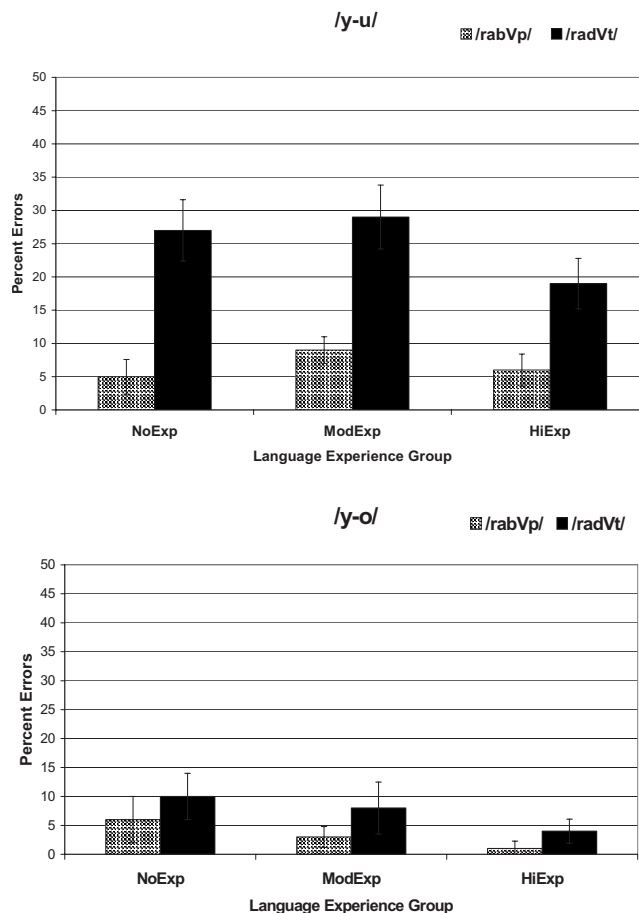


FIG. 1. Language and context effects on /y/ discrimination. Categorical discrimination of PF /y-u/ (top) and /y-o/ (bottom) in bilabial (/rabVp/) and alveolar (/radVt/) contexts by AE listeners with no French experience (NoExp), moderate French experience (ModExp), and extensive French experience (HiExp): percent errors and standard errors.

to native speakers of languages with front rounded vowel categories, [u] and [y] represent two different phonological categories. In English, on the other hand, the segments [y] and [u] may be allophones of the phonological category /u/. AE listeners, then, may tend to perceive high and mid front rounded vowels as more similar to AE /u/ or /ʊ/ (in which fronting can be expected) when the segments are in alveolar context than when they are in other contexts. Thus, they may confuse front rounded vowels with back PF vowels that also assimilate to back AE categories, especially in alveolar context.

The patterns with which L2 learners assimilate vowels as a function of L2 experience and consonantal context were investigated in a perceptual assimilation study by Levy (2009). AE listeners with no French experience (NoExp group), AE listeners with formal French classroom learning experience, but no immersion (ModExp), and AE learners with extensive classroom and immersion experience (HiExp group) participated. Listeners performed an assimilation task involving PF /y, œ, u, o, i, ε, a/ in bilabial /rabVp/ and alveolar /radVt/ contexts, presented in phrases. They were given a choice of 13 AE key words ("heed, hid, hayed, head, had, hod, hawed, hud, hoed, hood, who'd, hued, and herd") and were asked to select the word that contained the vowel

most similar to the target. They rated the vowel on a scale from 1–9 (“most foreign-sounding” to “most English-sounding”).

Levy (2009) found that front rounded vowels were assimilated primarily to back AE vowels (PF /y/ to palatalized AE /<sup>h</sup>u/, and PF /œ/ to AE /u/). Back rounded PF vowels were also assimilated to back AE vowel categories (PF /u/ to AE /u/, and PF /o/ to AE /u/ and /o/). No language experience effect was found for PF /y/ to AE /<sup>h</sup>u/ assimilation in alveolar context (NoExp=65%, ModExp=71%, and HiExp=61%). In bilabial context, on the other hand, listeners with extensive experience assimilated PF /y/ to AE /<sup>h</sup>u/ less often (72%) than listeners with no (80%) or only formal (85%) experience. For PF /œ/, assimilation patterns differed as a function of language experience and consonantal context. With extensive experience, /œ/ assimilated more often to AE /u/ (e.g., in alveolar context, NoExp=17% and HiExp=61%) or /ɜ:/ (e.g., in alveolar context, NoExp=0% and HiExp=20%) and less to /u/ (e.g., in alveolar context, NoExp=59% and HiExp=9%). Both front rounded vowels assimilated more often to AE /u/ in alveolar context (e.g., for PF /y/ assimilation to AE /u/, NoExp=31%) than in bilabial context (NoExp=7%).

The PAM (Best, 1995) predicts poor discrimination of contrasts that assimilate in a single-category pattern. Hence, according to the PAM, Levy’s (2009) finding of perceptual assimilation of front rounded vowels to back vowels was consistent with AE listeners’ greater difficulty distinguishing front rounded vowels from back rounded vowels than from front unrounded vowels reported by Levy and Strange (2008). However, AE listeners’ assimilation of front rounded vowels to back vowels was not consistent with Best’s *et al.*’s (1996) and Flege and Hillenbrand’s (1984) finding of relatively high accuracy in /y-u/ discrimination in naïve listeners and L2 learners. These inconsistencies may be thought of in terms of the Automatic Selective Perception model of speech perception (Strange and Shafer, 2008), which posits that L1 selective perceptual routines are relied on to a greater extent when task demands increase. It is possible that the tasks in Levy and Strange’s (2008) discrimination study and Levy’s (2009) assimilation study, involving vowels in phrases uttered by three speakers in continuous speech, were more demanding than tasks in earlier studies using citation materials uttered by a single speaker, for example, yielding poorer perceptual outcomes.

The present study investigated the effects of French language experience and consonantal context on AE listeners’ discrimination of L2 French vowels, extending Levy and Strange’s (2008) discrimination study in three ways: First, an additional group of listeners (ModExp) with just classroom experience was tested. Second, for a more comprehensive examination, discrimination of additional vowel pairs (front rounded vs back rounded /y-o/ and /œ-o/; and front rounded vs front unrounded /y-ɛ/, /œ-ɛ/, and /œ-i/) was targeted as well as the four also examined by Levy and Strange (2008), i.e., front rounded vs front unrounded /y-i/, front rounded vs back rounded, /y-u/ and /œ-u/, and front rounded vs front rounded /y-œ/. And finally, the same participants whose assimilation data were collected for Levy (2009)<sup>3</sup> were tested

on discrimination in order for assimilation-discrimination relationships to be examined, as described in Sec. III.

This study investigated (1) whether AE L2 learners of French had more difficulty discriminating PF front rounded vowels from PF front unrounded vowels or from PF back rounded vowels, (2) whether level of French experience affected the listeners’ discrimination accuracy, and (3) whether consonantal context affected their discrimination accuracy.

Because previous studies indicate that the discrimination of front vs back rounded vowels tends to be more difficult for AE listeners than does the discrimination of front rounded PF vowels vs other front vowels (Gottfried, 1984; Levy and Strange, 2008), more front vs back rounded vowel confusions were expected in the present experiment than front rounded vs unrounded vowel confusions. Overall, it was predicted that experienced L2 learners would demonstrate more accurate discrimination of contrasts than would inexperienced learners (Levy, 2009; Levy and Strange, 2008). Specifically, the HiExp group was expected to perform more accurately on the categorial discrimination task than the ModExp group, who was expected to perform more accurately than the NoExp group. However, based on the findings of Gottfried (1984) and Levy and Strange (2008), even the most-experienced AE listeners were predicted to have difficulty with the /y-u/ contrast. Other pairs expected to be difficult for the less-experienced listeners were /y-œ/ and /œ-o/, i.e., front rounded vowels paired with each other and front rounded vowels paired with vowels of a similar height. Consonantal context was expected to have a significant effect on discrimination, especially for inexperienced listeners, with front rounded vowels being less accurately discriminated in alveolar than in bilabial context.

## A. Method

### 1. Stimulus materials

The stimulus materials for this study were identical to those described by Levy (2009). In brief, three female adult native PF speakers who had lived in the United States for less than a year were recorded as they read nine PF vowels, blocked by bilabial /rabVp/ or alveolar /radVt/ context in the sentence: “J’ai dit neuf /raCVC/ à des amis.” (I said nine /raCVC/ to some friends.) A Shure microphone fed the signal to a Soundblaster Live Wave sound card via an Earthworks microphone preamp. The digital files were segmented so that only “neuf /rabVp/ à des amis” and “neuf /radVt/ à des amis” remained, with the target front rounded /y, œ/ and /i, u, ɛ, o, a/ for comparison. Task verification was accomplished by three monolingual native PF speakers visiting the United States for less than a month, who performed the categorial discrimination task (described below). They made no (0) errors on the experimental pairs, a total of three errors (=3% errors per pair) on the non-experimental pairs PF /u-i/, /y-ɛ/, and /y-o/ and reported that they had no difficulty performing the task.

An acoustic analysis of the PF stimuli was conducted by Levy (2009) and compared to AE acoustic values in Strange *et al.*’s (2007) production study. Although a full description is beyond the scope of this article, the following should be

noted: In bilabial context, PF /y/ approximated PF and AE /i/ far more than it approximated PF and AE /u/. PF /œ/ was intermediate between front AE vowels and back AE vowels. In alveolar context, PF /y/ still approximated PF /i/ more than it approximated PF /u/. However, in alveolar context, both PF and AE vowels /u/ and /o/ were fronted compared to their counterparts in bilabial context. PF /œ/ was only slightly fronted in this context. Thus, if the naïve and experienced participants had more difficulty discriminating PF /y/ from back than from front vowels, acoustics alone could not explain their patterns.

## 2. Participants

The three groups of participants in this experiment were the same as those described by Levy (2009). All 39 participants were raised in monolingual English-speaking households in the United States. The NoExp group was comprised of 13 native AE speakers, ages 20–40 years, who were living in New York City and had never studied French, nor lived in a French-speaking country, nor interacted significantly with French speakers. The ModExp group were 13 native AE speakers, ages 22–37 years, who were living in New York City, and had studied French in classroom settings, but had minimal French immersion experience. They had started learning French at a mean age of 16.1 years (SD=2.8) for 2–4 years (mean=3 years and SD=0.8). They had not lived in a French-speaking country for more than 5 months. The HiExp group were 13 native AE speakers, ages 20–61 years, with extensive classroom and immersion French experience, who were speaking French regularly (range=2 h/week–100% of the time, median=15 h/week). They had studied French for a mean of 8 years (range=5–13 years and SD=2.4), starting no earlier than age 12 years (mean age of starting=14 years and SD=1.6). They had spent at least 1 year living in a French-speaking country in adulthood (range=1–16 years and median=1.4 years), and spoke French frequently around the time of the experiment. Participants passed a hearing screening at 20 dB.

## 3. Procedure

Participants listened to the discrimination stimuli presented by STAX Professional SR Lambda headphones connected to an amplifier (STAX Professional SRM-1/MK-2), receiving the signal from the Dell Dimension XPS B800 computer in a sound-attenuated chamber. The five experimental “one-feature” vowel pairs presented were PF /y-i/, /y-u/, /œ-ɛ/, /œ-o/, and /y-œ/. These were contrasts whose members differed in just one feature (e.g., rounded vs unrounded for PF /y-i/ or front vs back for PF /y-u/, high vs mid for PF /y-œ/). The six “two-feature” vowel pairs were PF /y-ɛ/, /œ-i/, /y-o/, /œ-u/, /u-i/, and /a-ɛ/, whose members differed by more than one feature (e.g., back rounded vs front unrounded for PF /u-i/). In addition, all of these pairs included at least one vowel with a “counterpart”<sup>4</sup> in AE (/i, u, ɛ, o/). The two-feature pairs PF /u-i/ and PF /a-ɛ/ were considered control pairs because they had counterparts in AE and did not include front rounded vowels. Two-feature vowel pairs were expected to be more accurately discriminated

overall by virtue of being phonologically “more different” than the one-feature pairs in PF. Four orders were possible for presentation of each A-B vowel pair: AAB, ABB, BBA, and BAA. Trials contained triads of stimuli uttered by three different speakers in random order, blocked by consonantal context, with an equal number of correct A and B responses. Conditions were counterbalanced such that all nine tokens of each vowel occurred in each contrasting pair an equal number of times.

The stimuli were randomized and presented using the “Paradigm Discrim” program (by Bruno Tagliaferri). The pairs were arranged into AXB trials. Subjects were instructed to click on “1” if the vowel in the second stimulus was the same vowel in the first, and “3” if it was the same as the vowel as in the third. Prior to testing, AE subjects were given task familiarization in which they were asked to discriminate 18 trials of vowel pairs involving AE /ɛ/, /a/, /œ/, and /i/ vowel pairs in the AXB paradigm. Participants were permitted no more than two errors on task familiarization in order to continue with the experiment. All participants met these criteria.

The AE task familiarization was followed by French stimulus familiarization. Stimulus familiarization was identical to the experimental task. Following the stimulus familiarization in one context, listeners heard 4 blocks of 24 experimental trials in that context, then 1 block of stimulus familiarization trials in the other context, followed by 4 blocks of 24 experimental trials in that context. Each listener completed 12 judgments for each of the five one-feature pairs in each consonantal context, resulting in 60 one-feature trials per context. Six judgments were completed for each of the six two-feature pairs, resulting in 36 judgments on the two-feature pairs. Thus the experiment consisted of a total of 96 triads in each context. The inter-stimulus interval was 500 ms and trials were self-paced.

## B. Results

### 1. Data analysis

Discrimination scores were derived by tallying errors over trials for each contrast in each context and converted to percentages of errors of total number of trials. An error was defined as responding 3 when the trial was AAB or 1 when the trial was ABB.

### 2. Language experience and consonantal context effects

For an overview of categorical discrimination findings, Table I presents the percent errors by each language experience group (NoExp, ModExp, and HiExp across the top row) for each contrasting vowel pair, with consonantal contexts combined. The discrimination scores for the vowel pairs are listed beginning with scores for the experimental pairs, followed by the control pair scores. The individual experimental pairs are discussed in Secs. II B 3–II B 6 with regard to the language experience and consonantal context effects revealed. The overall discrimination errors for the experimental pairs decreased with language experience (mean=13%, 11%, and 3% for NoExp, ModExp, and HiExp, respectively).

TABLE I. Categorical discrimination of PF vowel pairs summed over /rabVp/ and /radVt/ contexts by AE listeners with no French experience (NoExp), moderate French experience (ModExp), and extensive French experience (HiExp): Percent errors and standard error of the mean (in percent) are given.

PF vowel pairs (Expt.)		No Exp	Mod Exp	Hi Exp
		% Error	% Error	% Error
High front rounded vs back rounded	y-u	16	19	12
	y-o*	8	5	3
Mid front rounded vs back rounded	æ-o	33	22	4
	æ-u*	33	29	6
High front rounded vs front unrounded	y-i	7	1	0
	y-ε*	1	1	1
Mid front rounded vs front unrounded	æ-ε	3	2	1
	æ-i*	0	3	1
High front rounded vs mid front rounded	y-æ*	17	16	4
Control pair	a-ε*	19	10	2
Control pair	u-i*	1	1	0

\* Note that vowel pairs with asterisks were two-feature vowel pairs and were presented for 12 judgments per participant. (The others were one-feature pairs, presented for 24 judgments.)

Vowel pairs were not equally difficult to discriminate, with the NoExp group making 0%–33% errors, depending on which contrast was presented.

For an analysis of whether PF front rounded vowels were more often confused with PF back rounded or PF front unrounded vowels, the discrimination data were divided into two scores: Total percent of errors made on pairs containing a front rounded vowel and a front unrounded vowel (PF /y-i/, /y-ε/, /æ-ε/, and /æ-i/) and total percent errors made on pairs containing a front rounded vowel and a back rounded vowel (PF /y-u/, /y-o/, /æ-o/, and /æ-u). When front rounded and unrounded vowels were contrasted, listeners in all groups made few errors (3%, 2%, and 1% for the NoExp, ModExp, and HiExp groups, respectively). When front and back rounded vowels were contrasted, on the other hand, listeners made far more errors (22%, 19%, and 6% for the NoExp, ModExp, and HiExp groups, respectively), as predicted.

Because listeners in all three groups made almost no errors on front rounded vs unrounded pairs, the remaining analyses focused on discrimination of front rounded vowels paired with back rounded vowels and with each other. On the four front rounded vs back rounded pairs (PF /y-u/, /y-o/, /æ-o/, and /æ-u/), the NoExp group made the most errors (22%), followed by the Mod Exp group (19%), followed by the HiExp group (6%). Because of heterogeneity of variance, nonparametric statistics were performed. As described below, a Kruskal–Wallis one-way analysis of variance (ANOVA) was implemented to examine the language experience effects on each of the four vowel pairs and Mann–Whitney U-tests provided pairwise comparisons for language experience and consonantal context.

### 3. Discrimination of PF /y-u/ and /y-o/: Language experience and consonantal context

Figure 1 presents mean errors for discrimination of pairs involving the front rounded vowel /y/ contrasted with the two back rounded vowels /u/ and /o/. The top graph shows percent errors (*Y*-axis) in discrimination of the /y-u/ pair by the NoExp, ModExp, and HiExp Groups (along the *X*-axis).

Scores for each language group are divided into bilabial (left checkered bar) and alveolar (right solid bar) contexts. For the high vowel pair PF /y-u/, listeners performed above chance, but not significantly differently across groups: NoExp = 16%, ModExp = 19%, and HiExp = 12%. A Kruskal–Wallis one-way ANOVA by language group confirmed that the language experience effect was not statistically significant [ $p = 0.22$ ]. This is consistent with Levy and Strange's (2008) finding that advanced listeners of French fared no better than listeners with no French experience for this vowel pair—a contrast that is particularly resistant to improvement. A Mann–Whitney U-test revealed the consonantal context main effect to be statistically significant [ $U = 252, p < 0.001$ ] at a two-tailed significance level, on the other hand, as predicted from Levy's (2009) perceptual assimilation findings, with more difficulty revealed in alveolar context than in bilabial context.

The bottom graph in Fig. 1 presents the data for the PF /y-o/ contrast. As the figure shows, few errors were made by any group on this pair (NoExp = 8%, ModExp = 5%, and HiExp = 3%). With so few errors, no significant experience effect [ $p = 0.21$ ] or consonantal context effect [ $U = 656, p = 0.15$ ] was present.

### 4. Discrimination of PF /æ-o/ and /æ-u/: Language experience and consonantal context

Figure 2 presents discrimination results for pairs involving /æ/ and back rounded vowels. For the PF /æ-o/ contrast (upper graph), the NoExp group made the most errors (26% in bilabial and 39% in alveolar context), followed by the ModExp group (15% in bilabial and 29% in alveolar context), followed by very few errors by the HiExp (1% in bilabial and 6% in alveolar context). A Kruskal–Wallis one-way ANOVA by language group revealed a main effect of language experience [ $p < 0.001$ ], with increased experience being associated with fewer errors in discrimination for this vowel pair. A Mann–Whitney U-test indicated that the ModExp group made significantly fewer errors than did the NoExp group ( $U = 45, p = 0.04$ , two-tailed); thus, formal in-

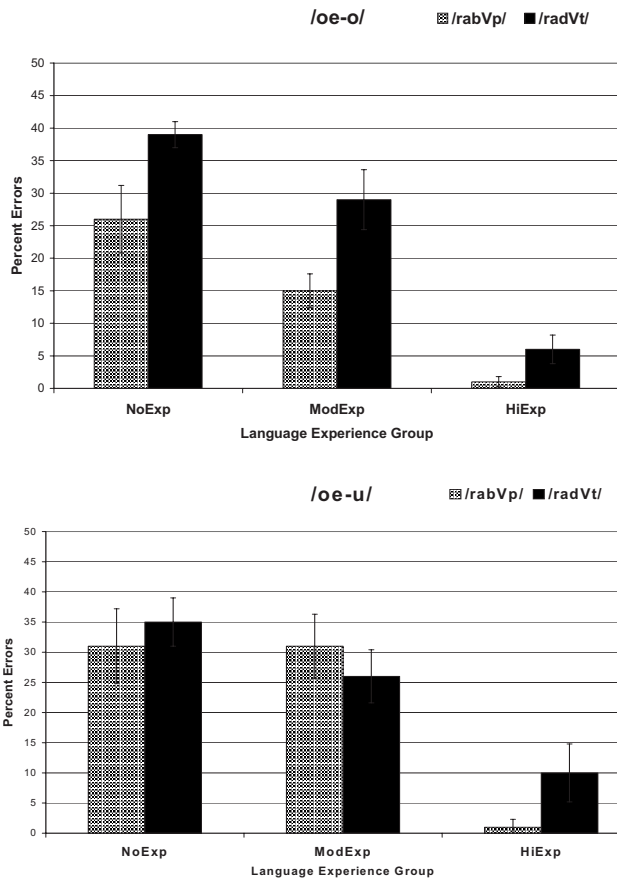


FIG. 2. Language and context effects on /æ/ discrimination. Categorical discrimination of PF /œ-o/ (top) and /œ-u/ (bottom) in bilabial (/rabVp/) and alveolar (/radVt/) contexts by AE listeners with no French experience (NoExp), moderate French experience (ModExp), and extensive French experience (HiExp): percent errors and standard errors.

struction was associated with (marginally) more accurate discrimination for this vowel pair. The HiExp group performed significantly more accurately than did the ModExp group ( $U=10, p<0.001$ ); thus, extensive language instruction and immersion were associated with fewer errors than was formal instruction without immersion. The prediction of a consonantal context effect, based on assimilation differences as a function of context for PF /æ/, was also borne out ( $U=505, p<0.01$ , two-tailed), with more errors in alveolar context than in bilabial for all groups.

The bottom graph in Fig. 2 displays discrimination results for the PF /œ-u/ contrast. A Kruskal–Wallis one-way ANOVA by language group indicated a main effect of language experience [ $p<0.001$ ]. The NoExp group made the most errors (31% in bilabial and 35% in alveolar context), followed by the ModExp group (31% in bilabial and 26% in alveolar context), followed by very few errors by the HiExp (1% in bilabial and 10% in alveolar context). Thus, as predicted, a language effect was present, with increased experience associated with fewer errors in discrimination of this vowel pair. However, for this pair only, the immersion group performed more accurately than the other groups ( $U=13, p<0.001$ , two-tailed). The formal experience group (ModExp) did not perform significantly more accurately than the NoExp group ( $U=13, p=0.39$ , two-tailed). An unex-

pected finding for this pair was the lack of a significant context effect [ $U=688, p=0.45$ ], despite differences in assimilation of /æ/ as a function of context.

### 5. Interaction of vowel pair, language group, and consonantal context for PF /y-u/ and /y-i/

When consonantal context was taken into consideration, the only score to reach above 6% errors in pairs involving front rounded vs unrounded vowels was the score of 10% errors for the PF /y-i/ pair in bilabial context by the NoExp group. Despite the low error rate, this contrast merits examination in light of the interaction of vowel pair, language experience, and consonantal context.

In Levy (2009), a subgroup of NoExp listeners perceptually assimilated PF /y/ to AE front unrounded /i/. (The other language experience groups rarely assimilated /y/ to /i/ in either context.) An interaction was found in the present study, primarily for one individual with no French experience, in which PF /y/ was assimilated to AE /i/ more often in bilabial context than in alveolar context. The interaction in discrimination is consistent with the interaction found in assimilation in Levy (2009). In bilabial context, the NoExp group made more errors for PF /y-i/ (10%) than for PF /y-u/ (5%), whereas in alveolar context, they made more errors for the PF /y-u/ pair (27%) than for the PF /y-i/ pair (4%). As in the assimilation task, this pattern was primarily due to one participant, who made 33.3% errors on PF /y-i/ in bilabial and 0% errors in alveolar context.

A closer examination of that listener’s perceptual assimilation and discrimination patterns provides an example of the PAM (Best, 1995) or the PAM-L2 (Best and Tyler, 2007) being predictive on an individual level: The NoExp listener perceptually assimilated all PF /y/ vowel stimuli in bilabial context to AE /i/ (100% of responses—more than all other listeners). As predicted by the PAM, he had discrimination difficulty (33% errors) with the PF /y-i/ contrast in bilabial context—the highest percentage of errors of any participant on this pair. In alveolar context, on the other hand, he perceptually assimilated PF /y/ exclusively to back vowels (39% to AE /u/, 50% to AE /ʊ/, and 6% to AE /ɨ/—never to AE /i/). As predicted, in alveolar context, he discriminated PF /y-i/ far more accurately (0% errors) than the PF /y-u/ contrast (25% errors). Thus, for this individual listener, the PAM predicted discrimination performance from perceptual assimilation patterns, and most effectively when consonantal context was taken into account.

### 6. Discrimination of PF front rounded vowels /y-œ/: Language experience and consonantal context

As shown in Fig. 3, NoExp listeners made the most errors in differentiating the PF /y-œ/ pair (12% errors in bilabial and 22% in alveolar context), followed by ModExp (10% errors in bilabial and 21% in alveolar context), followed by HiExp (3% errors in bilabial and 4% in alveolar context). A Kruskal–Wallis one-way ANOVA by language group indicated a main effect of language experience [ $p<0.001$ ]. A Mann–Whitney U-test indicated no significant difference between performance of the NoExp group and the

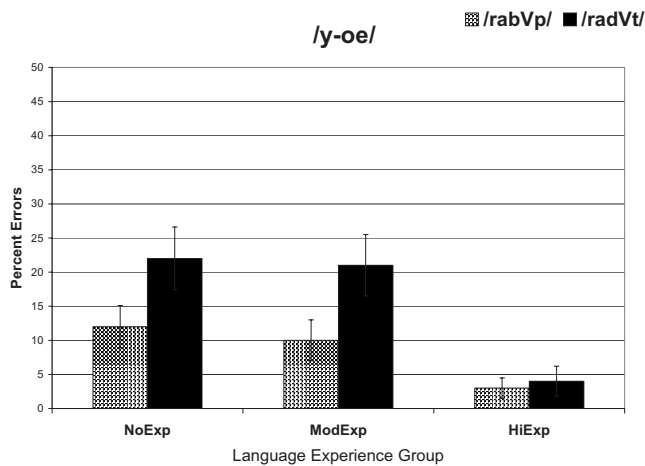


FIG. 3. Categorical discrimination of PF front rounded vowels /y-œ/ in bilabial (/rabVp/) and alveolar (/radVt/) contexts by AE listeners with no French experience (NoExp), moderate French experience (ModExp), and extensive French experience (HiExp): percent errors and standard errors.

ModExp group, [ $U=77$ ,  $p=0.69$ ], but a significant difference between ModExp and HiExp performance, [ $U=21$ ,  $p<0.001$ ]. According to a Mann-Whitney U-test, consonantal context only approached statistical significance [ $U=572$ ,  $p=0.05$ ] for this vowel pair.

### 7. Discrimination of control pairs PF /a-ɛ/ and PF /u-i/

The control pairs PF /a-ɛ/ and PF /u-i/ were, by definition, expected to result in few discrimination errors, based on the assumption that these vowels would fall into a two-category assimilation pattern. For the PF /a-ɛ/ pair, the groups made more errors than expected (NoExp=19%, ModExp=10%, and HiExp=2%). Levy (2009) indicated that the NoExp group perceived both PF /a/ and /ɛ/ as most similar to AE /æ/ some of the time; thus, it appears that listeners without immersion experience may have assimilated these segments in a single-category assimilation pattern instead. The control contrast PF /u-i/ was indeed discriminated without difficulty by all language groups (NoExp=1%, ModExp=1%, and HiExp=0% errors), indicating that listeners were on task.

## C. Discussion

In summary, AE listeners had more difficulty discriminating PF front rounded vowels from PF back rounded vowels than from PF front unrounded vowels. Overall, listeners who had formal-plus-immersion experience with French performed significantly more accurately than those without L2 French experience and those with only formal French instruction experience. Only the PF /y-u/ vowel pair remained relatively difficult for highly experienced listeners. Discrimination of pairs involving the mid front rounded vowel /œ/ with back rounded vowels was more accurate with greater L2 experience, especially with extensive formal-plus-immersion experience. Listeners made more errors with the /œ-u/ pair than with the /œ-o/ pair, despite the height difference in the first pair. For the /y-œ/ pair, an experience effect was also evident in the non-immersion vs immersion groups.

Overall, discrimination of front vs back rounded vowels and discrimination of front rounded vowels from each other was significantly less accurate in alveolar context than in bilabial context. The context effect was evident in both pairs involving vowels of a similar height (PF /y-u/ and /œ-o/). These findings had been predicted based on previous literature, including Levy's (2009) assimilation results. Contrary to expectations, no context effect was found for front rounded vowels paired with vowels of a different height. For the PF /y-o/ pair, this may be attributed to too few errors to reveal a significant interaction. The lack of a context effect in the PF /œ-u/ pair is less interpretable.

Stimuli in nearly all previous studies of the perception of front rounded vowels by AE listeners have been vowels preceded and/or followed by alveolar consonants (e.g., Best *et al.*, 1996; Flege, 1987; Flege and Hillenbrand, 1984; Gottfried, 1984; Polka, 1995; Polka and Bohn, 1996) or produced in isolation (e.g., Gottfried, 1984; Rochet, 1995; Stevens *et al.*, 1969). Results from the present experiment suggest that replications of such studies, but using stimuli in which the vowels are produced in other consonantal contexts, may reveal different results. In bilabial context, for example, AE listeners are likely to make fewer discrimination errors for pairs involving front rounded vowels than indicated in previous literature, although some naïve individuals may have more difficulty discriminating the PF /y-i/ contrast in bilabial than in alveolar context.

The overall effect of more accurate discrimination by the HiExp group than by the ModExp was not true for every vowel pair. Both formal and immersion experience in late L2 learners were associated with increased accuracy in perception of non-native contrasts. For the PF /y-u/ vowel pair, formal experience alone was not associated with greater accuracy, consistent with Levy's (2009) finding of no experience effect for perceptual assimilation of PF /y/ to AE /<sup>h</sup>u/ in alveolar context (a pattern that would predict two-category assimilation, thus higher discrimination accuracy), but not with the finding of an experience effect of decreased PF /y/ to AE /<sup>h</sup>u/ assimilation by the HiExp group in bilabial context. In the present study, listeners immersed in French for several years performed essentially the same as those with no French experience, lending support to studies that point to the PF /y-u/ contrast as one particularly resistant to perceptual learning by AE listeners (e.g., Gottfried, 1984; Levy and Strange, 2008).

Compared to naïve listeners, listeners with formal training alone discriminated the PF mid-vowel pair /œ-o/ more accurately, and listeners with extensive formal training and immersion performed with the greatest accuracy. No higher discrimination accuracy for the PF vowel pairs /œ-u/ and /y-œ/ was associated with merely formal training, but extensive training and immersion were associated with significantly more accurate discrimination. That discrimination accuracy with formal instruction only was not greater than with no French exposure supports the notion that, to be most effective, language instruction programs must include more than the typically administered foreign language requirements in United States schools.



### III. TESTING THE PAM ON L2-VOWEL LEARNING

In Sec. II, discrimination results were, for the most part, predicted based on the perceptual assimilation patterns reported in Levy (2009), with confusions arising when the front rounded and back rounded PF vowels in a pair assimilated the same back AE categories, which occurred most often in alveolar context. On an individual level, it was shown that a participant with no French experience, who assimilated PF /y/ to front vowels in bilabial context and to back vowels in alveolar, also had more difficulty discriminating PF /y-i/ in bilabial context than in alveolar context.

This section reports a more systematic examination of discrimination accuracy for the vowel pairs tested in relation to the same listeners' assimilation patterns, accomplished through the cross-language assimilation overlap method. This method was used to examine the relationship (i.e., correlation) between degree of overlap in assimilation (i.e., how often two non-native vowels perceptually assimilated to the same native category) and the discriminability of vowel pairs in order to test the predictions generated by the PAM (Best, 1995) for L2 vowel learning (Best and Tyler, 2007).

That is, by quantifying perceptual assimilation overlap (e.g., for the PF /y-u/ pair, how often tokens of both PF /u/ and PF /y/ assimilated to the same AE vowel category /u/), it was possible to place contrasting pairs along a continuum from most similar to least similar. This permitted more finely grained predictions to be made about relative discrimination accurately for vowel pairs. Additionally, for the purposes of the present study, it was not evident how to characterize the assimilation of AE /u/ and /<sup>h</sup>u/ response categories in Levy (2009). It was not clear whether this was two-category (palatalized /u/ vs nonpalatalized /u/), category goodness (allophonic variation), or single-category (phonological /u/) perceptual assimilation.

It was hypothesized that (1) vowel pairs whose members assimilated to separate categories (by each language experience group) would be discriminated more accurately (by the same language experience group) than those pairs whose members assimilated to the same categories, an outcome predicted by the PAM (Best, 1995) for naïve listeners and the PAM-L2 (Best and Tyler, 2007) for L2 learners, and that (2) the more trials in which an individual assimilated both members of a vowel pair to the same native category (i.e., the higher the overlap score), the less accurate the individual's discrimination would be for that vowel pair. Both hypotheses were expected to be true for all three language experience groups and in both consonantal contexts.

#### A. Cross-language assimilation overlap by language experience group

##### 1. Data analysis

In testing the first hypothesis, that vowel pairs whose members assimilated to separate categories would be more discriminable than those whose members did not, the cross-language assimilation overlap method proceeded as follows:<sup>5</sup> Vowel pairs were the sampling variable. For this analysis, an overlap score was obtained for each vowel pair within each language group. The overlap was operationally defined as the

smaller percentage of responses when two members of a PF pair were perceptually assimilated to a particular AE vowel category. For the /y-u/ experimental vowel pair in bilabial context, for example, when PF /u/ was presented to NoExp listeners in the perceptual assimilation task, the modal response (90.2%) was /u/. When /y/ was presented, 6.8% of stimuli were categorized as /u/; thus for 6.8% of the stimuli (the portion that overlaps between 90.2% and 6.8%, i.e., the smaller percentage), perception of /y/ and /u/ overlapped. In addition, the NoExp group categorized PF /u/ as /<sup>h</sup>u/ for 6.4% of the stimuli, which overlapped with the modal choice of /<sup>h</sup>u/ (79.9%) when PF /y/ was presented. Both /y/ and /u/ also were perceived as closest to /i/ for an overlap of 0.4%, and both were perceived as /ʊ/ for an overlap of 1.7%. Thus, when 6.8%, 6.4%, 4%, and 1.7% (the overlap percentages when both stimuli were assimilated to the same AE vowel) were summed, the result was a total overlap score of 15.3% for the perception of /y-u/ by the NoExp group in bilabial context.

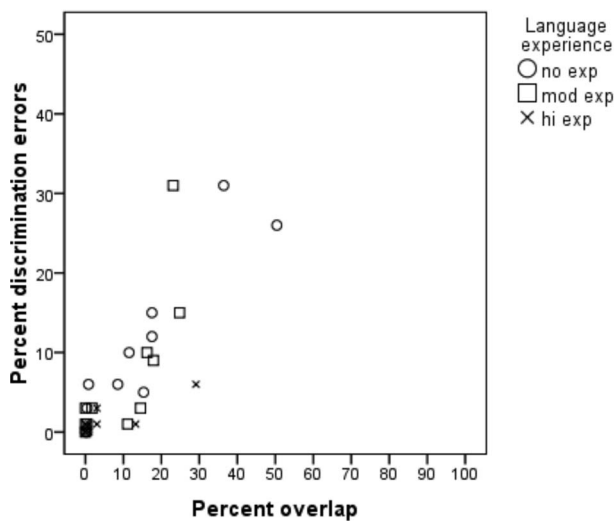
Overlap scores were tallied for the remaining experimental vowel pairs in each consonantal context within each language group and then ranked from lowest to highest. Finally, the discrimination error scores associated with each vowel pair were correlated with total overlap score for each pair. (In the above example, the NoExp group made 5% discrimination errors for /y-u/ in bilabial context, which was compared with the percentage overlap score of 15.3% for that pair.) Nonparametric correlations (Spearman rank order) were performed because the perceptual assimilation results could not be considered interval measures. The higher the overlap score, the higher the percent errors were expected to be revealed in discrimination. Thus, when overlap scores for each pair were ranked from lowest to highest, discrimination error results were also predicted to be ordered from lowest to highest.

## 2. Results

The Appendix lists the cross-language assimilation overlap score and the categorial discrimination percent errors for each language experience group, arranged by overlap score (in ascending order) for each group in each consonantal context. A pattern of more discrimination errors with higher overlap scores is evident, with contrasts involving front rounded vowels paired with back vowels and with each other generally revealing more overlap and more discrimination errors than the other pairs. For example, for NoExp listeners in alveolar context, the scores ranged from 65.9% overlap and 39% discrimination errors (for PF /œ-o/) to 0% overlap and 0% discrimination errors (for PF /œ-i/).

Figure 4 graphs the correlation between cross-language assimilation overlap and discrimination performance for vowel pairs in bilabial /rabVp/ context (A) and in alveolar /radVt/ context (B). Along the x-axis are the cross-language assimilation overlap scores, while the y-axis represents the percent errors in discrimination (up to chance of 50%). Each point on the graph represents a group's response to a particular vowel pair. Data for the NoExp Group are represented as Xs, whereas data for the ModExp group are represented by squares, and for the HiExp group, by circles. As shown in

## A. Bilabial context



## B. Alveolar context

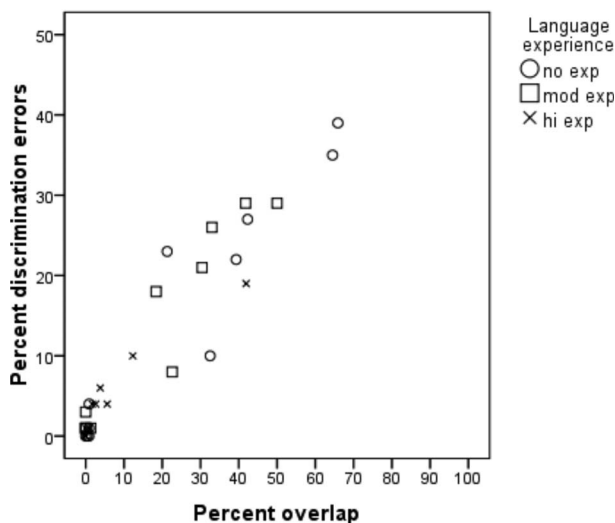


FIG. 4. Scatterplot of relationship between cross-language assimilation overlap patterns and percent errors in categorial discrimination in bilabial /rabVp/ context (a) and in alveolar /radVt/ context (b) by AE listeners with no French experience (no exp), moderate French experience (mod exp), and extensive French experience (hi exp) with vowel pairs /y-u/, /æ-o/, /y-o/, /æ-u/, /y-i/, /y-e/, /æ-e/, /æ-i/, /a-e/, /u-i/, and /y-æ/ as sampling variables.

Fig. 4(A), for the NoExp group, as perceptual overlap increased in bilabial context, so did discrimination errors. A Spearman rank order correlation confirmed a strong correlation between overlap scores and discrimination errors ( $\rho = 0.92$ ,  $p < 0.001$ ). Thus, for this naïve group of listeners, perceptual assimilation patterns were highly predictive of discrimination performance on French vowel contrasts in bilabial context, as posited by the PAM (Best, 1995).

For the ModExp group, the correlation for bilabial context data was also statistically significant ( $\rho = 0.84$ ,  $p = 0.001$ ). Thus, for these L2 learners with formal French instruction, the PAM (Best, 1995) and PAM-L2 (Best and Tyler, 2007) successfully predicted relative vowel discrimination difficulty. Results were also significant for the HiExp group ( $\rho = 0.68$ ,  $p < 0.05$ ). However, this correlation is not particularly

informative, as there were so few errors to interpret for that group. As the figure shows, most of the Xs representing the HiExp group cluster and overlap below 6% in perceptual assimilation and below 4% errors in discrimination, with only one outlying vowel pair, PF /y-u/, revealing higher overlap (29%) and discrimination error (6%) scores.

Turning to alveolar context, Fig. 4(B) reveals more variability in discrimination errors and perceptual assimilation overlap (i.e., larger spread) for all groups than was seen in bilabial context, reflecting the perceptual difficulties encountered by AE listeners in this context. With more errors to work with, the correlation coefficients were higher than for the bilabial context comparison for the three language experience groups (NoExp  $\rho = 0.96$ ,  $p < 0.001$ ; ModExp,  $\rho = 0.93$ ,  $p < 0.001$ ; and HiExp  $\rho = 0.95$ ,  $p < 0.001$ ). Thus, in support of the first hypothesis, the PAM (Best, 1995) and PAM-L2 (Best and Tyler, 2007) predicted relative accuracy in discrimination of vowel pairs from their assimilation patterns, not only for naïve learners of a language, but also for intermediate and advanced adult L2 learners.

## B. Cross-language assimilation overlap by individuals

### 1. Data analysis

A further correlational analysis was conducted in order to test hypothesis 2, that individuals' difficulty in discriminating a vowel pair would be related to their cross-language assimilation overlap of that vowel pair's members. As opposed to the previous analysis in which vowel pairs were ranked according to their group overlap scores, in this analysis, individuals' overlap scores in each consonantal context for PF /y-u/, /æ-u/, /y-æ/, and /æ-o/ were ranked and compared to their discrimination scores for each of these contrasts. These vowel pairs were chosen as these were the contrasts involving front rounded vowels that continued to pose the most difficulties discrimination (i.e.,  $\geq 10\%$  discrimination errors in at least one consonantal context by ModExp) despite language experience, presenting an opportunity to test the PAM (Best, 1995) quantitatively for individual L2 learners (Best and Tyler, 2007).

For this analysis, each listener's perceptual overlap score was tallied for each vowel contrast, defined here as the percent of times that, given a particular vowel contrast, the listener perceptually assimilated both vowel pair members to the same native phone. For example, on the /æ-o/ contrast in alveolar context, a ModExp listener perceptually assimilated PF /æ/ to AE /o, u, ʌ/ categories on 22%, 39%, 28%, and 11%, of trials, respectively. This listener assimilated PF /o/ to AE /o, u, ʌ/ on 67%, 17%, and 17% of trials, respectively. To calculate the overlap score for this listener for this vowel pair in alveolar context, the smaller percentages when both PF speech sounds assimilated to the same AE category (/o/ = 22%, /u/ = 17%, /ʌ/ = 17%, and /ʌ/ = 0%) were summed (=56%). The overlap score was compared to that listener's discrimination score for the same vowel pair in the same consonantal context (in this case 42% errors).

## 2. Results

A Spearman rank order correlation confirmed a correlation between individuals' overlap scores and discrimination errors for PF /y-u/, /æ-u/, /y-æ/, and /æ-o/ ( $\rho=0.68$ ,  $p < 0.001$ , two-tailed) with all groups and both consonantal contexts included. Correlations were statistically significant for each language experience group when consonantal contexts were combined (NoExp  $\rho=0.67$ ,  $p < 0.001$ ; ModExp  $\rho=0.46$ ,  $p < 0.001$ ; and HiExp  $\rho=0.58$ ,  $p < 0.001$ ) and for each consonantal context when language experience groups were combined (bilabial  $\rho=0.60$ ,  $p < 0.001$ ; alveolar  $\rho=0.69$ ,  $p < 0.001$ ). Results indicated that the more often (i.e., the more trials in which) individuals assimilated two members of a PF vowel pair to a single native category, the more discrimination errors they incurred for that vowel pair, providing quantitative support for the PAM (Best, 1995) and its extension to L2 learners (Best and Tyler, 2007).

## C. Discussion

### 1. Quantifying perceptual assimilation patterns

The cross-language assimilation overlap method provided a measure of the accuracy of PAM (Best, 1995) and PAM-L2 (Best and Tyler, 2007) predictions. This method examined the frequency with which two members of a vowel pair both assimilated to a particular native category and compared that frequency to the same group's or individual's discrimination accuracy for the pair in question. This method revealed that, generally, the more often L2 vowels in a pair were assimilated to the same native category by a particular group or individual, the less accurately the contrast was discriminated by that group or individual, results that support the PAM's position that discriminability is predictable from assimilation patterns.

This method does not necessarily differentiate categorizable-uncategorizable patterns from two-category (or from uncategorizable-uncategorizable) assimilation patterns in that it examines only whether one vowel in a pair assimilated to the same native category as did the other vowel. Moreover, it does not capture within-category differences. For example, it does not factor in the goodness ratings that differentiate single-category from category-goodness assimilation patterns. However, as listeners used a limited range of ratings in Levy (2009), and as the assimilation types were not self-evident in this study, the absence of ratings information was not expected to affect the findings meaningfully and the lack of reliance on assimilation types may have been an advantage of the quantification method for this study.

### 2. Limitations

A limitation of the study is that, although cross-speaker tasks were implemented, stimuli were uttered by only three native PF speakers; thus, the overlap and discrimination scores obtained from the listeners' responses may not be generalizable. However, as the predictions were tested based on assimilation and discrimination responses to the same data set, replications of this study are expected to yield different scores, but similar relationships between cross-language assimilation overlap and discrimination accuracy.

It should also be noted that response choice alternatives in an assimilation task may affect response patterns (in this quantification and in more traditional qualitative methods) and, thus, the resulting correlations with discrimination results. For example, had AE /<sup>h</sup>u/ not been a response alternative in Levy (2009), listeners might have assimilated more PF /y/ stimuli to AE /u/ than they did with the palatalized option, resulting in greater overlap in for the PF /y-u/ pair.

And finally, despite the significant correlations found in the present study, categorization models may not capture the complexity of non-native listeners' perceptual sensitivity, as demonstrated by Iverson *et al.*'s (2008) study of the categorization of English /w-v/ by native speakers of Sinhala, German, and Dutch speakers. Native speakers of Sinhala and German have one native phoneme similar to English /w/ and /v/, yet German speakers discerned the English /w-v/ distinction more successfully. Listeners were clearly sensitive to distinctions that were not necessarily reflected in their categorization patterns. Iverson *et al.* (2008) suggested that distortions in perceptual space also contribute to L2 learning.

## IV. GENERAL DISCUSSION

Examining AE listeners' overlap in perceptual assimilation of non-native and L2 PF vowels in relation to the same listeners' discrimination errors yielded significant correlations in bilabial and alveolar contexts. These findings provide preliminary support for predictions generated by the PAM (Best, 1995) to be extended to the domain of listeners in the more advanced stages of L2 learning, as proposed by Best and Tyler (2007). As this was the first study to test the PAM's predictions on vowel learning using a quantitative measure, replication of such studies using the cross-language assimilation overlap and other methods is needed to support this conclusion.

Findings from the reported discrimination experiment and other studies (e.g., Gottfried, 1984; Levy, 2009; Levy and Strange, 2008; Strange *et al.*, 2001, 2009) suggest that contextual variation in the phonetic realization of vowels across languages impacts L2 vowel learning. Levy's (2009) assimilation study indicates that PF /y/ will be perceived as most similar to AE /u/ more often in alveolar context than in bilabial context, leading to more PF /y-u/ discrimination difficulty in alveolar context (with more assimilation overlap) than in bilabial context, as found in the present study. Similarly, when surrounded by bilabials, PF /æ/ is likely to be perceived as more similar to AE /u/, whereas in alveolar context, PF /æ/ is likely to be assimilated to AE /u/. The PAM, taking consonantal context into consideration, would thus predict PF /æ/ to be more difficult to differentiate from PF /u/ in alveolar context (in a single-category assimilation pattern) than in bilabial context (in a two-category assimilation pattern), a prediction supported in the present study. As listeners assimilate PF /æ/ less to AE /u/ and more to other AE vowels (e.g., /ʊ/ and /ɜ:/), contrasts involving this vowel will assimilate in a two-category pattern, incurring less overlap, and discrimination accuracy is predicted to increase. However, discrimination may still be less accurate in alveo-

lar context than in bilabial, even for highly experienced L2 learners. These predictions, too, were borne out in the present study.

Consequences for speech production are addressed by [Flege's \(1995\) SLM](#), which posits that when L2 segments are encountered, they are classified as "identical," "similar," or "new," relative to the listener's native phonological inventory. Viewed from this perspective, the PF vowel /u/ is classified as a similar vowel by AE speakers, resulting in inaccurate pronunciation ([Flege, 1987](#)). [Flege \(1987\)](#) posited that PF /y/ is categorized as a new vowel, although it might be confused with PF /u/ in the initial stages of speech learning. With L2 experience, individuals learn to distinguish PF /y/ from AE /u/, as a new category is established; thus, /y/ may be produced in a near-native manner. Results from the present study support [Flege's \(1987\)](#) claim in bilabial, but not in alveolar, context. The context-specific categorization patterns found in this and [Levy and Strange \(2008\)](#), as well as in [Levy's \(2009\)](#) assimilation study, suggest an allophonic level of representation in equivalence classification, wherein the consonantal context may determine whether listeners perceive a vowel as new or similar. Thus, PF /y/ may be similar to AE /u/ in alveolar context and new in bilabial context.

Perceptual training protocols that take consonantal context into consideration might better assess listeners' perceptual difficulties with vowels and gain effectiveness by targeting those contexts in which listeners have the most difficulty. These measures might help determine whether stubborn contrasts, such as PF /y-u/ for AE listeners, may ever be mastered. In training L2 learners, the cross-language assimilation overlap method may provide fruitful information for mapping the L2 sounds onto the learners' native categories in more productive ways. That is, examining the assimilation patterns that are associated with more accurate discrimination may lead to training protocols in which, for example, the similarities between PF /æ/ and AE /ɜ/ are emphasized to AE learners of French in the hopes that such training will help them along the steep learning curve away from /æ-u/ confusion. Furthermore, studies may ask whether perceptual training alone will result not only in improved perceptual skills, but also in more intelligible production, as has been shown in a handful of studies of Japanese listeners' perceptual training on AE consonants ([Bradlow et al., 1997, 1999](#)) and AE vowels ([Strange and Akahane-Yamada, 1997](#)), as well as in perceptual training for children with phonological disorders ([Rvachew, 1994](#)). Taking into account the complex contextual variability that exists in individuals' languages is expected to result in more beneficial assimilation patterns and more accurate discrimination and comprehension as the individuals learn an L2.

## ACKNOWLEDGMENTS

This research was supported by a grant to the author (NIH-NIDCD Grant No. 1F31DC006530-01). The author is indebted to Winifred Strange. Great thanks also to Allard Jongman, Geoffrey Stewart Morrison, and an anonymous reviewer for their constructive comments. Many thanks also to Loraine Obler, Martin Gitterman, Catherine Best, James Jen-

kins, Kanae Nishi, Valeriy Shafiro, Franzo Law II, Natalia Martínez, Gary Chant, Bruno Tagliaferri, Victoria Hatzelis, Ana de la Iglesia, and the Teachers College Speech Production and Perception Laboratory, for their helpful contributions.

## APPENDIX: OVERLAP AND DISCRIMINATION SCORES

Cross-language assimilation overlap (Overlap) scores and categorical discrimination (CD) percent errors for vowel pairs in /rabVp/ and /radVt/ contexts by AE listeners with no (NoExp), moderate (ModExp), and extensive French experiences (HiExp).

	Vowel pair bp context			Vowel pair dt context		
	Overlap	CD % errors		Overlap	CD % errors	
NoExp	/æ-i/	0	0	/æ-i/	0	0
	/u-i/	0.4	0	/æ-ɛ/	0.4	0
	/y-ɛ/	0.4	3	/y-ɛ/	0.8	0
	/æ-ɛ/	0.8	6	/u-i/	0.8	1
	/y-o/	8.5	6	/y-i/	0.9	4
	/y-i/	11.5	10	/a-ɛ/	21.3	23
	/y-u/	15.3	5	/y-o/	32.5	10
	/y-æ/	17.5	12	/y-æ/	39.3	22
	/a-ɛ/	17.5	15	/y-u/	42.3	27
	/æ-u/	36.4	31	/æ-u/	64.5	35
ModExp	/æ-o/	50.4	26	/æ-o/	65.9	39
	/u-i/	0	0	/u-i/	0	1
	/y-ɛ/	0	1	/y-ɛ/	0	1
	/æ-i/	0	3	/æ-ɛ/	0	1
	/y-i/	0.4	1	/æ-i/	0	3
	/æ-ɛ/	1.7	3	/y-i/	1.3	1
	/a-ɛ/	11.1	1	/a-ɛ/	18.4	18
	/y-o/	14.5	3	/y-o/	22.6	8
	/y-æ/	16.2	10	/y-æ/	30.4	21
	/y-u/	17.9	9	/æ-u/	33.0	26
HiExp	/æ-u/	23.1	31	/y-u/	41.8	29
	/æ-o/	24.8	15	/æ-o/	50.0	29
	/y-i/	0	0	/y-i/	0	0
	/æ-i/	0	0	/æ-i/	0	1
	/u-i/	0	0	/u-i/	0	0
	/æ-ɛ/	0	1	/y-ɛ/	0	0
	/y-ɛ/	0	1	/æ-ɛ/	0.9	1
	/a-ɛ/	0.9	0	/y-o/	1.7	4
	/y-o/	0.9	1	/a-ɛ/	2.6	4
	/æ-o/	3.0	1	/æ-o/	3.8	6
/y-æ/	3.0	3	/y-æ/	5.6	4	
/æ-u/	13.2	1	/æ-u/	12.3	10	
/y-u/	29.1	6	/y-u/	41.9	19	

<sup>1</sup>Front rounded PF /ø/ and /œ/ are rarely contrastive in PF. For the purposes of this paper, /œ/ represents the mid front rounded vowel.

<sup>2</sup>In a categorial, i.e., cross-speaker, task, the speakers differ across the three stimuli. The listeners must thus make decisions based on speaker-independent categories (Beddor and Gottfried, 1995).

<sup>3</sup>Except for one HiExp listener, participants in the present experiment (and in Levy, 2009) differed from those in Levy and Strange, 2008. As Levy and Strange's (2008) study had taken place 3 years prior to the present experiment, no learning effect was expected.

<sup>4</sup>The term "counterpart" is used loosely here to denote a speech sound for which the other language has a speech sound that is transcribed identically in broad phonetic transcription. It is acknowledged that similarly transcribed sounds may differ in their distributions of acoustic properties and that speech perception models cannot yet predict how non-native speech sounds will be mapped onto native categories (Harnsberger, 2001).

<sup>5</sup>A table detailing the computation of overlap is available via e-mail from the author.

- Beddor, P. S., and Gottfried, T. L. (1995). "Methodological issues in cross-language speech perception research with adults," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 207–232.
- Beddor, P. S., Harnsberger, J. D., and Lindemann, S. (2002). "Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates," *J. Phonetics* **30**, 591–627.
- Best, C. T. (1995). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 171–204.
- Best, C. T., Faber, A., and Levitt, A. (1996). "Assimilation of non-native vowel contrasts to the American English vowel system," *J. Acoust. Soc. Am.* **99**, 2602.
- Best, C. T., McRoberts, G. W., and Goodell, N. M. (2001). "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," *J. Acoust. Soc. Am.* **109**, 775–794.
- Best, C. T., McRoberts, G. W., and Sithole, N. M. (1988). "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants," *J. Exp. Psychol.* **14**, 345–360.
- Best, C. T., and Strange, W. (1992). "Effects of phonological and phonetic factors on cross-language perception of approximants," *J. Phonetics* **20**, 305–330.
- Best, C. T., and Tyler, M. D. (2007). "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, edited by O.-S. Bohn and M. J. Munro (John Benjamins, Amsterdam), pp. 13–34.
- Bohn, O.-S., and Steinlen, A. K. (2003). "Consonantal context affects cross-language perception of vowels," Proceedings of the 15th International Congress of Phonetic Sciences, pp. 2289–2292.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1997). "Training Japanese listeners to identify English /r/ and /l/: Some effects of perceptual learning on speech production," *J. Acoust. Soc. Am.* **101**, 2299–2310.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1999). "Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production," *Percept. Psychophys.* **61**, 977–985.
- Clopper, C. G., Pisoni, D. B., and de Jong, K. (2005). "Acoustic characteristics of the vowel systems of six regional varieties of American English," *J. Acoust. Soc. Am.* **118**, 1661–1676.
- Flege, J. E. (1987). "The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification," *J. Phonetics* **15**, 47–65.
- Flege, J. E. (1995). "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 233–277.
- Flege, J. E., and Hillenbrand, J. (1984). "Limits on phonetic accuracy in foreign language speech production," *J. Acoust. Soc. Am.* **76**, 708–721.
- Gottfried, T. L. (1984). "Effects of consonant context on the perception of French vowels," *J. Phonetics* **12**, 91–114.
- Guion, S. G., Flege, J. E., Akahane-Yamada, R., and Pruitt, J. C. (2000). "An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants," *J. Acoust. Soc. Am.* **107**, 2711–2724.
- Harnsberger, J. D. (2001). "On the relationship between identification and discrimination of non-native nasal consonants," *J. Acoust. Soc. Am.* **110**, 489–503.
- Hillenbrand, J. M., Clark, M. J., and Nearey, T. M. (2001). "Effects of consonant environment on vowel formant patterns," *J. Acoust. Soc. Am.* **109**, 748–763.
- Iverson, P., Ekanayake, D., Hamann, S., Sennema, A., and Evans, B. G. (2008). "Category and perceptual interference in second-language phoneme learning: An examination of English /w/-v/ learning by Sinhala, German, and Dutch speakers," *J. Exp. Psychol.* **34**, 1305–1316.
- Kewley-Port, D., Burkle, T. Z., and Lee, J. H. (2007). "Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners," *J. Acoust. Soc. Am.* **122**, 2365–2375.
- Kuhl, P. K., and Iverson, P. (1995). "Linguistic experience and the 'perceptual magnet effect,'" in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 121–154.
- Levy, E. S. (2009). "Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French," *J. Acoust. Soc. Am.* **125**, 1138–1152.
- Levy, E. S., and Law, F., II (2008). "Production of Parisian French front rounded vowels by second-language learners," *J. Acoust. Soc. Am.* **123**, 3078.
- Levy, E. S., and Strange, W. (2008). "Perception of French vowels by American English adults with and without French language experience," *J. Phonetics* **36**, 141–157.
- Manuel, S. Y. (1999). "Cross-language studies: Relating language-particular patterns to other language-particular facts," in *Coarticulation: Theory, Data and Techniques*, edited by W. J. Hardcastle and N. Hewlett (Cambridge University Press, New York), pp. 179–198.
- Oh, E. (2008). "Coarticulation in non-native speakers of English and French: An acoustic study," *J. Phonetics* **36**, 361–384.
- Polka, L. (1995). "Linguistic influences in adult perception of non-native vowel contrasts," *J. Acoust. Soc. Am.* **97**, 1286–1296.
- Polka, L., and Bohn, O.-S. (1996). "A cross-language comparison of vowel perception in English-learning and German-learning infants," *J. Acoust. Soc. Am.* **100**, 577–592.
- Rochet, B. L. (1995). "Perception and production of second-language speech sounds by adults," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Timonium, MD), pp. 379–410.
- Rvachew, S. (1994). "Speech perception training can facilitate sound production learning," *J. Speech Hear. Res.* **37**, 347–57.
- Stevens, K. N., Liberman, A. M., Studdert-Kennedy, M., and Öhman, S. (1969). "Cross-language study of vowel perception," *Lang Speech* **12**, 1–23.
- Strange, W., and Akahane-Yamada, R. (1997). "Effects of identification training on Japanese adults' perception of American English vowels," *J. Acoust. Soc. Am.* **102**, 3137.
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., and Nishi, K. (2001). "Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners," *J. Acoust. Soc. Am.* **109**, 1691–1704.
- Strange, W., Bohn, O.-S., Nishi, K., and Trent, S. A. (2005). "Contextual variation in the acoustic and perceptual similarity of North German and American English vowels," *J. Acoust. Soc. Am.* **118**, 1751–1762.
- Strange, W., Levy, E. S., and Law, F. F., II (2009). "Cross-language categorization of French and German vowels by naïve American listeners," *J. Acoust. Soc. Am.* **126**(3), 1461–1476.
- Strange, W., and Shafer, V. L. (2008). "Speech perception in second language learners: The reeducation of selective perception," in *Phonology and Second Language Acquisition*, edited by J. G. Hansen Edwards and M. L. Zampini (John Benjamins, Amsterdam), pp. 153–191.
- Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., and Nishi, K. (2007). "Acoustic variability within and across German, French and American English vowels: Phonetic context effects," *J. Acoust. Soc. Am.* **122**, 1111–1129.
- Tranel, B. (1987). *The Sounds of French: An Introduction* (Cambridge University Press, New York).