

*This is the peer reviewed version of the following article:*

**Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*, 66(4), 741-773.**

*which has been published in final form at <http://dx.doi.org/10.1111/lang.12161>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.*

## **EMPIRICAL STUDY**

### **The Role of Inhibitory Control in Second Language Phonological Processing**

Isabelle Darcy,<sup>a</sup> Joan C. Mora,<sup>b</sup> and Danielle Daidone<sup>a</sup>

<sup>a</sup>Indiana University and <sup>b</sup>University of Barcelona

This study investigated the role of inhibition in second language (L2) learners' phonological processing. Participants were Spanish learners of L2 English and American learners of L2 Spanish. We measured inhibition through a retrieval-induced inhibition task. Accuracy of phonological representations (perception and production) was assessed through a speeded ABX categorization task and a delayed sentence repetition task. We used a measure of L2 vocabulary size to tease out L2 proficiency effects. Higher inhibitory control was related to lower error rate in segmental perception. Inhibition was also related to consonant but not to vowel production

accuracy. These results suggest a potential role for inhibition in L2 phonological acquisition, with inhibition enhancing the processing of phonologically relevant acoustic information in the L2 input, which in turn might lead to more accurate L2 phonological representations.

**Keywords** inhibitory control; retrieval induced inhibition; vowel and consonant production; vowel perception; English; Spanish

### **Author Note**

This research would not have been possible without the help of many people. We are indebted to Shiri Lev Ari and Sharon Peperkamp (LSCP, Paris) for sharing the E-prime script of the inhibition task with us. We also thank the speakers who recorded the stimuli: Paola Rodrigues, Tanya Flores, Diana Arroyo, Ana Fernandez, Maggie Peters, and Fiona Pannett. We thank Amanda Rabideau (University of Utah), Elena Safronova (Universitat de Barcelona), and Eva Cerviño-Povedano (Universitat de Barcelona) for help with testing and data processing in Bloomington and Barcelona. For their decisive help in making the testing in Seville possible, we thank Marina Barrio Parra and M. Heliadora Cuenca Villarín, as well as Ron Roosevelt (Universidad de Sevilla). For their institutional, financial, and logistical support, we are indebted to Carmen Muñoz (Universitat de Barcelona) and Kathleen Bardovi-Harlig (Indiana University, Bloomington). We also thank Stephanie Dickinson from the Indiana University Statistical Consulting Center for her help with statistical analyses. For stimulating discussions, we thank the members of the Second Language Psycholinguistics Laboratory (Indiana University, Bloomington), especially Ryan Lidster and Jeffrey Holliday, and the audience at New Sounds 2013 in Montréal. We also thank Pavel Trofimovich and the *Language Learning* editorial team,

and three anonymous reviewers for their excellent suggestions. We further acknowledge grant support: Grant-in-Aid, Indiana University Bloomington; Grants FFI2013-47616-P (Ministerio de Economía y Competitividad) and 2014SGR1089 (Generalitat de Catalunya).



This article has been awarded an Open Materials badge. All materials are publicly accessible in the IRIS digital repository at <http://www.iris-database.org>. Learn more about the Open Practices badges from the Center for Open Science: <https://osf.io/tvyxz/wiki>

Correspondence concerning this article should be addressed to Isabelle Darcy, Department of Second Language Studies, Indiana University, Memorial Hall 301, 1021 E. Third St, Bloomington, IN 47405, USA. E-mail: [idarcy@indiana.edu](mailto:idarcy@indiana.edu)

## Introduction

In the foreign language classroom, students often struggle with pronunciation and listening, and the challenge of acquiring the phonological system of a second language (L2) is met with variable outcomes. Certain learning conditions are more favorable than others, such as longer and more intensive exposure to L2 input (e.g., Flege, Yeni-Komshian, & Liu, 1999) and more frequent use of the L2 (e.g., Guion, Flege, & Loftin, 2000). Students' first language (L1) background also plays a role in L2 acquisition (e.g., Flege, Bohn, & Jang, 1997). Yet, even when these variables are kept identical, such as in the case of a homogenous student population, large individual differences remain in production and perception (e.g., Pallier, Bosch, & Sebastián-Gallés, 1997). Factors underlying individual differences examined in previous research include, among others, motivation (Moyer, 1999), singing talent and musical ability (Christiner & Reiterer, 2013; Slevc & Miyake, 2006), sound processing ability (Golestani & Zatorre, 2009; Lengeris & Hazan, 2010), and cognitive skills, such as working memory (e.g., Reiterer et al.,

2011), phonological short-term memory (Cerviño-Povedano & Mora, 2011; Darcy, Park, & Yang, 2015; MacKay, Meador, & Flege, 2001), attention (Darcy, Mora, & Daidone, 2014; Guion & Pederson, 2007; Safronova & Mora, 2013), and inhibition (e.g., Lev-Ari & Peperkamp, 2013).

Although some of these cognitive factors have been shown to be globally related to bilingual speech processing, it is still not well understood how they specifically relate to the perception and production of L2 segments (vowels and consonants) by instructed adult L2 learners. The goal of this study was to further explore this relationship, focusing on inhibitory control, because stronger inhibitory control could reduce interference from the L1, and thus enhance the processing of acoustic-phonetic information in the input. This in turn could lead to more accurate L2 phonological representations, as reflected in the more accurate perception and production of L2 segments. Late L2 learners ( $N = 34$ ) from two different L1 backgrounds took part in an inhibitory control task, and in tasks of perception and production of various L2 contrasts, to examine whether individual performance in one measure relates to performance in the other.

### **Inhibitory Control and Language Processing**

Inhibition is an important executive function encompassing a variety of response-related selection processes that involve multiple inhibitory systems (see Kok, 1999, and Nigg, 2000, for taxonomies of inhibitory processes). For example, Friedman and Miyake (2004) identified three main inhibition-related functions: (a) the ability to suppress a dominant, automatic, prepotent response (Prepotent Response Inhibition), (b) the ability to resolve interference from information irrelevant to the task at hand (Resistance to Distractor Interference), and (c) the ability to resist interference from information that was previously relevant to the task but has since become irrelevant (Resistance to Proactive Interference). An important distinction to be made between

these three inhibitory control processes is that the first two generally involve relatively conscious, effortful control processes that lead to the active suppression of a response, whereas the third results directly from language processing (see also Miyake et al., 2000). Arguably, all three are involved in L2 acquisition at various stages and at various levels of processing. More importantly for the purposes of the present study, Miyake and Friedman's (2012) more recent findings suggest that individual differences in executive functions, such as *updating* (monitoring and updating of working memory contents) and *shifting* (switching between mental sets), reflect individual differences in inhibitory control, thus underscoring the potential of inhibition for subsuming general individual differences in executive control.

In bilingual research, inhibition has been proposed as the cognitive control mechanism responsible for preventing the selection of nontarget language words during speech production (Green, 1998). This is achieved by bilinguals inhibiting the activation of nontarget lexical representations when speaking the selected target language, so that inefficient inhibitory control would lead to interference from the language not in use (e.g., Spivey & Marian, 1999). Several studies have tested the consequences of cross-language interference at the phonetic level within a language switching paradigm for bilingual speakers. A consistent finding in this research is that L2-to-L1 and L1-to-L2 interference in phonetic production (e.g., voice onset time [VOT] in the production of oral stops) occurs when switching between languages and is modulated by the speakers' degree of bilingual experience and proficiency in the test languages (Filippi, Karaminis, & Thomas, 2014; Goldrick, Runnqvist, & Costa, 2014; Olson, 2013). Similarly, Filippi, Leech, Thomas, Green, and Dick (2012) simultaneously presented two sentences (one in Italian and one in English) to Italian-English bilinguals and English controls and found that bilinguals were better able to resist sentence-level interference during comprehension.

Inhibition can also result effortlessly from lexical processing itself by inducing a decrease in activation levels through lexical retrieval. Such retrieval-induced inhibition tasks (Anderson, R. Bjork, & E. Bjork, 1994; Perfect, Moulin, Conway, & Perry, 2002; Veling & Knippenberg, 2004) have shown that retrieving items from memory (e.g., *apple*) belonging to a given lexical category (e.g., *fruit*) causes inhibition (i.e., suppression of activation) of nonretrieved items (e.g., *pear*) in the same category, leading to the participant's inability to recall an inhibited item (forgetting) or to slower retrieval. Interestingly, retrieval-induced inhibition has also been shown to work across languages, modulating cross-language interference and inhibiting learners' L1 phonology. For example, Levy, McVeigh, Marful, and Anderson (2007) found that the more often L1 English learners of L2 Spanish repeatedly named objects in Spanish (e.g., repeating *culebra* "snake" one, five, or 10 times), the harder it became for them to recall the corresponding English name (*snake*) when cued by a rhyming English word (e.g., *break*).

Individuals, whether fully functional bilinguals in two or more languages or less proficient late L2 learners, may vary in their inhibitory control capacity. Several studies have shown that individual differences in inhibitory control explain variations in performance on language processing tasks. For example, bilinguals with more efficient inhibitory control exhibit less cross-language competition in picture naming tasks (Linck, Hoshino, & Kroll, 2008) and during spoken bilingual word recognition, compared to bilinguals with less efficient control (Mercier, Pivneva, & Titone, 2014). Bilinguals with better control also obtain shorter switch costs when shifting between languages (Linck, Schwieter, & Sunderman, 2012), and they are more efficient in planning and producing L2 speech irrespective of L2 proficiency (Pivneva, Palmer, & Titone, 2012).

Importantly for the current study, inhibitory control has also been linked to the amount of influence that the phonology in bilinguals' one language has on the phonology in their other language. Lev-Ari and Peperkamp (2013) demonstrated a relationship between inhibition and L1 phonological attrition, such that speakers with higher inhibitory control scores showed less L2 influence in their L1 VOT perception and production. The researchers employed a retrieval-induced inhibition task to measure the inhibitory skill of L1-dominant English-French bilinguals who resided in France and used both English and French daily. The lower the speakers' inhibitory skill was, the more they produced and perceived VOT in English voiceless stops (/p, t, k/) in a French-like manner; that is, greater inhibitory skill allowed speakers to avoid L2 use effects in their L1.

Although Lev-Ari and Peperkamp (2013) investigated inhibitory control and its relationship to L1 phonology, their findings have important implications for L2 phonological processing. It is possible that strong inhibitory control could also minimize the effects of L1 phonology on L2 phonology. Put differently, being able to suppress the L1 more robustly could help L2 users reduce interference from their L1 segment categories during L2 use, thus allowing for more accurate perception and production of L2 segments. To our knowledge, no study thus far has shown a link between inhibitory control and perception or production of difficult L2 phonological contrasts.

### **The Present Study**

Given that inhibitory control has been linked to L2 acquisition and processing, but the relationship between L2 speakers' inhibitory control ability and their L2 phonological processing is still largely unknown, we set out to examine the relationship between the strength of L2 learners' inhibitory control and their accuracy in perceiving and producing L2 segments. We

hypothesized that stronger inhibitory control would correlate with more accurate L2 perception and production, since learners who are better able to suppress their L1 may have less influence from L1 phonological categories when speaking and listening to the L2. We also obtained measures for several demographic variables (e.g., age of first exposure to the L2, length of residence abroad, current L2 use) for all participants in order to facilitate analyses of individual differences.

To obtain generalizable findings regarding inhibition and phonological processing that are independent of the specific experimental items, we tested two groups of late adult L2 learners differing in their L1s (L1 English learners of Spanish and L1 Spanish learners of English) using the same stimuli, which contained items in their L1 and L2 (Spanish and English). This bidirectional approach, in which the two groups completed the same tasks in their respective L2, served as an internal control for the stimuli and allowed us to be more confident in the generalizability of our findings across different groups of L2 learners. We chose to measure phonological processing in both perception and production, using an ABX task for perception (Analysis 1) and a delayed sentence repetition task for production (Analysis 2), as described below. We used a retrieval-induced inhibition task to assess learners' inhibitory control and performed analyses of individual learner differences, targeting the relationship between measures of phonological processing and inhibition scores while taking into account learners' proficiency as measured through a vocabulary knowledge test.

### **Participants and Overall Design**

A total of 81 participants were tested: 35 L1 Spanish learners of English in Seville (Spain), 26 L1 English learners of Spanish in Bloomington, Indiana (USA), and 20 controls to provide performance baselines for the tasks (10 native speakers of each L1, tested in the same locations

as the L2 learners). Participants were tested as part of a larger project that included other cognitive measures not reported here.<sup>1</sup> The order of tasks was the same for all participants with slight adjustments where necessary (e.g., only the learners completed L2 vocabulary tasks) and, in general, included the production task first, followed by the cognitive tasks, the perception task, and the vocabulary test at the end. Participants had to pass a pure-tone audiometry test at octave frequencies between 500 and 8,000 Hz at 20 dB HL (Reilly, Troiani, Grossman, & Wingfield, 2007) in order to be included in the analysis. A total of 11 L2 English learners and four L2 Spanish learners were excluded because they had failed the audiometry test. Additionally, one L2 English participant was excluded because he was an early Catalan/Spanish bilingual, and two L2 Spanish participants were excluded because their native speaker (NS) status was unclear or because they reported to have had very early exposure to a language other than English. Further, two NS controls (one English, one Spanish) were also excluded from the analysis, either because their NS status was unclear or because of the presence of a speech/hearing pathology. In total, data from 58 participants were included: 23 L2 English participants, 20 L2 Spanish participants, and 15 NS controls (eight English, seven Spanish). Table 1 summarizes descriptive statistics for the main background variables for each group.

**Table 1** Means (standard deviations in parentheses) for participants' demographic variables

Variable	L2 English ( $n = 22$ )	L2 Spanish ( $n = 20$ )	NSs ( $n = 15$ )
Age at testing	23.1 (4.8)	19.7 (1.0)	21.6 (2.9)
Motivation (1–9)	7.5 (0.7)	7.4 (0.9)	6.3 (1.2)
Current L2 use (max. 36)	16.4 (5.6)	9.1 (6.8)	4.6 (6.6)
Self-rating (1–5)	4.0 (0.5)	4.0 (0.6)	1.9 (0.9)
Residence abroad (weeks)	4.5 (8.8)	10.8 (27.4)	0.3 (1.1)
Years of study	11.7 (2.5)	8.7 (2.8)	4.0 (3.8)
Age of first L2 exposure	8.0 (3.1)	8.5 (4.3)	10.1 (3.5)
Age of first L2 use	12.9 (5.0)	10.6 (3.9)	14.0 (4.6)
Gender (female)	16	17	10
Handedness (left-handed)	3	4	0

A composite motivation score was obtained by averaging each participant's ratings on nine motivation items presented through 9-point Likert scales (1 = *strongly agree*, 9 = *strongly disagree*). The questions asked about the degree with which participants felt motivated to learn or use the L2 (see Appendix S1 in the Supporting Information online); a higher score indicates greater motivation. Current L2 use was determined via an average score (0–36) obtained by adding up participants' selected level of intensity of L2 use (0 = 0%, 1 = 1–25%, 2 = 26–50%, 3 = 51–75%, 4 = 76–100%) in nine contexts of language use (e.g., with friends, at home/work, media). L2 self-evaluation was a self-reported estimate of participants' ability to speak spontaneously, understand, read, and write the L2, using the following descriptions (recoded as numeric scores): very poorly (1), poorly (2), passably (3), well (4), very well (5). A mean self-

evaluation score was obtained by averaging the four scores for each participant. Proficiency was assessed as vocabulary size using the X-Lex test (see below).

## **Analysis 1: Perception**

### **Materials and Procedure**

To assess L2 perceptual processing, we administered a speeded ABX categorization task (e.g., Gottfried, 1984). In each trial, participants heard three stimuli in a row and had to choose if the last token (X) was more similar to the first token (A) or to the second token (B). To increase task demands, the stimuli consisted of trisyllabic nonwords in both Spanish and English with the structure CV.'CV.CV(C), such as [fa'neða] or [fə'ni:dɪf]. Stimuli in both languages were recorded by two female early balanced bilinguals (Mexican Spanish/American English) so that participants could hear both test and control stimuli in the same voice. Furthermore, physically different tokens were used in each trial: One voice was used for the A and B tokens, and the other for the X token. Two different voices were used within each trial to ensure that participants had to make a decision according to the phonological category of the stimuli rather than relying on low-level acoustic traces, as the exact acoustic properties of the stimuli differed across speakers. All participants heard all Spanish and English stimuli in two separate blocks of 64 trials each. The design was such that nonnative contrasts for L2 learners of English were native contrasts for L2 learners of Spanish and vice versa (see Table 2). For example, /i-ɪ/ (as in *feet-fit*), which is phonemically contrastive in English but not in Spanish, was a nonnative contrast for the L2 English group; therefore, the results of the L1 English (L2 Spanish) group for that contrast served as a baseline. All contrasts were produced using appropriate phonetic realizations for each language (see Appendix S2 in the Supporting Information online).

**Table 2** Contrasts used in the ABX task and associated subconditions

Group	Vowel contrasts		Consonant contrasts		Common contrasts	
L2 English	e–e̞	<i>vowelS</i> native	d–ɾ	<i>consS</i> native	a–ɪ	<i>ctrlVS</i>
	i–ɪ	<i>vowelE</i> nonnative	ʃ–tʃ	<i>consE</i> nonnative	t–d	<i>ctrlCS</i>
L2 Spanish	e–e̞	<i>vowelS</i> nonnative	d–ɾ	<i>consS</i> nonnative	a–ɪ	<i>ctrlVE</i>
	i–ɪ	<i>vowelE</i> native	ʃ–tʃ	<i>consE</i> native	t–d	<i>ctrlCE</i>

*Note.* *consE* = English consonant; *vowelE* = English vowel; *consS* = Spanish consonant; *vowelS* = Spanish vowel; *ctrlCE* = control English consonant; *ctrlVE* = control English vowel; *ctrlCS* = control Spanish consonant; *ctrlVS* = control Spanish vowel.

We predicted that group performance would be more accurate for native contrasts over nonnative ones. However, since nativeness covaried with stimulus language (i.e., phonetic realizations were L1 specific), it was important to add a control condition where only stimulus language varied, not nativeness, by choosing contrasts common to both languages. By determining that stimulus language did not influence performance in the control condition, we could attribute differences in performance in the test condition to the nativeness of the contrasts. Conversely, if L1-specific phonetic realization triggered an advantage in performance in the control condition, it would be more difficult to clearly attribute potential accuracy differences to the nativeness of the contrasts.

For L2 Spanish, the perception of nonnative /e–e̞/ (as in *reno-reino* “reindeer-kingdom”) was hypothesized to be difficult for this group of learners based on perceptual mapping data (Morrison, 2006). The perception of the nonnative /d–ɾ/ test contrast (as in *cada-cara* “each-expensive”) has been shown to pose difficulty to low-intermediate learners (Rose, 2010). For L2 English, the nonnative contrast /i–ɪ/, which is subject to developmental stages of acquisition, has previously been found to be perceived either inaccurately or not at all by L2 English learners (e.g., Morrison, 2009). The nonnative /ʃ–tʃ/ contrast (as in *ship-chip*) has been shown to not be

accurately realized in production (e.g., Anrriich, 2007) and was therefore hypothesized to be difficult in perception. In total, four nonword pairs per condition were tested. Each pair was repeated in four combinations (ABA, ABB, BAA, and BAB), yielding a total of 128 trials, 64 for each stimulus language. Trials were assigned to two blocks according to stimulus language (English-Spanish or vice versa), and block order was counterbalanced across participants. Within each block, trials were randomized. If a participant made no response within 2,500 milliseconds, the next trial was initiated. The task was administered on a PC through headphones using the presentation software DMDX (Forster & Forster, 2003).

## **Results**

Results for the 58 participants were screened for outliers by examining individual performance in the control condition. For each group, one learner and one NS whose performance in the control condition was below two standard deviations from their respective group mean were excluded from analyses. This left 22 L2 English learners and six Spanish NSs, as well as 19 L2 Spanish learners and seven English NSs for analysis. The proportion of errors (%) and mean reaction time (RT) for each participant were computed across four trials per item (with four items per condition) for each of the eight subconditions.

### *Preliminary Analyses*

In order to first ascertain whether the L2 learners performed like the NS controls when processing native stimuli, we compared the error rate and RT performance for both L2 learners and NS controls for the same L1 contrasts. For each language (L1 English, L1 Spanish), two mixed-effects models were fitted in SPSS 22 for both error rate and RTs, using group (NS, L2 learner), condition (test, control), and stimulus language (native, nonnative) as fixed-effects factors. Participant was declared as a random factor. For both error rate and RT in each language,

the three-way interaction between the fixed factors was not significant. This suggests that within each language, performance of the two groups was not modulated by the combination of condition and stimulus language. More specifically, for either dependent variable, there was no difference between groups for either test or control condition when the stimuli were native (all  $p > .10$ ) as shown in Table 3.

**Table 3** Estimated mean differences ( $\Delta M$ ) in mean error rates (%) and RTs (ms) between learners and controls for perception of English and Spanish stimuli

Stimuli	Variable	Condition	$\Delta M$	SE	$p$	95% CI
English	Error	Control	-2.9	2.8	.299	-8.5, 2.7
		Test	-1.2	2.8	.659	-6.8, 4.3
	RT	Control	-16.8	90.1	.853	-201.7, 168.1
		Test	-74.9	90.1	.413	-259.8, 110.0
Spanish	Error	Control	0.5	3.2	.870	-5.7, 6.8
		Test	1.9	3.2	.541	-4.3, 8.2
	RT	Control	12.4	116.5	.916	-226.1, 250.9
		Test	-5.4	116.5	.964	-243.8, 233.1

*Note.*  $\Delta M$  = mean difference, SE = standard error, CI = confidence interval.

For the English stimuli, the learners made on average 2.1% fewer errors than the NSs for control and test stimuli; this difference was not significant. The 95% confidence interval (CI) for this estimated difference was similar for both conditions. Similarly, for the Spanish stimuli, the learners made on average 1.2% more errors than the NSs for control and test stimuli, a difference that was also nonsignificant. The CI for this estimated difference was similar for both conditions.

The mean differences in RT (with negative values indicating that the NSs are slower by that amount [in milliseconds] than the learners) suggest that the two groups processed the L1 stimuli with similar RTs. The difference was not significant in any condition, and 95% CIs for the control and test mean differences overlap in both conditions. This result implies that the L2 learners, while processing the L1 stimuli, performed like NSs. The L2 learners' performance on L1 stimuli can therefore be considered an internal control condition, serving as the basis for comparing the two learner groups to each other using the same stimuli. Full descriptive statistics for participants' performance in the perception task are summarized in Appendix S3 in the Supporting Information online.

#### *Overall Error Rate and Reaction Time*

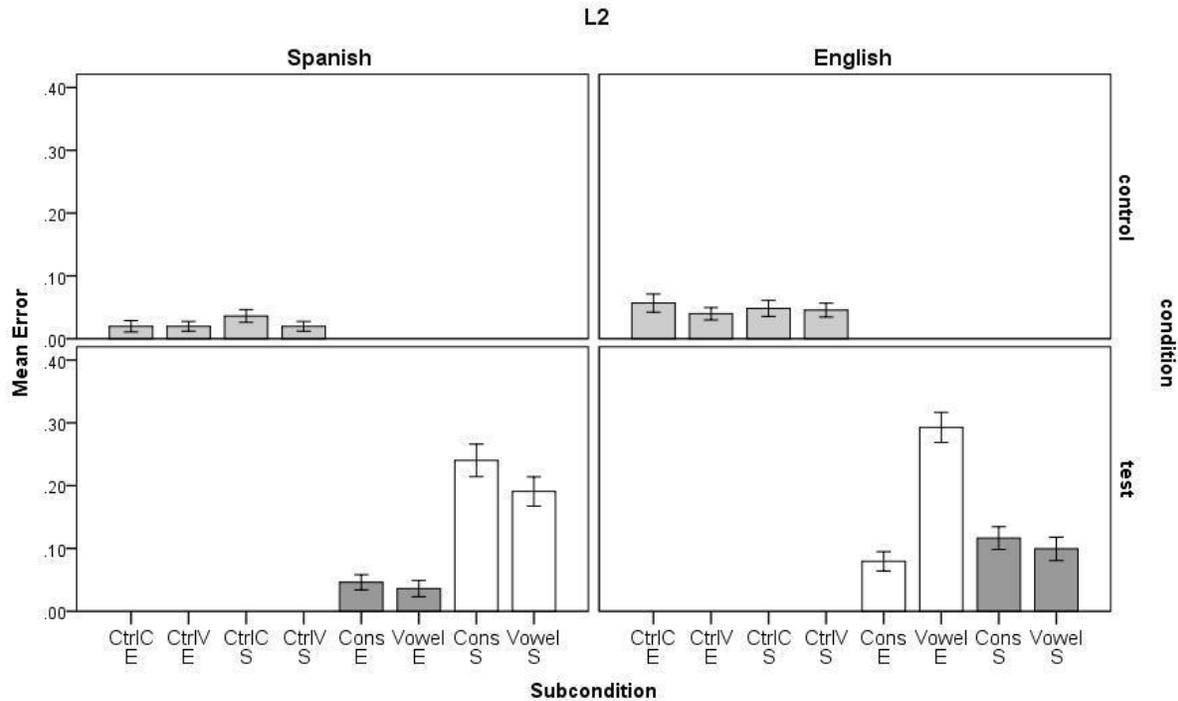
A linear mixed-effects model was fitted for the error and RT patterns. The factors L1 Group (English, Spanish), stimulus language (Spanish, English), and condition (control, test) were used as fixed effects; participants and items were used as random effects. The parameter estimates for each model (error, RT) are presented in Appendix S4 in the Supporting Information online. Type III tests of fixed effects for error rate revealed that there was a significant triple interaction between L1, stimulus language, and condition,  $F(1, 1262) = 58.40, p < 0.001$ . When listening to either language, an effect of L1 was only visible in the test condition, not in the control condition (both  $p > .10$ ). For English test stimuli, English L1 mean error was 4.1%, Spanish L1 mean error was 18.6%, and the mean difference was 14.5 (CI = 10.6–18.4). For Spanish test stimuli, Spanish L1 mean error rate was 10.8%, English L1 rate was 21.5%, and the mean difference was 10.8 (CI = 6.9–14.6, both  $p < .001$ ). Similarly, for both groups, stimulus language significantly affected performance only in the test condition (both  $p < .001$ ), not in the control condition. Put differently, in the test condition only, when listening to the L1 stimuli and native contrasts,

participants made fewer errors than when listening to the L2 stimuli and nonnative contrasts (see Appendix S3).

For RT, type III tests of fixed effects revealed that the three-way interaction between L1, stimulus language, and condition was also significant,  $F(1, 1265) = 15.7, p < .001$ . For both groups, stimulus language had a significant effect on RT in the test condition (both  $p < .001$ ), such that participants were slower for L2 stimuli/nonnative contrasts compared to L1 stimuli. In the control condition, stimulus language did not affect performance in the L2 Spanish group ( $p > .10$ ), but it did affect performance in the L2 English group ( $p < .001$ ). L1 Spanish participants were slower to respond to L2 English stimuli in both the test and control conditions. Reasons for slower RTs in this control condition are unclear, and an analysis of individual performance showed large variability in participants' mean RT in this condition. Because variability did not seem consistent across conditions, we decided against using RTs to examine individual differences. Therefore, all subsequent analyses involve error rate data only.

#### *Error Rate Across Vowel and Consonant Contrasts*

The following analyses compared participants' error rates across separate vowel and consonant contrasts, using a linear mixed-effects model. The factors group (L2 English, L2 Spanish) and subcondition (consE, vowelE, consS, vowelS, ctrlCE, ctrlVE, ctrlCS, ctrlVS, as indicated in Table 2) were modeled as fixed effects. Participant was used as a random effect. Parameter estimates for the model (error) are presented in Appendix S5 in the Supporting Information online, and Figure 1 displays the error rates in each of the eight subconditions.



**Figure 1** Mean proportion error for each learner group in each subcondition (vowels and consonants). White bars represent nonnative contrasts, dark bars represent native contrasts (test condition). Light grey bars in the top panels represent common contrasts (control condition). Error bars enclose  $\pm 1$  SE.

Type III tests of fixed effects for error rates revealed that there was no significant main effect of L1,  $F(1, 39) = 2.30, p > .10$ . Both groups' overall error rate was very similar, and confidence intervals overlapped for L2 Spanish ( $M_{err} = 7.6\%$ ,  $CI = 5.4\text{--}9.8$ ) and L2 English ( $M_{err} = 9.7\%$ ,  $CI = 7.7\text{--}11.8$ ). However, there was a significant main effect of subcondition,  $F(7, 1254) = 35.4, p < .001$ , mainly driven by the four test subconditions featuring nonnative contrasts, for which error rate was much higher overall ( $M_{err} = 13.8\%$ ) than for the control subconditions ( $M_{err} = 3.6\%$ ). The interaction between group and subcondition was also significant,  $F(7, 1254) = 27.4, p < .001$ . Univariate tests revealed that performance of the two groups differed only for the test subconditions *consS*, *vowelE*, and *vowelS* (all  $p < .001$ ), such that each group performed

more accurately on the native relative to the nonnative stimuli. In other words, the L2 English group outperformed the L2 Spanish group in the Spanish conditions (*consS*, *vowelS*), whereas the L2 Spanish group outperformed the L2 English group in the English condition (*vowelE*). The test subcondition *consE* did not show a significant group effect because both groups performed with similar accuracy in that condition (see Figure 1). Performance of both groups was similar in all control subconditions (all  $p > .10$ ).

Pairwise comparisons (with Sidak adjustment for multiple comparisons) were carried out to examine differences between subconditions within the L2 Spanish group. These tests showed that none of the control subconditions differed from the *consE* or *vowelE* subconditions, but that all control subconditions triggered significantly fewer errors than the nonnative *consS* or *vowelS* subconditions (all  $p < .01$ ). The test subconditions *consE* and *vowelE* did not differ from each other ( $M_{\text{err}} = 4.6\%$ ,  $CI = 1.0\text{--}8.2$ ; and  $M_{\text{err}} = 3.6\%$ ,  $CI = 0\text{--}7.2$ , respectively), but they triggered significantly fewer errors than the nonnative *consS* and *vowelS* subconditions ( $M_{\text{err}} = 24.0\%$ ,  $CI = 20.4\text{--}27.6$ ; and  $M_{\text{err}} = 19.1\%$ ,  $CI = 15.5\text{--}22.7$ , respectively). Finally, performance in these two nonnative *consS* and *vowelS* subconditions did not differ ( $p > .10$ ), but was significantly less accurate than in all other conditions (all  $p < .01$ ).

For the L2 English group, a similar picture emerged, with one crucial difference: Pairwise comparisons showed that none of the control subconditions differed from the native *vowelS* subcondition. Unexpectedly, however, the native *consS* subcondition ( $M_{\text{err}} = 11.6\%$ ,  $CI = 8.3\text{--}15.0$ ) elicited significantly less accurate performance, compared to three out of the four control conditions. This can be explained by the fact that the /d-t/ contrast used in this condition is generally difficult to perceive (Daidone & Darcy, 2014). All control subconditions triggered fewer errors ( $M_{\text{err}} = 4.0\text{--}5.7\%$ ,  $CI = .6\text{--}9.0$ ) than the nonnative *vowelE* subcondition ( $M_{\text{err}} =$

29.3%, CI = 25.9–32.6; all  $p < .01$ ). However, none of the control subconditions differed from the *consE* subcondition ( $M_{\text{err}} = 8.0\%$ , CI = 4.6–11.3). Performance on the native *consS* and *vowelS* subconditions did not differ ( $M_{\text{err}} = 11.6\%$ , CI = 8.3–15.0; and  $M_{\text{err}} = 9.9\%$ , CI = 6.6–13.3, respectively), but both triggered significantly fewer errors than the nonnative *vowelE* subcondition ( $M_{\text{err}} = 29.3\%$ , CI = 25.9–32.6). Unexpectedly, however, like the control conditions, the two native test subconditions also did not differ from the nonnative *consE* subcondition, which was quite accurate ( $M_{\text{err}} = 8.0\%$ , CI = 4.6–11.3). Finally, participants were significantly more accurate on the *consE* than on the *vowelE* subcondition, and while the error rate on *consE* did not differ from any other subconditions (except *vowelE*), *vowelE* differed significantly from all others. As shown in Figure 1, for the L2 English group, error rate for the nonnative test consonant contrast /ʃ-ʧ/ (*consE*) was as low as for the native contrasts.

## Discussion

The source of observed difficulties in perception can be attributed to the nonnativeness of the contrasts tested. When a contrast was common to the L1 and L2, error rates were low, irrespective of the language in which the stimuli were recorded. However, the bilingual ABX task revealed specific difficulties with nonnative contrasts for both learner groups, with the exception of the English consonant contrast /ʃ-ʧ/, which did not pose difficulties for the L2 English learners. In fact, the ranges and standard deviations of individual performances in the nonnative vowel conditions were very similar for both L2 groups (see Table 4) but differed in the consonant conditions. Because the consonant conditions for the L2 English group exhibited less individual variation than the other conditions, for the purpose of this study, we used each individual's performance (error rate) in the nonnative *vowel* conditions as a measure of L2 perceptual processing. However, since the two means in the vowel conditions differed (as shown

by the nonoverlapping CIs), even though the range of performance was the same, we computed a *z* score for the individual error rate (see below).

**Table 4** Descriptive statistics for nonnative vowel and consonant perception (percent error) and inhibition (score)

Group	Measure	<i>M</i>	<i>SD</i>	<i>Range</i>	<i>95% CI</i>
L2 English	English vowel	29.3	22.0	0.0–75.0	25.9, 32.6
L2 Spanish	Spanish vowel	19.1	20.0	0.0–75.0	15.5, 22.7
L2 English	English consonant	8.0	14.0	0.0–50.0	4.6, 11.3
L2 Spanish	Spanish consonant	24.0	23.0	0.0–75.0	20.4, 27.6
L2 English	Inhibition	1.1	0.3	0.4–1.7	0.9, 1.2
L2 Spanish	Inhibition	1.0	0.1	0.8–1.3	0.9, 1.1

*Note.* CI = confidence interval.

## Analysis 2: Production

### Materials and Procedure

We examined the same contrasts as those used in perception through a delayed sentence repetition task (e.g., Trofimovich & Baker, 2006), either in the L2 (for the learners) or in the L1 (for the controls). There were four pairs of sentences for each contrast, for a total of 16 sentences per language (see Appendix S6 in the Supporting Information online). Participants sat in front of a computer screen in a sound-isolated recording booth, wearing high quality dynamic stereo headphones. Materials were displayed on the screen from slides in Microsoft PowerPoint. Participants' productions were recorded on a computer or through a digital recorder at a sampling rate of 22,050 Hz with a 16-bit resolution in a mono channel. The audio files were then

spliced into individual sentences for acoustic analysis in Praat (Boersma & Weenink, 2013). In each trial, participants heard a question (prompt), followed after 250 milliseconds by an answer (response). After a 500 millisecond delay, the prompt was presented again, and the participants had to repeat aloud the response heard previously. The written sentences appeared on the screen together with the first auditory presentation of the prompt-response pair, and disappeared for the second presentation of the prompt and for the recording of the answer. All L2 learners received instructions in their respective L1 and completed a warmup prompt in the L1 before moving on to the L2. Controls completed the set in their L1. Participants were allowed to repeat trials once in case of hesitations or misremembered words (in these cases, only the second recording was analyzed). This task took about 5–7 minutes to complete. The sentence prompts and the corresponding responses in both languages were recorded by the same two female bilinguals and were normalized for amplitude. In half the sets, one voice was used for the prompt token and the other was used for the response tokens; the reverse was done for the remaining sets.

### **Data Analysis**

For the L2 Spanish learners, we predicted that both the monophthong /e/ (e.g., *pena* ['pena] “shame”) and the diphthong /ej/ (e.g., *peina* ['peina] “[he/she] combs”) would be produced similarly (Morrison, 2006), more specifically like a long, diphthongized vowel similar to the English /eɪ/ (e.g., ['pe'na]). That is, compared to NSs, we expected that L2 learners’ monophthongs would display more formant movement and be longer; conversely, L2 learners’ diphthongs were expected to display less formant movement. Hence, the overall vowel duration was measured, and three measurement points were located 20%, 50%, and 80% into the vowels. A duration ratio (diphthong/monophthong) was calculated to assess differences in duration between diphthongs and monophthongs while controlling for individual variation in speech rate.

The mean values for F1, F2, and  $f_0$  (fundamental frequency) were extracted from a 10 millisecond window centered at the three measurement points. These frequency measures were first converted to Bark, and then a Bark-distance metric was applied, subtracting B0 from B1 (B1-B0) for vowel height and B1 from B2 (B2-B1) for degree of vowel fronting, where “B” stands for Bark-converted frequency (Hz) values (Baker & Trofimovich, 2005; Bohn & Flege, 1990). The amount of formant movement in the vowel was obtained by computing the Euclidean distance between the 20% and the 50% measurement points and between the 50% and the 80% measurement points. The two Euclidean distances were added. This spectral distance score was used as a measure of formant movement, as represented in the Bark-normalized vowel space.

For the /d-r/ contrast (e.g., *cada* ['kaða] “each” vs. *cara* ['kara] “face”), we predicted that learners’ productions would be ambiguous at best between /d/ and /r/. Specifically, we expected /d/ (spirantized [ð]) to be realized without the spirantization. For the tap, we expected a number of realizations, from the English-like alveolar approximant to an English-like /d/. Given the variety of possible realizations, it was not practical to obtain specific acoustic measures. Instead, we visually and auditorily examined the spectrograms and made a categorical decision about the accuracy of the target realizations of intervocalic /d/ and intervocalic /r/. The L2 Spanish productions were scored as accurate only when they presented the auditory and acoustic characteristics found in the native Spanish productions: the /d/ produced as a spirantized [ð] and the /r/ produced as a single-closure tap with a very short constriction duration. For each participant’s four tokens of /d/ and four tokens of /r/, every correct production was given one point, and every incorrect production was given a score of zero, thus yielding a score out of eight. An interrater reliability analysis revealed 95.6% agreement between our categorical production scores and those of a L1 Spanish naïve listener (Cohen’s  $\kappa = .92$ ).

For the L2 English learners, in the case of the /i-ɪ/ contrast, we predicted that both duration and spectral realizations would differ from NS productions. Because Spanish learners of English have been shown to rely primarily on duration cues in distinguishing this tense-lax vowel contrast, unlike native English speakers who rely primarily on spectral cues (e.g., Morrison, 2008, 2009), a duration ratio (tense/lax) was computed as a measure of accuracy in temporally differentiating the two vowels. To obtain spectral measures, F1, F2, and  $f_0$  were extracted from a 15 millisecond window centered at the midpoint of the steady-state portion of the second formant of the vowel. The Euclidean distance between the contrasting vowels in a Bark-normalized vowel space was taken as a measure of accuracy in spectrally differentiating the two vowels. For the /j-ɟ/ contrast, we predicted that both sounds might be realized as the affricate /tʃ/. Spectrograms were visually and auditorily examined to make a categorical decision about the accuracy of production. The L2 English productions were scored as accurate if produced as palatoalveolar and exhibited presence (/tʃ/) or absence (/tʃ/) of a closure. As in the analysis of the /d-r/ tokens, this yielded a score out of eight. Interrater reliability reached 99.5% between our categorical production scores and those of a L1 English naïve listener (Cohen's  $\kappa = .98$ ).

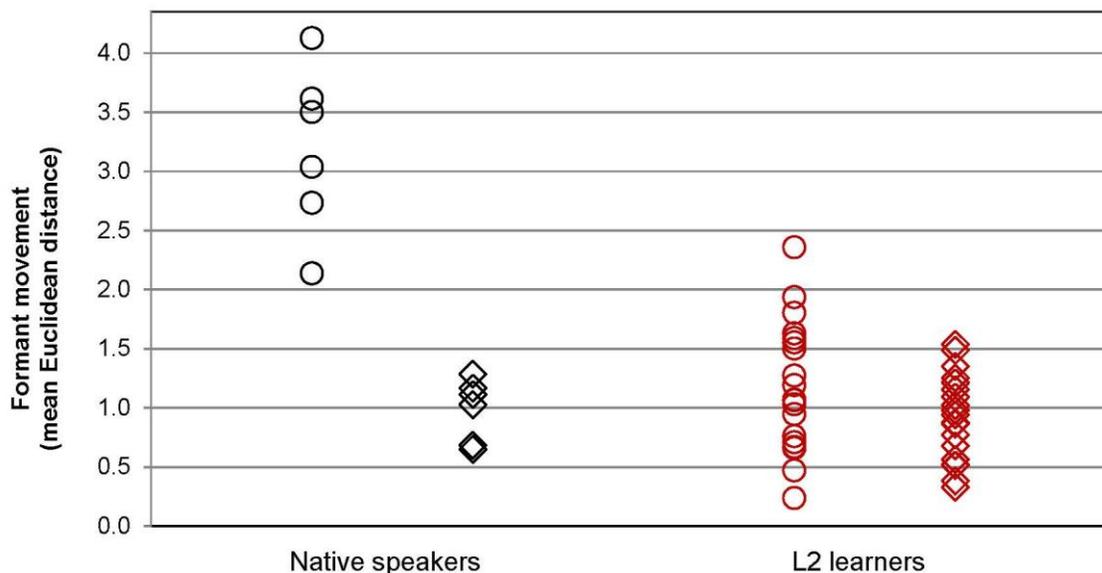
## **Results**

### *Spanish Contrasts*

The L2 Spanish learners ( $n = 19$ ) produced Spanish monophthongs with significantly longer duration ( $M = 97$  ms,  $SD = 14$ ) than native controls ( $n = 6$ ) did ( $M = 84$  ms,  $SD = 7$ ), Mann-Whitney  $U = 20.0$ ,  $z = -2.35$ ,  $p = .017$ ,  $r = .47$ . However, both speaker groups produced diphthongs of similar duration ( $M = 113$  ms,  $SD = 15$ ; and  $M = 113$  ms,  $SD = 13$ , respectively), Mann-Whitney  $U = 62.0$ ,  $z = .32$ ,  $p = .78$ ,  $r = .06$ . Overall, diphthongs were produced with

significantly longer durations than monophthongs by both the L2 Spanish learners,  $t(18) = 4.95$ ,  $p < .001$ ,  $r = .58$ , and native controls, Wilcoxon  $T < .001$ ,  $p = .028$ ,  $r = .90$ . However, the L2 Spanish learners produced the diphthong-monophthong contrast with significantly smaller duration ratios ( $M = 1.16$ ,  $SD = .14$ ) than the Spanish natives did ( $M = 1.35$ ,  $SD = .16$ ), Mann-Whitney  $U = 93.0$ ,  $z = 2.29$ ,  $p = .021$ ,  $r = .46$ , suggesting that, as expected, they produced a smaller duration distinction in the production of this vowel contrast compared to native controls.

The two groups differed in the amount of formant movement during vowel production. The L2 Spanish learners were found to produce the Spanish diphthong /ej/ with much *less* formant movement ( $M = 1.18$ ,  $SD = .55$ ) than native controls ( $M = 3.19$ ,  $SD = .71$ ), as shown in Figure 2, Mann-Whitney  $U = 113.0$ ,  $z = 3.56$ ,  $p < .001$ ,  $r = .71$ . However, contrary to our predictions, the L2 Spanish learners did not produce the Spanish monophthong /e/ as an English-like diphthongized vowel. Instead, they produced both vowels as a monophthongal /e/, with very little formant movement (1.18 for diphthongs vs. .92 for monophthongs,  $SD = 0.36$ ),  $t(18) = 1.92$ ,  $p = .071$ ,  $r = .17$ , but distinguished them through duration. For the /d-t/ contrast, the L2 Spanish learners obtained an average score of 4.88 ( $SD = 2.38$ ,  $range = 0-7$ ), representing 52.4% of targetlike productions, whereas the NSs performed at ceiling ( $M = 7.89$ ,  $SD = .41$ ).



**Figure 2** Amount of formant movement for diphthongs (circles) and monophthongs (diamonds) for native speakers (left) and L2 learners (right).

### *English Contrasts*

For the production of English /i-ɪ/, the L2 English learners ( $n = 22$ ) differed from the L1 English controls ( $n = 7$ ) in that they produced these vowels with less distinct quality and duration. The L2 English learners produced /i/ significantly higher and more fronted than /ɪ/, in terms of B1-B0 values (1.56 vs. 1.86 for /i/ and /ɪ/, respectively),  $t(21) = -4.27$ ,  $p < .001$ ,  $r = .46$ , and B2-B1 values (10.56 vs. 10.16 for /i/ and /ɪ/, respectively),  $t(21) = 4.48$ ,  $p < .001$ ,  $r = .48$ . However, they did not show sufficiently large spectral distances in the production of the contrast to be comparable to the L1 English speakers' productions; the L2 English learners ( $M = .61$ ,  $range = .13-1.39$ ,  $SD = .35$ ) obtained significantly smaller mean spectral distance scores (Euclidean distances in Bark) than the L1 English speakers did ( $M = 3.87$ ,  $range = 3.12-5.12$ ,  $SD = 0.76$ ), Mann-Whitney  $U = 154.0$ ,  $z = 3.92$ ,  $p < .001$ ,  $r = .73$ . Similarly, the L2 learners' duration ratios ( $M = 1.08$ ,  $SD = .17$ ) were much smaller than the L1 English speakers' ratios ( $M = 1.20$ ,  $SD =$

.15), but the difference did not reach significance, Mann-Whitney  $U = 109.0$ ,  $z = 1.63$ ,  $p = .11$ ,  $r = .30$ . The learners produced these vowels with almost distinct durations ( $M = 87$  ms,  $SD = 18$  for /i/;  $M = 82$  ms,  $SD = 13$  for /ɪ/) approaching significance,  $t(21) = 2.04$ ,  $p = .054$ ,  $r = .16$ . For the NSs, the difference was larger ( $M = 101$  ms,  $SD = 21$  for /i/;  $M = 85$  ms,  $SD = 17$  for /ɪ/). Overall, the results indicate that the L2 English learners had difficulty with producing a clear distinction between the two English vowels. For the /ʃ-tʃ/ contrast, the L2 English learners obtained an average score of 7.02 ( $SD = 1.23$ ,  $range = 4-8$ ), representing 86.1% of targetlike productions, whereas the NSs' productions were 100% accurate, as expected.

## Discussion

This analysis examined vowel and consonant production for the same contrasts as the ones tested in perception, and the results matched the perception data. The learners' difficulties were pronounced for both vowel contrasts. The learners were less accurate at spectrally differentiating the tense/lax and monophthong/diphthong contrasts, compared to the native controls. For consonants, the results were mixed. While the /d-r/ contrast presented difficulties for many L2 Spanish learners, the fricative-affricate /ʃ-tʃ/ contrast, which only appeared word-initially in our materials, was not problematic for the L2 English learners. One possibility is that the fricative-affricate contrast was easier for the English learners because [ʃ] is used in Andalusian Spanish as a result of the weakening of /tʃ/ (e.g., *muchacho* [mu'ʃaʃo] "boy," Hualde, 2005). However, speakers of Andalusian Spanish using /ʃ/ are also constantly exposed to standard Spanish (using only /tʃ/), and their phonology might include two variants of the same category determined dialectally ([tʃ] and [ʃ]), possibly facilitating the production of this English contrast. Another potential reason for the accuracy differences observed between vowels and consonants is that our scoring system left less room for individual differences in consonants with a score out of eight,

as compared to detailed acoustic analyses of vowels. Therefore, it was decided to keep both scores separate, with no composite production score calculated for vowels and consonants.

### **Analysis 3: Retrieval-Induced Inhibition**

#### **Materials and Procedure**

In the present study, individual differences in inhibitory control were measured as retrieval-induced inhibition (Anderson et al., 1994; Veling & van Knippenberg, 2004). The inhibitory control task, based on the task in Lev-Ari and Peperkamp (2013), was administered through E-prime in the participants' L1 only and took about eight minutes to complete. Participants memorized six words of three different categories (vegetables, occupations, or animals) presented visually on the screen and then practiced only half the items from two categories (e.g., *tomato*, *nurse*) by typing them several times on the screen. This increased the level of activation of the practiced items and was expected to cause the inhibition (i.e., decrease in the level of activation) of the unpracticed items in these practiced categories. By contrast, the items from the unpracticed category were not inhibited and served as control items. Participants were then tested on the recognition of the practiced items as well as two types of unpracticed items (in addition to new distractors to make the task more meaningful): (a) those from the two practiced categories (inhibited) and (b) those from the unpracticed category (control). Participants with greater inhibitory skill were expected to bring the unpracticed items in practiced categories to lower activation levels, with the consequence that retrieval RTs during recognition would be longer for these items than for practiced and control items.

## Results and Discussion

Given that inhibition of lexical items depends on the practice of some items, participants who did not recall two or more items out of six in the practiced categories were excluded from the analyses (one L2 Spanish learner, leaving 19 participants; six L2 English learners, leaving 17). Therefore, the data for 36 L2 learners were analyzed. Following Lev-Ari and Peperkamp (2013), median recognition RTs were computed for each participant. The median RT was faster for the practiced items ( $Mdn = 807$  ms) than for the inhibited items ( $Mdn = 985$  ms), a significant difference,  $t(35) = -3.05$ ,  $p = .002$ ,  $r = .21$ . However, despite the tendency for the inhibited items to be slower than the nonpracticed control items ( $Mdn = 933$  ms), this difference was not significant in the group analysis,  $t(35) = -.60$ ,  $p > .10$ ,  $r = .01$ . This comparison still confirmed that the nonpracticed items from the practiced categories had been inhibited, as shown by their significantly longer RTs as compared to the practiced items. Following Lev-Ari and Peperkamp (2013), an inhibitory control score was obtained by dividing the median RT for inhibited items by the median RT for control items, such that the higher the score above 1, the stronger the inhibition. The bottom two rows in Table 4 present descriptive statistics for each group. One extreme score (3.05) was removed from the L2 English group to make the distribution comply with criteria for normality. A  $t$  test (unequal variances, two-tailed) confirmed that both learner groups did not differ significantly in inhibition scores,  $t(19) = .73$ ,  $p > .10$ . Thus, inhibitory control skill, and the subsequent analyses of the participants' phonological processing in relation to it, was not L1 dependent, allowing us to merge both learner groups for analyses.

## **Analysis 4: Relationship Between Inhibition and Phonological Processing**

### **Data Analysis**

Because our goal was to examine factors related to L2 phonological processing, only the data from the L2 learners were included in these analysis. Out of the 43 L2 learners tested initially, only 34 (16 L2 English, 18 L2 Spanish) with complete data in all tasks were included in the correlation analyses. One difficulty with collapsing data from two different L2 groups is that they may vary in L2 proficiency. The L2 learners' overall proficiency was measured through a test which was designed to assess receptive lexical knowledge in learners (X-Lex, providing a vocabulary size estimate of 0–5,000 words, Meara, 2005) and whose scores have been shown to relate to L2 proficiency levels (Miralpeix, 2012). Vocabulary size was then included as a covariate in the analyses, in order to tease out this variable from other effects. Full descriptive statistics for 43 learners' background variables appear in Appendix S7 in the Supporting Information online.

A series of Mann-Whitney  $U$  tests showed that the L2 English learners were slightly older in age at testing,  $U = 242.50$ ,  $z = 3.54$ ,  $p < .001$ ,  $r = .61$ , used their L2 more often,  $U = 234.50$ ,  $z = 3.14$ ,  $p = .001$ ,  $r = .54$ , had studied their L2 for a longer period of time,  $U = 226.50$ ,  $z = 2.87$ ,  $p = .003$ ,  $r = .49$ , and had a greater L2 vocabulary size,  $U = 277.00$ ,  $z = 4.59$ ,  $p < .001$ ,  $r = .79$ , compared to the L2 Spanish learners. However, these learner groups did not differ in motivation, self-assessed ability in L2 speaking and listening, or the age at which they first started to learn and use their L2. In sum, the L2 English learners appeared to be slightly more proficient in their L2 than the L2 Spanish learners, a group difference we controlled for by using proficiency scores. The descriptive statistics for the target variables used in the correlation analyses are shown in Table 5.

**Table 5** Means (standard deviations in parentheses) for L2 learners' cognitive and phonological measures

Variable	L2 English ( <i>n</i> = 16)	L2 Spanish ( <i>n</i> = 18)	All learners ( <i>n</i> = 34)
Inhibition score	1.05 (0.30)	1.01 (0.11)	1.03 (0.21)
ABX error (vowels)	0.29 (0.12)	0.20 (0.14)	0.24 (0.14)
ABX <i>z</i> score	2.98 (1.50)	1.36 (1.63)	2.12 (1.75)
Consonant production (max. 8)	6.91 (1.27)	4.09 (2.45)	5.41 (2.42)
Vowel production <i>z</i> score	-3.68 (0.43)	-2.54 (0.78)	-3.08 (0.86)

The perception scores used for each participant were *z* scores for the error rate in the perception of nonnative vowels in Analysis 1; these *z* scores were based on the mean and standard deviations for the NSs of the learners' respective L2 (e.g., for the L2 English learners, the values were taken from the English NSs). The production score for consonants corresponded to the number of correct productions (out of eight) for nonnative contrasts by each participant. Given that the NSs' mean accuracy was at ceiling (*SD* = 0), we did not compute a *z* score for consonant production accuracy. Finally, to avoid issues related to comparing raw spectral distance scores (formant movement for diphthongs vs. spectral quality for /i-ɪ/) for the vowel production measure, we computed a *z* score for each participant based on the means and standard deviations for the NSs of the learners' respective L2, thus making individual scores more comparable.

## Results

We first examined how the scores in perception related to the scores in production.<sup>2</sup> Given our modest sample size, we also report marginally significant correlations. As shown in Table 6, the

L2 learners' self-estimates of listening and speaking ability were correlated with each other, but they showed no relationship with their performance in perception and production. By contrast, our perception measures (raw error rate and  $z$  scores for ABX) correlated with the consonant production score, such that a learner with a lower error rate (or lower, more nativelike  $z$  score) was also more accurate in consonant production. Both the ABX  $z$  score and consonant production measures were also marginally related with vowel production scores, in what at first glance looks like the opposite of what was predicted.

**Table 6** Partial correlations (Pearson) between production and perception measures, controlling for proficiency with X-Lex scores

Measure	1	2	3	4	5	6
1. Self-rated speaking	—					
2. Self-rated understanding	0.43**	—				
3. ABX $z$ score	-0.08	0.16	—			
4. ABX error (vowels)	-0.09	0.15	0.99**	—		
5. Consonant production	0.16	-0.06	-0.33*	-0.40*	—	
6. Vowel production	0.21	0.03	-0.24†	-0.19	-0.26†	—

*Note.*  $n = 33$ . † $p < .10$ , \* $p < .05$ , \*\* $p < .01$ , one-tailed.

The vowel scores were also  $z$  scores (mostly negative for the learners), indicating that the learners obtained significantly smaller mean spectral distance scores (Euclidean distances in Bark) relative to the NSs. By contrast, the ABX  $z$  score was mostly larger than 0 because the learners made more errors on average relative to the NSs. Therefore, a score closer to 0 in both cases indicated more targetlike performance. Thus, the relationship was as expected in both

cases: The lower (closer to 0) the  $z$  score in ABX for perception, the higher (closer to 0) it was for vowel production (also a  $z$  score). Put differently, the fewer errors someone made in ABX perception, the more nativelike his or her vowel production was. Similarly for consonants, higher accuracy in producing consonants was associated with more nativelike vowel production.

**Table 7** Partial correlations (Pearson) between phonological performance and inhibition scores for all learners, controlling for proficiency with X-Lex scores

Phonological measure	Inhibition
ABX error (vowels)	-0.42**
ABX $z$ score	-0.39*
Consonant production	0.34*
Vowel production	-0.04

*Note.*  $n = 33$ . \* $p < .05$ , \*\* $p < .01$ , one-tailed.

Table 7 summarizes the results of correlation analyses between inhibition scores and the four phonological accuracy scores (controlling for proficiency). These correlations show that error rate and  $z$  scores in ABX perception were significantly negatively correlated with inhibition scores, which indicates that individuals with higher inhibition scores also obtained lower error rates and lower (more targetlike)  $z$  scores in vowel perception. For the consonant score, the same relationship emerged in a partial correlation: A significant association was found, such that the learners with a higher inhibition score also obtained higher accuracy scores in consonant production. No clear pattern emerged between vowel production and inhibition scores. Two regression analyses (summarized in Table 8) showed that the inhibition score was a significant

predictor of ABX perception error rate and perception  $z$  scores. Inhibition explained about 18% of the overall variance in perception error rates ( $p = .018$ ) and perception  $z$  scores ( $p = .029$ ) when controlling for proficiency. This relationship is illustrated graphically in Appendix S8 in the Supporting Information online.

**Table 8** Results of hierarchical regressions using inhibition as predictor of ABX error rate and ABX  $z$  scores, controlling for proficiency with X-Lex scores

Criterion variable	Predictor	$R^2$	$B$	95% CI	$t$	$p$
ABX error rate (vowels)	X-Lex	0.00	0.02	0.00, 0.00	0.13	.901
	Inhibition	0.18	-0.42	-0.49, -0.05	-2.51	.018
ABX $z$ score	X-Lex	0.04	0.16	0.00, 0.01	0.93	.359
	Inhibition	0.18	-0.38	-6.01, -0.35	-2.29	.029

## Discussion

Overall, inhibition correlated with perception and production scores in the expected direction. Once proficiency was controlled (through partial correlations), individuals with higher inhibition scores were found to have both lower error rates and more nativelike  $z$  scores in the vowel perception task. There was also a significant relationship between inhibition scores and higher accuracy in consonant production. Surprisingly, no relationship emerged between vowel production accuracy and inhibition scores. In sum, the relationship between inhibition and perception appeared stronger than the relationship between inhibition and production.

This differential effect of inhibition on perception and production may be due to the inherently different nature of the processing mechanisms which govern speech perception and

production and which might involve different inhibition types. Our inhibition task falls under the “resistance to proactive interference” type outlined by Friedman and Miyake (2004). This type of inhibition might be highly relevant for category learning, helping learners form accurate vowel categories, for instance, while resisting interference from L1-specific memory traces during perception. We interpret the relationship between inhibition scores and ABX perception to mean that those who demonstrated higher inhibition scores may have used this ability in the past to support the learning of L2 segmental categories. Thus, the learners with stronger inhibitory control may have had an advantage over those with poorer inhibitory control during the course of L2 acquisition; this advantage would have enabled learners to develop more accurate representations for L2 segments by virtue of their greater capacity to avoid L1 interference during L2 phonological processing.

For production, in addition to having accurate segmental categories, dominant motor (articulatory) plans from the L1 must be inhibited while speaking. Therefore, it is perhaps not so surprising that the relationship between production and inhibition was less strong in our data; this is because we measured inhibition through a task evaluating resistance to proactive interference. We might have observed a different relationship if we had also used a measure of motor-based inhibition (e.g., the Simon task). In addition, given the possibility that the target consonant contrasts were not ideal to reveal sufficient variation among learners, especially among the L2 English learners, we interpret the consonant results with caution; this issue needs to be reexamined in future research.

However, the obtained perception-production difference in inhibitory control may be unrelated to L1 inhibition effects in L2 learning.<sup>3</sup> It is possible that our inhibition measure reflected the participants’ ability to inhibit the wrong response alternative in the ABX task, that

is, within the L2. Hence, it is theoretically possible that the inhibition relationship was stronger for perception than for production because the ABX task required more, while the production task required less, inhibition within the L2. This hypothesis could explain how those learners with stronger inhibition scores also made fewer errors in the vowel perception task. If correct, this alternative might involve specifically the inhibition of a dominant response: The ABX task structure presented participants with two alternative responses (A, B) and participants' decision in selecting the correct alternative for X was based, among other factors, on acoustic/phonetic similarity. Hence, upon hearing the X token (if X = A), a participant had to reject the competing alternative (B). Phonetically speaking, therefore, B was more different from X than A. It is not immediately clear how rejecting the wrong (i.e., acoustically more dissimilar) response alternative might involve using inhibitory control, as the wrong response does not automatically qualify as the dominant response that is to be inhibited. If this were the case, we should have also observed a similar result pattern in the control condition. This was not the case in our data, but the control condition was not designed to allow for this analysis. In addition, it is also not immediately clear—if ABX scores are to be explained via inhibition of a dominant response—why the ABX scores would correlate with the kind of inhibition we measured in our task, involving resistance to proactive interference (see Friedman & Miyake, 2004, showing that dominant response inhibition was not correlated with resistance to proactive interference).

In sum, to fully rule out an alternative explanation for our findings, it would be necessary to use a variety of inhibition tasks, which measure resistance to a dominant response, as well as resistance to proactive interference (which we argue might be related in important ways to resisting interference from the L1 during acquisition). In addition, a task such as ABX should be designed to allow for subsequent analysis of the specific latency and error patterns in the various

stimulus configurations, to evaluate the extent to which inhibition within the L2 might play a role above and beyond the phonological categorization performance that might result from inhibition of the L1 phonological system. In particular, such an analysis could indicate whether individual differences in inhibitory control would result in a task effect or, as in our main interpretation, would reflect a learning facilitation mechanism.

Last but not least, we also observed a general lack of relationships between inhibition and phonological measures on the one hand and the other demographic variables on the other, which may be due to the nature of our participants' context of learning. Having acquired their L2 in a foreign language classroom (mostly after age 8), the learners received (and continue to receive) input that is fundamentally different in both quality and quantity from learners in an immersion setting (Piske, 2007). Because learners in foreign language settings have less input with which to develop their L2 phonological categories, small differences in their age, L2 proficiency, exposure, or motivation may not be enough to affect phonological category development.

### **General Discussion and Conclusion**

The present study investigated whether language learners with better inhibitory skill are also better equipped to acquire a new phonological system. We hypothesized that late L2 learners' enhanced ability to keep their two phonological systems apart through inhibitory control would minimize L1 interference at all levels of phonological processing and would consequently enhance this processing, leading to the formation of more targetlike representations for L2 segments. Our prediction was that stronger inhibitory control would correlate with greater accuracy in segmental tasks, thereby indicating that L2 learners with more efficient inhibitory control may be better able to suppress the influence of L1 phonological categories and, hence,

develop more accurate L2-specific categories during acquisition. We found that when L2 proficiency was statistically controlled, inhibition was in fact correlated in the expected direction with learners' performance in the perception task and with their consonant production. Vowel production did not show any relationship with inhibitory control. Overall, our findings indicate that inhibitory control might be implicated in L2 phonological processing and might contribute to explaining in part the large variation found in L2 learners' perception and production performance.

Taken together, the results of this study provided support for the contribution of inhibitory control to L2 speech processing in production and perception in a foreign language learning context. Despite the fact that all participants were acquiring their L2 under the constrained input conditions of a foreign language classroom, they still exhibited large intersubject differences in their L2 perception and production, which might partly be explained by individual differences in inhibitory control. This learning context differs radically from that in previous research targeting the effects of inhibition on phonological processing in bilinguals. In Lev-Ari and Peperkamp's (2013) study, where individual differences in inhibitory control were found to explain intersubject variation in how strongly bilinguals' native English VOT was affected by L2 French use, participants were living in a L2 environment. In the foreign language classroom context, which is characterized by L1 dominance and limited L2 exposure and use, learners' inhibitory control capacity might play a bigger role and make a more important difference in how accurately L2 sound representations develop than in an immersion setting, where the much larger L2 exposure and use may contribute more fundamentally to L2 phonological development, compared to learners' cognitive skills. In a L2 environment such as immersion, the L1 might be inhibited more easily and be therefore less accessible, and while this

may result in more accurate L2 performance, it may not necessarily lead to bilinguals developing stronger inhibitory control. This, we believe, might depend on how much a given learner switches between the two languages. In fact, Linck et al. (2008) found evidence of classroom L2 learners in a L1 environment significantly outperforming learners in an immersion setting in a nonlinguistic inhibitory control task, such as the Simon task. This suggests that individual differences in inhibitory control might explain a great deal of variance in L2 speech perception and production despite the limited L2 exposure and use typical of classroom environments.

The mechanism underlying this relationship can, at this point, only be speculative. Our approach relates individual differences in inhibitory control to learners' ability to selectively activate one phonological system and deactivate the other for the purpose of enhancing phonological processing and acquisition. Accordingly, temporarily suppressing the activation of a L1 system, at any level of phonological processing, would allow learners to minimize L1 perceptual interference, thus facilitating the encoding of more accurate L2 phonological representations. Consequently, inhibitory control could be an essential component of executive function facilitating L2 phonological acquisition. Assessment of learners' inhibitory control is also dependent on the task used to measure it (e.g., verbal vs. nonverbal) or the type of inhibitory control (e.g., intentional vs. unintentional) that the task taps into. The kind of inhibition we examined is unintentional and appears to fall under the resistance to proactive interference type (Friedman & Miyake, 2004). In order to increase the reliability of these findings, it is important to obtain measures through tasks requiring intentional as well as unintentional inhibition, and to identify robust tasks through study replication.

Our findings point to learners experiencing a specific benefit in terms of segmental category acquisition from having strong inhibitory control. Do these results mean that having

strong inhibitory control is always “good” for L2 learners? Perhaps not. Depending on the specific situation, strong inhibitory control may not always be linked to a positive outcome: This directionality again depends on the language task performed and on the kind of inhibition involved, as well as on the degree of language switching required by the task or the language context or situation. We predict that learners with weaker inhibitory control might be more comfortable switching between their two languages, and situations that require such switches might favor speakers who overall keep both languages coactivated (and therefore inhibit either language less). Conversely, learners with strong inhibition skills are predicted to be less efficient at switching between languages (showing larger switch costs) but would potentially have more clearly separated phonological systems and more precise L2 phonological representations (see Lev-Ari & Peperkamp, 2013, for a related claim). This is the first study targeting the role of inhibitory control as a means of suppressing L1 interference in L2 speech perception and production for learners in a foreign language context. Our results add to the growing literature on inhibition and language, and show the widespread and generalized nature of the interplay between language and inhibitory control.

## **Notes**

1 Participants took part in an audiometry test, an inhibitory control task, an attention control task, an ABX categorization task, a production task, and a working memory task. They also took a vocabulary test and filled out a background questionnaire.

2 All variables included in the analysis were normally distributed; one extreme value for inhibition in the L2 English group was removed, resulting in  $N = 33$ . Age, self-ratings, and length of residence abroad were not normally distributed and were not included.

3 We thank Pavel Trofimovich for pointing this out.

### References

- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063–1087. doi:10.1037/0278-7393.20.5.1063
- Anrrich, G. M. (2007). *Substitutions for English consonants by adult speakers of Cuban Spanish* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.  
Retrieved from <http://search.proquest.com/docview/304874204>
- Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*, 1–27.  
doi:10.1177/00238309050480010101
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.52) [Computer software]. Retrieved from <http://www.praat.org>
- 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.  
doi:10.1017/S0142716400008912
- Cerviño-Povedano, E., & Mora, J. C. (2011). Investigating Catalan learners of English over-reliance on duration: Vowel cue weighting and phonological short-term memory. In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in*

*the acquisition of second language speech: New Sounds 2010* (Vol. I, pp. 56–64).

Frankfurt am Main: Peter Lang.

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology (Cognitive Science)*, *4*(874), 1–11. doi:10.3389/fpsyg.2013.00874

Daidone, D., & Darcy, I. (2014). Quierro comprar una guitarra: Lexical encoding of the tap and trill by L2 learners of Spanish. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected proceedings of the 2012 Second Language Research Forum* (pp. 39–50). Somerville, MA: Cascadilla Proceedings Project.

Darcy, I., Mora, J. C., & Daidone, D. (2014). Attention control and inhibition influence phonological development in a second language. *Concordia Working Papers in Applied Linguistics*, *5*, 115–129.

Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, *40*, 63–72. doi:10.1016/j.lindif.2015.04.005

Filippi, R., Karaminis, T., & Thomas, M. S. (2014). Language switching in bilingual production: Empirical data and computational modelling. *Bilingualism: Language and Cognition*, *17*, 294–315. doi:10.1017/S1366728913000485

Filippi, R., Leech, R., Thomas, M. S., Green, D. W., & Dick, F. (2012). A bilingual advantage in controlling language interference during sentence comprehension. *Bilingualism: Language and Cognition*, *15*, 858–872. doi:10.1017/S1366728911000708

- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437–470. doi:10.1006/jpho.1997.0052
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*, 78–104. doi:10.1006/jmla.1999.2638
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods*, *35*, 116–124. doi:10.3758/BF03195503
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101–135. doi:10.1037/0096-3445.133.1.101
- Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, *25*, 1031–1036. doi:10.1177/0956797613520014
- Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of a second language phonology. *Brain and Language*, *109*, 55–67. doi:10.1016/j.bandl.2008.01.005
- Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, *12*, 91–114.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67–81. doi:10.1017/S1366728998000133
- Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning* (pp. 57–77). Amsterdam: John Benjamins. doi:10.1075/llt.17.09gui

- Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics*, *28*, 27–42.  
doi:10.1006/jpho.2000.0104
- Hualde, J. I. (2005). *The sounds of Spanish*. New York: Cambridge University Press.
- Kok, A. (1999). Varieties of inhibition: Manifestations in cognition, event-related potentials and aging. *Acta Psychologica*, *101*, 129–158. doi:10.1016/S0001-6918(99)00003-7
- Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *Journal of the Acoustical Society of America*, *128*, 3757–3768.  
doi:10.1121/1.3506351
- Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*, 320–331.  
doi:10.1016/j.wocn.2013.06.002
- Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language. The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science*, *18*, 29–34. doi:10.1111/j.1467-9280.2007.01844.x
- Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon*, *3*, 349–374. doi:10.1075/ml.3.3.06lin
- Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*, 651–662. doi:10.1017/S136672891100054X
- MacKay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*, 103–125. doi:10.1159/000028490

- Meara, P. M. (2005). *X\_Lex: The Swansea Vocabulary Levels Test (Version 2.05)* [Computer software]. Swansea: Lognostics.
- Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, *17*, 89–117. doi:10.1017/S1366728913000084
- Miralpeix, I. (2012). *X\_Lex and Y\_Lex: A validation study*. Paper presented at the 22nd VARG Conference (Vocabulary Acquisition Research Group), Newtown, UK.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8-14.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- Morrison, G. S. (2006). *L1 & L2 production and perception of English and Spanish vowels: A statistical modelling approach* (Unpublished doctoral dissertation). University of Alberta, Canada.
- Morrison, G. S. (2008). L1-Spanish speakers’ acquisition of the English /i/-/ɪ/ contrast: Duration-based perception is not the initial developmental stage. *Language and Speech*, *51*, 285–315. doi:10.1177/0023830908099067
- Morrison, G. S. (2009). L1-Spanish speakers’ acquisition of the English /i/-/ɪ/ contrast II: Perception of vowel inherent spectral change. *Language and Speech*, *52*, 437–462. doi:10.1177/0023830909336583

- Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*, 81–108. doi:10.1017/S0272263199001035
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220–246. doi:10.1037/0033-2909.126.2.220
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics, 41*, 407–420. doi:10.1016/j.wocn.2013.07.005
- Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*, B9–B17. doi:10.1016/S0010-0277(97)00030-9
- Perfect, T. J., Moulin, C. J., Conway, M. A., & Perry, E. (2002). Assessing the inhibitory account of retrieval-induced forgetting with implicit-memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 1111–1119. doi:10.1037/0278-7393.28.6.1111
- Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language learning* (pp. 301–304). Amsterdam: John Benjamins.
- Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology (Cognition), 3*(57), 1–18. doi:10.3389/fpsyg.2012.00057
- Reilly, J., Troiani, V., Grossman, M., & Wingfield, R. (2007). An introduction to hearing loss and screening procedures for behavioral research. *Behavior Research Methods, 39*, 667–672. doi:10.3758/BF03193038

- Reiterer, S., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology (Language Sciences)*, 2(271), 1–12. doi:10.3389/fpsyg.2011.00271
- Rose, M. (2010). Differences in discriminating L2 consonants: A comparison of Spanish taps and trills. In M. T. Prior (Ed.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 181–196). Somerville, MA: Cascadilla Proceedings Project.
- Safronova, E., & Mora, J. C. (2013). Attention control in L2 phonological acquisition. In A. Llanes Baró, L. Astrid Ciro, L. Gallego Balsà, & R. M. Mateus Serra (Eds.), *Applied linguistics in the age of globalization* (pp. 384–390). Lleida: Edicions de la Universitat de Lleida.
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency. Does musical ability matter? *Psychological Science*, 17, 675–681. doi:10.1111/j.1467-9280.2006.01765.x
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10, 281–284. doi:10.1111/1467-9280.00151
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30. doi:10.1017/S0272263106060013
- Veling, H., & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 315–318. doi:10.1037/0278-7393.30.2.315

## **Supporting Information**

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1.** Motivation Questionnaire (L1 English, L1 Spanish).

**Appendix S2.** Phonetic Realization of Example Stimuli Used in the ABX Task.

**Appendix S3.** Descriptive Statistics for Participants' Performance in the ABX Task.

**Appendix S4.** Parameter Estimates for ABX: L1 by Condition.

**Appendix S5.** Parameter Estimates for ABX: L1 by Subcondition.

**Appendix S6.** Production Sentences.

**Appendix S7.** Descriptive Statistics for Background Variables Used in Correlation Analyses.

**Appendix S8.** Associations Between Inhibition and Perception/Production Accuracy.