Article

The Perception of Postalveolar English Obstruents by Spanish Speakers Learning English as a Foreign Language in Mexico

Mariela López Velarde and Miquel Simonet *

Program in Second Language Acquisition and Teaching, University of Arizona, Tucson, AZ 85721, USA; marialopez@arizona.edu
*   Correspondence: simonet@arizona.edu

Abstract: The present study deals with the perception (identification and discrimination) of an English phonemic contrast (/t∫/–/∫/, as in *cheat* and *sheet*) by speakers of two Mexican varieties of Spanish who are learning English as a foreign language. Unlike English, Spanish does not contrast /t∫/ and /∫/ phonemically. Most Spanish varieties have [t∫], but not [∫]. In northwestern Mexico, [∫] and [t∫] find themselves in a situation of "free" variation—perhaps conditioned, to some extent, by social factors, but not in complementary distribution. In this variety, [∫] and [t∫] are variants of the same phoneme. The present study compares the perceptual behavior of English learners from northwestern Mexico, with that of learners from central Mexico, whose native dialect includes only [t∫]. The results of a word-categorization task show that both groups of learners find *cheat* and *sheet* difficult to identify in the context of each other, but that, relative to the other learner group, the group of learners in northwestern Mexico find this task to be particularly challenging. The results of a categorical discrimination task show that both learner groups find the members of the /t∫/–/∫/ contrast difficult to discriminate. On average, accuracy is lower for the group of learners in northwestern Mexico than it is for the central Mexicans. The findings suggest that the phonetic variants found in one's native dialect modulate the perception of nonnative sounds and, consequently, that people who speak different regional varieties of the same language may face different obstacles when learning the sounds of their second language.

Keywords: second language acquisition; phonology; discrimination; cross-linguistic assimilation; obstruent; affricate; fricative; dialect; English; Spanish

1. Introduction

Most people "have an accent" when speaking a language other than their native one(s). This has been widely documented, and we currently have a sizeable scientific literature describing and explaining this phenomenon—see the following reviews (Best and Tyler 2007; Bohn 2017; Broselow and Kang 2013; Chang 2019; Colantoni et al. 2015; Davidson 2017, p. 201; Eckman 2012; Flege 1995; Piske et al. 2001; Simonet 2016). Interestingly, "having an accent" is not restricted to speech production, but also manifested in perception. Current models of L2 speech acquisition account for those findings—typically from the perspective of perception and categorization—by postulating some sort of interaction between native and nonnative sounds in the representational network of bilinguals (Best and Tyler 2007; Escudero 2005; Flege 1995; van Leussen and Escudero 2015). L2 learners have an "accent" in their L2, these models state, because they already have internalized knowledge of a first language (L1). Native and nonnative sounds must find a way to co-exist, and this typically results in modifications to the nature of such sounds. In other words, L2 listeners assimilate the sounds of

their L2 in terms of the categories that are robustly represented in their phonology by the time they are learning the L2 (i.e., their L1), and they acquire these new sounds as a function of how they map them.

In English, /tʃ/ and /ʃ/ constitute a phonemic contrast, as seen in minimal pairs such as *sheet–cheat* and *chair–share*. Spanish does not have this contrast. Most varieties of Spanish have [tʃ] in their inventory, but they do not have [ʃ] (Hualde 2005, pp. 152–72). In spelling, /tʃ/ is systematically represented by the digraph <ch>, as in *charco* 'puddle' [ˈtʃaɾko] and *chamarra* 'jacket' [tʃaˈmara], and most Spanish speakers would consistently pronounce this phoneme as a postalveolar affricate, [tʃ]. It follows that, if they are to acquire the English /ʃ/–/tʃ/ contrast successfully, native Spanish speakers who possess this particular phonological system must develop a new phoneme, (/ʃ/), in opposition to one they can recycle from their native language, (/tʃ/); they must create a new contrastive category, and they must assign to it a new phonetic substance. Learning new sounds and new oppositions typically presents a significant phonological challenge (Best and Tyler 2007; Colantoni et al. 2015; Escudero 2005).

Native speakers of some regional varieties of Spanish, on the other hand, may have an acquisitional obstacle of a different nature. In some dialects, both [tʃ] and [ʃ] are found, but not in phonemic opposition. One such variety is spoken in northwestern Mexico, where people are known to pronounce Spanish words that have <ch> variably, with either [ʃ] or [tʃ] (Alessi Molina and Díaz 1994; Amastae 1996; Brown 1989; Carreón Serna 2007; Martín Butragueño 2009; Méndez 2017; Moreno de Alba 1994; Serrano Morales 2000, 2009). In northwestern Mexico, therefore, [ˈtʃaɾko] and [ˈʃaɾko] are common variants of the same word, *charco* 'puddle.' It seems to follow that, in order to acquire the English /ʃ/–/tʃ/ contrast successfully, native Spanish speakers from northwestern Mexico do not need to learn any new sounds. They already have both [tʃ] and [ʃ] in their inventory of phonetic categories. However, and this might be crucial, what they must do is learn that these two sounds are not variants of the same phoneme, like they are in their native Spanish variety, but separate phonemes (or separate contrastive categories). Learning new mappings between surface and underlying phonological representations presents a substantial acquisitional obstacle of a different kind (Barrios et al. 2016b).

The present study aims at contributing to the literature on the effects of native linguistic experience on the acquisition of L2 sounds. Most importantly, it examines the relative difficulty of developing new categories (new sounds) *versus* that of developing new phonemic contrasts between sounds one can reuse from one's native phonetic inventory (new mappings). The study is concerned with categorization patterns in the perception of an English phonemic contrast, /ʃ/–/tʃ/, by two groups of L1 Spanish learners of English who speak different regional varieties of their native language.

1.1. Cross-Linguistic Interactions in L2 Speech Acquisition

The fact that native and nonnative sounds interact in the bilingual mind is illustrated by a well-known example, that of Spanish-speaking learners of English, who tend to have difficulties with the English /i/–/ɪ/ and /æ/–/ɑ/ contrasts (Barrios et al. 2016a; Casillas 2015; Escudero and Boersma 2004; Flege et al. 1994, 1997; Flege and Bohn 1989; Kondaurova and Francis 2008; Morrison 2008, 2009). Spanish has five phonemic vowels, /i e a o u/, and Spanish-speaking learners of English tend to assimilate both English /i/ and /ɪ/ to a single native Spanish vowel, /i/ (e.g., Flege and Bohn 1989). This two-to-one cross-linguistic assimilation pattern creates an acquisitional obstacle for this learner population because it makes the two members of the English /i/–/ɪ/ contrast difficult to discriminate from each other (e.g., (Flege et al. 1994)). The English /ɑ/–/æ/ contrast also presents a challenge for Spanish-speaking learners, as both English vowels are cross linguistically assimilated to a single Spanish vowel, /a/ (Barrios et al. 2016a; Casillas and Simonet 2016). These findings indicate that the obstacles L2 learners encounter when acquiring the phonology of their L2 are, at least in part, determined by the listeners' native language background and the cross-linguistic assimilations established between L1 and L2 sounds.

Several theoretical accounts have attempted to explain the obstacles learners face during their acquisition of the L2 phonology. Two such models are the Perceptual Assimilation Model applied to L2 learning (PAM-L2) (Best and Tyler 2007) and the Second Language Linguistic Perception model

(L2LP) (Escudero 2005; van Leussen and Escudero 2015). Both of these frameworks postulate that the native and nonnative sounds of L2 learners interact (in some way). The manner in which this interaction is modelled, however, differs in the two accounts. The PAM-L2 proposes that difficulties arise as a function of the assimilability of L2 contrasts to L1 categories. Cross-linguistic assimilation is claimed to rely on the phonetic (in particular, articulatory) similarity between the L2 and the L1 sounds. The following are three of the possible assimilation patterns the PAM-L2 operationalizes: (i) When two L2 phones are cross linguistically assimilated or equated to two different L1 phonemes, a *two-category* assimilation (TC) is said to have occurred. In a TC scenario, discrimination of the two key L2 segments is predicted to be excellent, since the discrimination of the two corresponding L1 categories is assumed to be optimal. (ii) In contrast, if two L2 sounds are cross linguistically assimilated to the same L1 category and both are equally similar to the L1 sound, a *single-category* assimilation (SC) pattern occurs. The model predicts that, in cases of SC assimilation, the discrimination of the two key L2 phones is poor, as the two L2 sounds will have been categorized as variants of the same sound. This type of cross-linguistic assimilation pattern is particularly challenging for learners. (iii) A third type of assimilation pattern is called *category-goodness* assimilation (CG). In a CG pattern, two contrastive L2 sounds are assimilated to the same L1 category, but the cross-linguistic similarity is greater for one of the categories than for the other. In such a situation, discriminating between two key L2 categories is predicted to range from moderate to good, depending on the degree of category-goodness assimilation for each of the L2 segments.

The L2LP model differs from the PAM-L2 in some important ways. The L2LP claims that, at the initial stages of the L2 learning process, learners transfer or duplicate the entire L1 system to form an interlanguage system. Although the L2 system begins as a duplicate of the L1 grammar (a transferred grammar), this only occurs once, and it is subsequently handled as a separate phonological grammar. The novel system is equipped with the same learning mechanisms available in the L1, and it evolves as experience with the L2 increases. Technically, therefore, the L2LP rejects that the L1 and L2 phonological systems interact because it rejects that the two reside in a common representational network. Nevertheless, since the dedicated L2 system begins its course as an exact copy of the L1, the sound categories and mapping strategies learners developed for the L1 powerfully determine the manner in which L2 sounds are processed, perceived, produced, represented, and, ultimately, learned (or not learned). An aspect of the L2LP that resembles the PAM-L2 is that it operationalizes the existence of "cross-linguistic" *comparisons* in terms of L1 and L2 contrasts based on the phonetic (acoustic, in this case) similarity of L1 and L2 sounds. For instance, the cross-linguistic comparison described as single-category assimilation (SG) in the PAM-L2 is called *new scenario* in the L2LP, and both models predict that learning the L2 contrast in this particular scenario is challenging. What the PAM-L2 calls a two-category (TG) assimilation pattern, the L2LP calls a *similar scenario*; that is, when two L2 sounds each resemble a different L1 sound, learning these two categories is predicted to be easy. In sum, while some aspects of the PAM-L2 and the L2LP make them substantially different from each other, other principles of the two models are fundamentally identical.

Creating a new category during L2 acquisition is particularly difficult in cases in which two contrastive categories of the L2 are cross linguistically assimilated to a single category. A (perhaps) different kind of phonological obstacle presents itself when learners must develop new phonological mappings affecting sounds that already exist in their native inventory. For instance, Spanish has both [d] and [ð], but these are in complementary distribution—the two sounds are allophones (variants) of the same contrastive category, /d/. In English, on the other hand, these two sounds are contrastive, as illustrated by the minimal pair *den–then*. It follows that Spanish-speaking learners of English must develop new mappings between surface sounds that already exist in their inventory and new underlying representations (Barrios et al. 2016b). In other words, they must learn that two sounds that are linked to a single phonemic representation in their L1 are actually contrastive in their L2—they are linked to separate phonemes in the L2. It has been hypothesized that this type of acquisitional obstacle, called *allophonic split*, is particularly challenging for L2 learners (Eckman et al. 2001; Lado 1957).

This prediction derives from the finding that, in native speech, discriminating between contrastive categories is much easier than doing so between categories that are phonetically distinct but not contrastive; in other words, sounds that are not contrastive are perceived to be more similar to each other than contrastive sounds are (Barrios et al. 2016b; Johnson and Babel 2010). The literature on this learning scenario is scant but, interestingly, a recent study has shown that an obstacle of the kind described here does not "cause consistent difficulty for advanced L2 learners in perception" (Barrios et al. 2016b, p. 14). At this juncture, therefore, it is not known which of the learning scenarios—learning a new sound *versus* learning a new mapping—presents a greater challenge.

The present study is singularly placed to compare the relative difficulty of two of the learning scenarios discussed above: (i) the need to acquire a new sound category (that is, a new phoneme together with a new surface allophone), and (ii) the need to acquire a new mapping between surface and underlying representations (that is, a new phoneme for an already existing surface allophone).

1.2. Regional Dialects and L2 Speech Acquisition

The present study compares the perceptual behavior of two groups of Spanish-speaking learners of English. The two groups of learners differ in their region of origin; that is, they were brought up as speakers of two different geographical varieties of Mexican Spanish. The premise of our study is that the particular, specific L1 experience of L2 learners can determine, to some extent, the obstacles they encounter (and progression paths they take) when learning their L2. It follows that the phonology of the native dialect can modulate the acquisition of the phonology of the L2. A handful of recent studies have examined the potential role of regional dialect on L2 development. Some have explored the acquisition of different L2 dialects, that is, how people who speak the same language but are learning different varieties of the L2 differ in their linguistic behavior (Baker and Smith 2010; Escudero and Boersma 2004). Others have analyzed the potential effects of the native dialect on the acquisition of the L2, that is, how people who speak different varieties of the same language progress towards learning the same L2 (Chládková and Podlipský 2011; Escudero et al. 2012; Escudero and Chládková 2010; Mayr and Escudero 2010).

It has been demonstrated that people who speak the same native language but are exposed to different regional varieties of their L2 can face different cross-linguistic assimilation scenarios, leading to potentially different learning paths. For instance, Escudero and Boersma (2004) examined how two groups of Spanish-speaking learners of English perceived the English /i/–/ɪ/ contrast. One of the groups was learning English in Scotland, whereas the other was hypothesized to have been exposed mostly to the variety spoken in the South of England. This study found that learners in the two exposure groups behaved differently in their vowel categorization patterns. The authors attributed this to the acoustic properties of the particular target vowels involved, which led to different cross-linguistic assimilation patterns.

One's native dialect also seems to modulate the cross-linguistic assimilation patterns one will establish. For instance, Escudero et al. (2012) investigated the perceptual assimilation patterns displayed by Dutch-speaking L2 learners of English. It is well known that speakers of Dutch tend to have difficulties with the acquisition of the English /æ/–/ɛ/ contrast. Interestingly, Escudero and colleagues noted that, since the acoustics of the vowel categories of two regional varieties of Dutch (North Holland and Flanders) differ, it would be reasonable to predict diverging patterns of cross-linguistic assimilation for learners of English living in these two regions. Indeed, it was found that differences in the native-dialect phonetic system led to differences in cross-linguistic assimilation of vowels for learners in these regions. There is ample evidence that native phonology determines, to some extent, the learning paths of people acquiring a L2; evidently, the term "native phonology" refers to the individual phonological system of a given learner—their native phonological competence, which is based on their personal linguistic experience—and not to that of the "standard" dialect of a given learner's L1.

The present study compares the perceptual behavior of speakers of two regional varieties of Mexican Spanish when confronting an acquisitional challenge, the English /ʃ/–/tʃ/ contrast. Dialectological descriptions of the native phonology of speakers of these two regional varieties suggest differences that could lead to diverging patterns of cross-linguistic assimilation between their native consonants and those of English. This could lead to differences in their acquisitional obstacles and phonological learning paths.

1.3. Postalveolar Obstruents in Mexican Spanish

Dialectologists identify four regional varieties of Spanish in Mexico: central, coastal, northern, and peninsular (Yucatán) (Lope Blanch 1990; Martín Butragueño 2011, 2014; Moreno de Alba 1994). One of the phonological variables used to map the regional varieties of Mexican Spanish concerns the pronunciation of the first consonant in words such as *charco*, 'puddle', and *chamarra*, 'jacket.' In most dialects of Spanish, both in the Americas and in the Iberian Peninsula, this consonant is pronounced as a postalveolar affricate. This is true of most varieties of Mexican Spanish as well. Therefore, in most regions in Mexico, including the central highlands (the socially prestigious regional variety), *charco* 'puddle' is pronounced as [ˈtʃaɾko]. The speech of people born and raised in the northwestern Mexican states—including Sonora, Chihuahua, and Baja California Norte, among others—is characterized by a pattern of phonetic variation in which the postalveolar obstruent in *charco* 'puddle' may be pronounced as either [tʃ], an affricate, or [ʃ], a fricative. Therefore, in northwestern Mexican speech, *charco* 'puddle' is pronounced sometimes as [ˈtʃaɾko] and sometimes as [ˈʃaɾko]. Although variationist studies vary significantly in their reported fricativization rates ("fricativition" refers to the practice of pronouncing <ch> as [ʃ], a diachronic innovation), what seems clear is that, in this Spanish dialect, both the fricative and affricate variants of this variable are found (Alessi Molina and Díaz 1994; Brown 1989; Carreón Serna 2007; Méndez 2017).

Note that the use of the two variants of <ch> is not determined by a phonological rule; in other words, the two variants are not in complementary distribution, but in a scenario of "free" variation. In reality, the variation is not completely free: the investigations that have explored the phonetic variation that affects the pronunciation of <ch> have identified a number of social factors that may modulate to some extent variant choice. Among the social factors involved are age, level of education, and gender. Studies vary in their reported effects of gender (Carreón Serna 2007; Méndez 2017), and some claim that gender is meaningful only when it interacts with age (Jaramillo and Bills 1982). Age might also be relevant only when correlated with level of education (Jaramillo and Bills 1982). At any rate, what is important for our present purposes is that the alternation between [tʃ] as [ʃ] is not determined by a phonological process. The two sounds are neither contrastive nor in complementary distribution, since the same lexical item may be pronounced with either variant.

The constant exposure to the variation that affects <ch> has been found to affect northwestern, Mexican Spanish listeners' patterns of spoken word recognition. In a lexical access investigation, López Velarde and Simonet (2019) confirmed that listeners in northwestern Mexico are equally likely to accept Spanish word forms produced with either variant of <ch>. They also found that both variants are equally likely to prime listeners for the efficient recognition of spoken words (that is, [ˈʃaɾko] primes [ˈtʃaɾko] as much as [ˈtʃaɾko] does), which suggests that this group of listeners store Spanish words with both variants within the same abstract (or prototypical) mental representation. This study confirms that the two variants of <ch> are indeed allophones of the same phoneme. These findings suggest that people who experience sociophonetic variability in their speech community may store more than one phonetic variant in their long-term mental representation of words.

1.4. The Present Study

The current study focuses on a phonemic contrast of English—that between /tʃ/, as in *cheat*, and /ʃ/, as in *sheet*—and investigates the perceptual identification and discrimination patterns pertaining to this contrast displayed by two groups of L2 learners of English whose native language is Spanish. Our learner

sample was recruited from two dialectal regions, central and northwestern Mexico. More specifically, this study explores how speakers of northwestern Mexican Spanish, who are recurrently exposed to the sociophonetic variability that affects Spanish <ch> in their speech community, perceive the target English contrast (/ʃ/–/tʃ/), and how these perceptual habits differ (if at all) from those demonstrated by speakers of central Mexican Spanish, who lack experience with this specific variability pattern.

Even though we hypothesize that both of our target populations of L2 learners of English may find this contrast relatively difficult to master, we believe that the specific learning obstacles the two populations experience are different—and this, we hypothesize, results in different learning outcomes. On the one hand, we postulate that the English /ʃ/–/tʃ/ contrast will prove to be relatively difficult for central Mexican learners because [ʃ], as in *sheet*, does not correspond to any sound in their dialect of Spanish while [tʃ], as in *cheat*, does. It is possible that these learners assimilate the two phonemes of English to the same native category, /tʃ/. Nevertheless, since [tʃ] and [ʃ] are phonetically quite distinct, and since central Mexican Spanish has both affricates (/tʃ/) and fricatives (/s, f, h/), it could be the case that (adopting PAM's terminology) the English /ʃ/–/tʃ/ contrast presents a category-goodness (CG) assimilation pattern, one in which English /tʃ/ is assimilated to central Mexican Spanish /tʃ/ with a very high goodness of fit and English /ʃ/ also assimilates to this central Mexican Spanish phoneme but with a lower goodness of fit.

Northwestern Mexican Spanish speakers, unlike people from central Mexico, are exposed to two variants of <ch> in their native dialect, [tʃ] and [ʃ]. These are two allophones of the same phoneme. For this reason, we hypothesize that speakers of northwestern Mexican Spanish are likely to assimilate both English /tʃ/ and English /ʃ/ to the same native phoneme, and that the goodness of fit of these two cross-linguistic assimilation patterns is likely to be similarly high. This might create (adopting PAM's terminology again) a single-category (SG) assimilation pattern. Since CG assimilation patterns are expected to lead to better discriminability than SC ones (Best et al. 1988, 2001; Best and Tyler 2007), we hypothesize that the error rates in the identification and discrimination of our target English contrast will be larger for learners in northwestern Mexico than for central Mexican learners.

An alternative way to frame the learning scenario for the northwestern Mexican Spanish speakers is that of an allophonic split (Barrios et al. 2016b; Eckman et al. 2001). These learners must unlearn that both [tʃ] and [ʃ] are mapped onto the same phoneme (as they are in their native dialect of Spanish), and they must develop a new phoneme specific to the L2 to which only one of these two phonetic categories is mapped. In other words, speakers of northwestern Mexican Spanish must develop a new phonological, underlying category and remap their sound categories so that they are each assigned to a different contrastive unit. Are the northwestern Mexican Spanish speakers more or less likely than the central Mexicans to succeed in their discrimination of the members of English /ʃ/–/tʃ/ contrast? This is the fundamental research question that motivates the present study.

To compare the acquisition of our target English phonemic contrast with a contrast about which much is known, we selected a second phonemic contrast of English to serve as a control condition, that between /i/ and /ɪ/. It is well known that native speakers of Spanish find the *seat–sit* contrast very difficult to discriminate and, therefore, learn (Casillas 2015; Escudero and Boersma 2004; Kondaurova and Francis 2008; Morrison 2008, 2009); they also find the two members of the contrast very difficult to identify against each other. Therefore, the *seat–sit* contrast, tested in our experiments alongside our target contrast, *sheet–cheat*, serves as a control condition, one that should be similarly challenging for both of our target learner populations.

2. Method

2.1. Participants

The data were collected in two locations in Mexico, Hermosillo and Santiago de Querétaro, which are the two largest cities in the states of Sonora and Querétaro, respectively. The participants in Hermosillo were lifelong residents of the state of Sonora. The majority of the participants had lived

in Hermosillo from birth, and those born in other municipalities had moved to the city as children. Many of the participants tested in Santiago de Querétaro were not born in the city of Santiago, but they reported having moved there as children or as teenagers. The Querétaro residents in our sample who were not natives to Querétaro were born in other central states of the country, such as Guanajuato, Jalisco, Morelos, and Puebla. Particularly with respect to their treatment of /t∫/, as well as to that of many other sounds, the central highlands of Mexico form a single dialectal area. In sum, data were collected in two dialectal areas, the northwest (exemplified by Hermosillo, Sonora) and the central highlands (exemplified by Santiago de Querétaro, Querétaro).

A total of 88 people (44 from Sonora, 44 from Querétaro) participated in this study. Participants' ages ranged between 18 and 43 years old. All but five participants were college students at the time of testing, graduate or undergraduate. Three participants had graduated with a college degree, and two had not completed college and were working in the industry. The high number of college students or college graduates in our sample is due to our having recruited our participants in college settings, the *Universidad de Sonora* (Hermosillo) and the *Universidad Autónoma de Querétaro* (Santiago de Querétaro). The educational and professional profile of the participants is not fully representative of the general population native to these locations—highly educated people are overrepresented. The profile, however, might be representative of the narrower population, in these locations, who have learned English as a foreign language. At any rate, the social profile of the two dialectal groups does not differ—both groups consist of highly educated people who are learning English as a foreign language in a school setting. All participants study (or studied) English in college.

Participants responded to the Bilingual Language Profile questionnaire (Gertken et al. 2014) The questionnaire collects information regarding the listeners' linguistic background and L2 learning experience with a focus on attitudes, history, self-assessed proficiency, and daily usage of the two languages. The questionnaire produces a language dominance score along a spectrum centered around 0, which represents balanced bilingualism. The participants in our study are expected to be Spanish dominant; in our implementation of the survey, dominance in Spanish is captured with scores ranging between 0 and −218.

We also administered an English vocabulary-size test to assess the participants' English knowledge, the LexTALE. The LexTALE (www.lextale.com) is a standardized test designed to measure vocabulary size in language learners (Lemhöfer and Broersma 2012). To the extent that vocabulary size reflects overall knowledge of the language, the LexTALE provides an indicator of a person's knowledge of English. It seems reasonable to speculate that acquisition of phonemic, contrastive categories is based upon vocabulary knowledge (Simonet 2016). The test consists of 60 trials, comprising 40 English words and 20 nonwords, and these are presented to participants for them to make lexical decisions on. The resulting score is expressed in percent-correct units, and it is corrected for the unequal number of words and nonwords. In this study, the test was administered using PsychoPy 2 (Peirce et al. 2019). After responding to the BLP and the LexTALE, participants proceeded to complete the identification task followed by the discrimination task.

The two dialectal groups do not seem to differ with respect to their dominance scores, $t(85.1) = −0.608$, $p > 0.05$ [.54], 95% c.i. [−14.9, 7.9], Cohen's $d = −0.13$, but they do in regards to their English vocabulary size scores, $t(82.7) = −3.63$, $p < 0.001$ [0.0004], 95% c.i. [−11.86, −3.47], Cohen's $d = −0.77$. The average BLP score for the Querétaro group is −97.3 (*SD* = 28.3, *range* [−142.2, −18.1]), and the average for the Sonora group is −100.8 (*SD* = 25.5, *range* [−140.1, −41.5]). This confirms that all participants are dominant in Spanish. The average LexTALE score for the Querétaro group is 69.6 (*SD* = 10.8, *range* [53.7, 97.5]), and the average for the Sonora group is 61.9 (*SD* = 8.9, *range* [42.5, 81.25]). Thus, the Queretaroans have, on average, higher vocabulary size scores than the Sonorans, but there is much overlap between the two groups. On average, neither of the two groups are near ceiling (i.e., 90% or higher). In terms of their vocabulary size scores, both groups include a relatively wide range of learners.

2.2. Materials

2.2.1. Identification Experiment

The key data in this study were collected by means of two perception tasks, an identification task and a categorical discrimination task. In the identification task, participants were presented with 96 auditory stimuli consisting of one of four English words: *cheat*, *sheet*, *seat*, and *sit*. A total of 24 different iterations of each of the four words were played in random order to each participant. Listeners were asked to identify each stimulus by indicating, from a closed list of options, the lexical item each auditory stimulus corresponded to. Four options to select from were shown on a computer screen in alphabetical order, from left to right: *cheat*, *sheet*, *seat*, *sit*. The participants are hypothesized to misidentify *cheat* as *sheet*, *sheet* as *cheat*, *seat* as *sit*, and *sit* as *seat*. We have no hypothesis as to whether they would also misidentify *seat* or *sit* as *cheat* or *sheet*.

2.2.2. Discrimination Experiment

An ABX categorical discrimination task was designed to test two key contrasts: *sheet–cheat* (target) and *seat–sit* (control). In an ABX task, listeners hear a triad of auditory tokens (A, B, and X) presented in a sequence within the same trial and, upon hearing all three, they indicate whether the third token (X) matches either the first (A) or the second (B) item in the sequence. There were no "catch" trials in our version of the task, which means that there always was a correct answer. Importantly, all of the stimuli in each triad were acoustically different, including the two matching tokens, as each one of them had been recorded by a different talker. Under such conditions, comparisons cannot be based on acoustic memory, but must be based on phonological or lexical memory (participants are comparing abstract categories, not auditory tokens), which requires participants to access their mental representations to make their decisions. This is the reason why we refer to this task as a *categorical* ABX.

Each participant provided 48 observations to the data set: 24 trials focused on the *seat–sit* contrast (*sea*t–sit–seat [6]; seat–sit–sit [6]; sit–seat–sit [6]; sit–seat–seat [6]), and 24 focused on the sheet–cheat contrast (sheet–cheat–cheat [6]; sheet–cheat–sheet [6]; sheet–sheet–cheat [6]; sheet–cheat–cheat [6]). In 24 of the trials, the matching word was adjacent to the target word—it was in the second position. In other words, the target word was always in the third position of the sequence and, in cases of adjacency, the matching word was in second position. In 24 of the trials, the matching word was not adjacent to the target—it was in the first position. Everything else being equal, matching adjacent categories is expected to be easier than matching non-adjacent one (Best et al. 2001).

2.2.3. Auditory Stimuli

Four native English speakers, all of them women, served as talkers. Their productions were recorded in a sound-treated booth using professional recording equipment: a Shure SM10A head-mounted dynamic microphone and a Sound Devices USBPre2 audio interface connected to a laptop computer. Speech productions were digitized at 44.1 kHz, with 16-bit quantization. Sound files were normalized for intensity.

The talkers were asked to produce the target words by embedding them in a constant carrier phrase, "__ is the word." The materials were presented in random order to avoid any possible systematic effects of list intonation or exhaustion on the same lexical items. Talkers produced all target words four times (4 tokens × 4 iterations × 4 talkers = 64 items). One token of each target word per talker was selected (avoiding disfluencies and any extraneous noise) for a total of four target stimuli per lexical item.

2.3. Procedure

Participants completed the tasks individually. In Querétaro, participants were tested in a sound-attenuated booth, while in Sonora they were tested in a quiet library room. Stimuli were presented auditorily over a set of Audio Technica ATH-M50x closed-circumaural headphones connected to a laptop computer running PsychoPy2 (Peirce et al. 2019). Participants responded by pressing

a key on a Logitech G512 Lightsync RGB mechanical keyboard. Prior to the completion of the experimental tasks, the first author, a native Spanish speaker from Sonora, provided them with a general description of the tasks and their instructions. This conversation took place in Spanish. Before participating in any of the perceptual tasks, people completed the Bilingual Language Profile questionnaire (Gertken et al. 2014), then the LexTALE (Lemhöfer and Broersma 2012).

For the identification task, participants were instructed to listen to each stimulus, one per trial, and indicate their answer as quickly and accurately as possible by pressing one of four keys on the keyboard (1, 2, 3, or 4). Trials began with a red cross in the center of the screen for 250 ms, which was followed by a screen showing the four response options: *cheat*, *sheet*, *seat*, *sit*. Words were shown in capital letters. Numbers—that is, key codes—were presented in yellow and shown below their corresponding lexical item. Response options were shown for 2500 ms. Auditory stimuli were played 500 ms from the onset of the screen displaying the response options. Participants were allotted 2 s to enter a response. If participants did not provide a response within the allotted time, a new trial began, and the response was left empty.

In the ABX task, participants were asked to listen to all three sounds presented in the trial and only then respond by pressing either number 1 or number 2 on the keyboard to indicate whether they believed the third sound matched the first (1) or the second (2) one in the triad. The words 'first' and 'second' were shown on the computer screen in upper case and accompanied with their matching key codes, 1 or 2. Each trial began with the showing of a red cross in the center of the screen for 1000 ms. The first stimulus of the triad was played at the 1 s mark and was then followed by the second sound of the triad at the 2 s mark. The stimulus onset asynchrony of these two stimuli was thus set at 1 s. The stimulus onset asynchrony between the second and third stimuli was set at 1.5 s. Simultaneously with the playing of the third auditory stimulus in the triad, a screen showing the two response options was shown. Participants had 2 s to introduce their answer. If no answer was entered within this time, a new trial began, and the response was left empty.

2.4. Analysis

All statistical analyses were run in *R*, with packages *tidyverse* (Wickham 2017), *afex* (Singman et al. 2018), and *effsize* (Torchiano 2018). For reproducibility, readers may obtain the *R* scripts and synthetic data frames from either of the authors.

2.4.1. Identification

The analysis of the identification-task data was conducted in two steps. In the first step, we classified what participants heard (the lexical items the talkers had produced) as a function of what they responded (the lexical items the listeners had responded they had heard). This results in a contingency table. The original data set was comprised of 8448 rows, all of them listeners' responses to auditory stimuli. Nevertheless, a number of these observations were excluded from the analysis because the listener did not respond within their allotted time. There was a total of 424 not-responded-to trials, about 5% of the observations. The analysis was then conducted without such trials, with a data set containing 8024 observations.

In a second step, we simply calculated the proportion of times a given participant was accurate versus the times they were inaccurate. In order to prepare the data for the statistical analysis, we ran an arcsine transformation of the proportion-correct scores by participant and condition.

2.4.2. Discrimination

The original data set comprised a total of 4224 rows, 44 (listeners) × 2 (locations) × 48 (responses), of which 288 (6%) were empty, that is, trials that did not contain any information because the participant had failed to respond within the allotted time. An analysis of the participants' responses was then conducted after removing the empty observations from the data set, which results in a data frame comprised of 3936 observations. The analysis counted the proportion of correct responses per listener,

per condition (contrast type and adjacency). Then, the accuracy scores (or proportion of correct responses) were arcsine-transformed for the statistical analysis.

3. Results

3.1. Identification

The analysis of the identification data focuses on the proportion of times the auditory stimuli were identified as each of the four lexical items. Table 1 shows the proportion of responses, calculated only for the trials that were responded to, as a function of stimulus played (rows) and response given (column), further broken down by region-of-origin of the participants.

Table 1. Proportion of times each auditorily presented lexical item was identified as being an instance of one of four possible words (*cheat*, *sheet*, *seat*, *sit*), further broken down by region of origin, in Mexico, of the English learners (Sonora, Querétaro).

	Sonora				Querétaro			
	Cheat	*Sheet*	*Seat*	*Sit*	*Cheat*	*Sheet*	*Seat*	*Sit*
[t∫]*eat*	0.51	0.42	-	-	0.83	0.16	-	-
[∫]*eet*	0.39	0.57	-	-	0.20	0.73	-	-
s[i:]*t*	-	-	0.32	0.64	-	-	0.42	0.55
s[ɪ]*t*	-	-	0.64	0.32	-	-	0.55	0.40

Note: Rows represent the auditory stimuli played and columns represent the labels displayed on the screen and, thus, the responses available to the participants. Responses below 5% are not shown. Within each participant group, rows add up to 1.

As it may be observed in Table 1, the proportion values suggest that neither *seat* nor *sit* are likely to be categorized as neither *sheet* nor *cheat*. We may conclude that [s] is categorized as being distinct from both [t∫] and [∫], and that this is true for both groups of learners—when a word begins with [s] and ends with [t], only *seat* and *sit* are viable options. Equivalently, when a word begins with a postalveolar obstruent, either [t∫] or [∫] and end with [t], neither *seat* nor *sit* are viable options. We infer that it is reasonable for us to treat the *sheet–cheat* and the *seat–sit* contrasts as separate binary oppositions in our analysis. The scores in Table 1 also suggest that the identification of both *seat* and *sit* lead to a large number of categorization errors, and that both groups of learners are likely to confuse the two words with each other.

We now turn our attention to *cheat* and *sheet*. It appears that the identification patterns of these two words differ in the two groups of learners. In the case of Queretaroans, *cheat* and *sheet* do not appear to be very difficult to identify even in a task that plays these words in the context of each other—accuracy rates are relatively high, with 83% correct responses for *cheat* and 73% correct responses for *sheet*, but they can nevertheless be confused with each other at rates that are not negligible. On the other hand, the Sonorans made many categorization errors for *cheat* and *sheet*. Sonorans' accuracy rates are relatively low, with 51% correct responses for *cheat* and 57% for *sheet*.

Our statistical analysis focuses on accuracy rates. We select only the cells that may be interpreted as displaying 'correct' responses: [t∫]*eat* identified as *cheat*, [∫]*eet* identified as *sheet*, *s*[i:]*t* identified as *seat*, and *s*[ɪ]*t* identified as *sit*. This analysis ignores all other cells. The arcsine-transformed accuracy scores are analyzed with a mixed-design, two-way 2 × (4) ANOVA with Location (Querétaro, Sonora) as a between-subjects factor, and Item (*cheat*, *sheet*, *seat*, *sit*) as a within-subjects factor. The ANOVA yields main effects of Location, $F(1, 86) = 18.2$, $p < 0.0001$, $\eta^2 = 0.10$, and of Item, $F(2.13, 183.4) = 62.5$, $p < 0.0001$, $\eta^2 = 0.26$. It also yields a statistical interaction between the two factors, $F(2.13, 183.4) = 6.4$, $p < 0.05$ [.002], $\eta^2 = 0.03$. The results reveal that, as a group, the Sonorans are more likely to make categorization mistakes than the Queretaroans with this closed lexical set, but this further depends on the lexical item.

To investigate the interaction, we divide the data set into four subsets as a function of Item. The results reveal that, for both *seat* and *sit*, accuracy rates are comparable across the two learner groups: *sit*, $t(83.3) = -1.16$, $p > 0.05$ [0.25], 95% c.i. [−0.287, 0.075], Cohen's $d = -0.247$; *seat*, $t(66.5) = -1.97$, $p > 0.05$ [0.053], 95% c.i. [−0.252, 0.001], Cohen's $d = -0.42$. In other words, both groups of participants are similarly likely to be accurate when identifying these two words. On the other hand, accuracy rates are different across learner groups for both *cheat*, $t(83.9) = -8.08$, $p < 0.0001$, 95% c.i. [−0.501, −0.303], Cohen's $d = -1.722$, and *sheet*, $t(73.02) = -3.53$, $p < 0.001$ [0.0007], 95% c.i. [−0.316, −0.088], Cohen's $d = -0.753$. In both cases, the Queretaroans are more likely to be accurate than the Sonorans.

To summarize, identifying the two members of the *seat–sit* contrast appears to be similarly challenging for both groups of Spanish-speaking learners of Spanish, whereas asking participants to identify the two members of the *sheet–cheat* contrast is more likely to lead to errors for the Sonorans than for the Queretaroans. The results obtained in the identification task suggest the following hypotheses: (i) The Sonorans are just as likely as the Queretaroans to find the *seat–sit* contrast difficult to discriminate, and so we use this contrast as our control condition in the discrimination study; (ii) the Sonorans are likely to find the *sheet–cheat* contrast more difficult to discriminate than the Queretaroans.

3.2. Discrimination

Table 2 shows the untransformed proportion of correct responses by participant group and experimental condition. There are two experimental conditions in our design: (i) the lexical contrast tested in a given trial (*seat–sit*, *sheet–cheat*), and (ii) the adjacency condition between the target word and the matching one. When the matching stimulus is located in the first position in the triad, the matching and the target stimuli are not adjacent (*primacy* condition), whereas when the matching stimulus is found in the second position in the triad the matching and the target stimuli are adjacent (*recency* condition). Everything else being equal, recency trials are predicted to be easier to answer accurately than primacy ones, particularly for challenging contrasts (Best et al. 2001).

Table 2. Proportion of correct responses by learner group (Querétaro, Sonora), as a function of lexical contrast (*seat–sit*, /iː/-/ɪ/; *sheet–cheat*, /ʃ/-/tʃ/) and adjacency condition (primacy, recency).

	Sonora			**Querétaro**		
	Primacy	**Recency**	***M***	**Primacy**	**Recency**	***M***
/ʃ/-/tʃ/	0.48	0.64	0.56	0.61	0.74	0.68
/iː/-/ɪ/	0.79	0.82	0.81	0.80	0.85	0.83

Note: Primacy stands for trials in which target and matching stimuli are not adjacent; recency stands for trials in which target and matching stimuli are adjacent.

Firstly, we submit the arcsine-transformed proportion-correct scores to a mixed-design, three-way $2 \times (4) \times (2)$ ANOVA with Location (Querétaro, Sonora) as a between-subjects factor, and Contrast (*sheet–cheat*, *seat–sit*) and Adjacency (primacy, recency) as within-subjects factors. The ANOVA yields main effects of Location, $F(1, 86) = 5.7$, $p < 0.05$ [.02], $\eta^2 = 0.02$, Contrast, $F(1, 86) = 117.8$, $p < 0.0001$, $\eta^2 = 0.19$, and Adjacency, $F(1, 86) = 14.8$, $p < 0.001$ [.0002], $\eta^2 = 0.05$. Of these effects, the largest one is Contrast, then Adjacency, then Location. Importantly, there are two significant two-way interactions: Contrast by Adjacency, $F(1, 86) = 4.43$, $p < 0.05$ [.04], $\eta^2 = 0.009$, and Contrast by Location, $F(1, 86) = 6$, $p < 0.05$ [.02], $\eta^2 = 0.01$. There is no significant Location by Adjacency interaction and no significant three-way interaction.

The interactions are explored in three steps. Firstly, to explore the Contrast by Adjacency interaction, we average over Location and analyze the effects of adjacency for the two contrasts separately. This analysis pools the data for the two dialectal regions. According to a series of paired-sample t-tests, Adjacency does not trigger a significant effect for the *seat–sit* contrast, $t(87) = 2.05$, $p > 0.016$ [.04], 95% c.i. [0.002, 0.138], Cohen's $d = 0.199$, but it does for the *sheet–cheat* one, $t(87) = 3.9$, $p < 0.001$ [0.00019], 95% c.i. [0.084, 0.259], Cohen's $d = 0.396$. People are found to be less accurate in

primacy trials than in recency ones, but most obviously so in trials that test the *sheet–cheat* contrast than the other one. Secondly, to explore the Contrast by Location interaction, we average over Adjacency and analyze the effects of Contrast for the two learner groups separately. For both groups, the *sheet–cheat* contrast leads to significantly more response errors than the *seat–sit* contrast (Querétaro: $t(87) = -6.27$, $p < 0.0001$, 95% c.i. $[-0.27, -14]$, Cohen's $d = -0.595$; Sonora: $t(87) = -8.92$, $p < 0.0001$, 95% c.i. $[-0.397,$ $-0.252]$, Cohen's $d = -0.971$), but the effect is larger for the Sonorans than for Queretaroans. Finally, returning once more to the Contrast by Location interaction, the potential effects of Location for the two contrasts are analyzed separately. According to two Welch *t*-tests, there is no significant effect of Location for the *seat–sit* contrast, $t(172.5) = -0.616$, $p > 0.0125$ $[.54]$, 95% c.i. $[-0.108, 0.057]$, Cohen's $d = -0.093$, whereas the effect is statistically significant for the *sheet–cheat* contrast, $t(173.9) = -3.37$, $p < 0.0125$ $[.0009]$, Cohen's $d = -0.508$.

To summarize, both groups learners find both lexical contrasts relatively difficult to discriminate. Interestingly, the *sheet–cheat* contrast appears to be more challenging than the *seat–sit* contrast. Additionally, both groups are similarly accurate in their discrimination of the *seat–sit* contrast, which we are taking to be our control condition. The most important finding is that, for the *sheet–cheat* contrast, the Sonorans are more likely than the Queretaroans to make discrimination errors. From these analyses, one could infer that the discrimination of the /ʃ/–/tʃ/ contrast is more challenging for the Sonorans than it is for the Queretaroans. Recall, however, that, according to the LexTALE, the Queretaroans in our sample have, on average, a larger English vocabulary than the Sonorans. Thus, the finding could be due, to some extent, to an asymmetry in English vocabulary size rather than to their native dialect phonologies.

To address the possibility that vocabulary size, rather than native phonology, explains these findings, we select a subset of the 22 learners with the lowest LexTALE scores in the Querétaro group and the 22 learners with the highest LexTALE scores in the Sonora sample to form a subset comprising 44 learners, $\frac{1}{2}$ of the sample. In this subset, the vocabulary size of the Sonorans ($M = 68.6$, $SD = 5.9$) is larger than that of the Queretaroans ($M = 61$, $SD = 4.2$), according to a Welch *t*-test conducted on LexTALE scores, $t(37.8) = 4.89$, $p < 0.0001$, 95% c.i. $[4.47, 10.76]$, Cohen's $d = 1.47$. A mixed-design ANOVA with arcsine-transformed accuracy scores obtained in the discrimination task yields only main effects of Contrast, $F(1, 42) = 76.5$, $p < 0.0001$, $\eta^2 = 0.21$, and Adjacency, $F(1, 42) = 4.47$, $p < 0.05$ $[.04]$, $\eta^2 = 0.04$, and a Contrast by Adjacency interaction, $F(1, 42) = 5.91$, $p < 0.05$ $[.02]$, $\eta^2 = 0.02$, but no other main effects, and no further interactions. Importantly, there are no main effects of Location, $F(1, 42) = 0.023$, $p > 0.05$ $[.64]$, $\eta^2 = 0.002$, and Location does not interact with Contrast, $F(1, 42) = 1.92$, $p > 0.05$ $[.17]$, $\eta^2 = 0.007$. Overall, participants are more accurate in their discrimination of the *seat–sit* contrast ($M = 0.81$, $SD = 0.2$) than of the *sheet–cheat* contrast ($M = 0.59$, $SD = 0.24$). They are also more accurate in recency trials ($M = 0.75$, $SD = 0.23$) than in primacy trials ($M = 0.65$, $SD = 0.25$). Additionally, the effects of adjacency are more robust in the *sheet–cheat* contrast than in the control contrast. In sum, an analysis of a subset of data in which the Sonorans have larger English vocabularies than the Queretaroans fails to reveal any differences between the groups in regard to their discrimination of the *sheet–cheat* contrast (or the *seat–sit* one, for that matter). Apparently, for the Sonorans to match the Queretaroans in their discrimination of the *sheet–cheat* contrast, they must have larger vocabularies than them.

In a final analysis, we explore the potential effects of vocabulary size on the discrimination of the target contrast by means of two linear regression models. These analyses address the following question: Do learners with larger vocabularies show increased sensitivity to the target consonant contrast? To conduct these comparisons, we select one of the experimental conditions—the most "difficult" one, the arcsine-transformed *sheet–cheat* contrast in primacy trials—so that we obtain a single accuracy score per participant, and then correlate this variable with the learners' LexTALE scores. We conduct two regression analyses, one per participant group. Whereas the regression model analyzing the Querétaro data yields a significant finding, $F(1, 42) = 11.2$, $p = 0.002$, adjusted $R^2 = 0.191$, the one analyzing the Sonorans does not, $F(1, 42) = 0.03$, $p < 0.001$, adjusted $R^2 = -0.023$. In other

words, the Queretaroans with larger vocabularies seem to be more sensitive to the *sheet–cheat* contrast than the ones with smaller vocabularies, whereas no such relation exists for the Sonorans. According to a series of one-sample *t*-tests, in this particular experimental condition, the Sonorans (as a group) are not found to have accuracy rates higher than chance (0.5 proportion-correct scores), $t(43) = -0.46$, $p > 0.025$ [0.64], 95% c.i. [0.69, 0.84], while the Queretaroans are, $t(43) = 3.03$, $p < 0.025$ [.002], 95% c.i. [0.84, 1.007].

4. Discussion

4.1. Summary of Findings

The present study reported on two categorization experiments aimed at investigating the acquisition of the English /ʃ/–/tʃ/ (*sheet–cheat*) contrast by two groups of foreign-language learners, both with Spanish as their L1. The first group was recruited in central Mexico (Querétaro), where the local dialect uses [tʃ] but not [ʃ]. The second group was recruited in northwestern Mexico (Sonora), where the local dialect uses both [tʃ] and [ʃ] in free variation. Alongside the target contrast, we investigated the English /i/–/ɪ/ contrast as a control condition.

The findings of a perceptual identification experiment indicated that the learners found the members of the /ʃ/–/tʃ/ contrast difficult to identify when played in the context of each other. On average, 29% of the relevant stimuli were identified incorrectly. Interestingly, the Sonorans (40% mean error rate) made significantly more errors than the Queretaroans (18% mean error rate). The findings of a categorical discrimination experiment demonstrated that the learners found the members of the target contrast hard to discriminate from each other. Overall, the mean error rate for the /ʃ/–/tʃ/ contrast was 38%. Importantly, the groups differ from each other in their discrimination of the /ʃ/–/tʃ/ contrast. In particular, the average error rate of the Sonorans in discrimination of this contrast was as high as 44% while that of the Queretaroans was 32%.

The learner groups differed from each other in their identification and discrimination accuracy of the /ʃ/–/tʃ/ contrast, but not in that of the /i/–/ɪ/ contrast. Taken together, the findings of these two experiments suggest that the Sonorans find the acquisition of the English /ʃ/–/tʃ/ contrast more difficult than the Queretaroans.

4.2. Interpretation and Implications

The experimental evidence suggests that the English /i/–/ɪ/ contrast is particularly difficult for native Spanish speakers who are learning English as a L2 (Escudero and Boersma 2004; Kondaurova and Francis 2008; Morrison 2008, 2009). Studies demonstrate that Spanish-speaking learners of English tend to associate both of the English vowels to their native /i/, perhaps initially merging the three phonetic categories into one. Spanish speakers can certainly overcome this initial obstacle and can acquire this contrast. Nevertheless, even when they do, their representation of the /i/–/ɪ/ distinction is likely to be based on duration, not spectrum; that is, whereas native, monolingual speakers of English distinguish these two vowel categories based on their spectral properties (i.e., first- and second-formant frequencies), native Spanish speakers who have been able to acquire this contrast are likely to base their distinction on a seemingly parasitic correlate, duration (Kondaurova and Francis 2008; Morrison 2008). The results pertaining to both of our participant groups demonstrate that the identification and discrimination of items implementing the /i/–/ɪ/ contrast are challenging for this population.

Let us now focus on an unexpected secondary finding pertaining to the control contrast. It seems that learners were more likely to be accurate in the discrimination task than in the identification one. For the /i/–/ɪ/ contrast condition, the participants were above chance in their discrimination patterns while they were at chance (as a group) when asked to identify *sit* and *seat*. In fact, some participants were more likely to be wrong than right in the identification experiment. One possible interpretation is that our learners have been able to develop separate phonetic categories for /i/ and /ɪ/ while not having learned which word has which sound category. The results of the discrimination task suggest that

participants can distinguish the two categories, though neither perfectly nor consistently; those of the identification task, on the other hand, suggest that participants do not associate the phonetic categories with lexical items. Insofar as a phoneme is a phonetic category associated to a particular lexical set (i.e., a category included in a phonolexical representation) (Simonet 2016), one could say that our participants may have developed (fuzzy) separate phonetic, but not *phonemic*, categories for /i/ and /ɪ/. The input they have received may have allowed them to form two phonetic categories, perhaps by means of distributional acoustic learning (Escudero and Williams 2014; Wanrooij et al. 2013); it may not have been sufficient, however, for them to form accurate, detailed phonolexical representations that include those categories. Indeed, experimental evidence suggests that forming phonetic categories in a L2 and associating them with phonolexical representations involve different stages of learning (Amengual 2016; Díaz et al. 2012; Sebastián-Gallés and Díaz 2012).

One factor that may have affected participants' identification patterns in the present study has to do with sound-to-spelling correspondences—recall that participants were asked to identify auditory stimuli in terms of visual ones. In Spanish, the phoneme-to-grapheme correspondence of /i/ consistently matches <i>. Note that the English lexical items chosen for the present study, *seat* and *sit*, had <i> corresponding with /ɪ/, not /i/. It is possible that the learners were variably affected by the Spanish spelling conventions during their identification of *sit*. This, together with their not having developed separate *phonemic* categories for these sounds, could explain the pattern of results. At any rate, the crucial finding pertaining to the /i/–/ɪ/ contrast is that the performance of the two groups of participants in the current study was comparable. We are justified to consider this contrast our control condition.

We now discuss the findings pertaining to the target contrast, /ʃ/–/tʃ/. We hypothesized that the perceptual behavior of our two groups of learners would differ in terms of their categorization of the English /ʃ/–/tʃ/ contrast. This hypothesis was based on the characteristics of the phonology of their native dialect of Spanish. An interesting finding was that both groups of participants seemed to find the English contrast somewhat challenging to acquire. This was not unexpected, as no variety of Spanish has a comparable contrast. The Queretaroans speak a variety of Spanish that uses [tʃ], but not [ʃ], whereas the Sonorans speak a variety that uses both [tʃ] and [ʃ] in free (i.e., not phonologically conditioned) variation. Whereas, as mentioned, all participants found the target contrast somewhat challenging to discriminate, the Queretaroans were found to be, on average, more accurate in their perceptual performance involving this particular contrast than the Sonorans. We propose that the acquisitional obstacles encountered by these two groups of learners, and thus their learning paths, differ on substantial grounds.

Let us first discuss the case of the Queretaroans. As mentioned in the Introduction, most current models of L2 phonological acquisition postulate that learners develop connections between the sound categories of their L2 and those of their L1. One such model is the PAM-L2 (Best and Tyler 2007) and another is the L2LP (Escudero 2005; van Leussen and Escudero 2015). The Queretaroans in our sample may have assimilated both English categories, /tʃ/ and /ʃ/, to the closest category in their L1, /tʃ/. Since Spanish has /tʃ/ but also has voiceless fricatives (/s f h/, just not /ʃ/), one could hypothesize that the acquisitional obstacle encountered by these learners is not insurmountable, provided that these learners extrapolate this aspect of their native system. Speakers of this dialect possess the capacity to represent affricates as being distinct from fricatives. We postulate that the Queretaroans in our sample have assimilated both English /tʃ/ and /ʃ/ to the same native phoneme but, crucially, have done so at different degrees of goodness of fit. Thus, whereas English /tʃ/ may be strongly assimilated to Spanish /tʃ/ (i.e., the match is close to perfect), the interlingual assimilation of English /ʃ/ to Spanish /tʃ/ may be rather poor (i.e., the match is imperfect). In the terminology used in the PAM, this would be an instance of CG assimilation. In cases of CG assimilation, the perceptual discrimination of the two members of the L2 contrast is expected to range from moderate to good. The Queretaroans' discrimination of the English /ʃ/–/tʃ/ contrast is indeed moderate. In sum, we believe that the obstacle the Queretaroans face when learning the English /ʃ/–/tʃ/ contrast is having to develop a new phonetic category, separating

from an initial stage in which the two sounds as categorized as instances of the same sound. Postulating that the Queretaroans assimilate both English /tʃ/ and /ʃ/ to the same native phoneme, but at different degrees of fit, explains both the presence of the obstacle and its size (i.e., moderate).

Let us now discuss the case of the Sonorans. In Sonora, as well as in other northwestern Mexican states, the local Spanish dialect uses both [ʃ] and [tʃ] in free variation. Crucially, these phonetic variants are not in complementary distribution, but occupy the same segmental slots in the same lexical items. The obstacle for these particular learners lies in the fact that they must unlearn the phonological mapping patterns of their native dialect—the mapping between phonetic categories (or allophones) and phonolexical representations (or phonemes)—before they can learn those of English. Assuming that in the first stage of L2 acquisition learners transfer their L1 competence into their L2 system (Escudero 2005; van Leussen and Escudero 2015), Sonorans begin their development at a stage in which [ʃ] and [tʃ] are equivalent at some level of representation. Therefore, they must first unlearn that [ʃ] and [tʃ] are variants of the same phoneme before they can learn that these two sounds are contrastive in English and, thus, associated with different lexical sets. The acquisitional obstacle may be formalized in two ways.

The first way in which the acquisitional obstacle encountered by the Sonorans may be formalized makes use of the same theoretical constructs we have used to explain the behavior of the Queretaroans, interlingual phonetic category assimilations (Best and Tyler 2007; Escudero 2005). It appears that the lexical distribution of phonetic categories determines to some extent peoples' perceptual behavior. In particular, phonetic categories in complementary distribution are less likely to be perceived as being distinct from each other than categories that are contrastive in the lexicon (Johnson and Babel 2010). If two sound categories are in free variation, it is even more likely that they will be perceived as being perceptually very similar to each other. Sonorans, therefore, are likely to perceive [ʃ] and [tʃ] as being perceptually more similar (to each other) than speakers of other Spanish dialects. If we extrapolate this to interlingual interactions in L2 acquisition, we postulate that Sonorans have assimilated both English /tʃ/ and /ʃ/ to the same native phoneme, and that both English categories are optimal matches to this native phoneme. The situation for the Sonorans could be one of SC assimilation (Best et al. 2001; Best and Tyler 2007), a *new scenario* (Escudero 2005). In cases of SC assimilation, discriminability is predicted to be very poor. Indeed, the discrimination of the English /ʃ/–/tʃ/ contrast by the Sonorans is very poor.

The second way in which the acquisitional obstacle encountered by the Sonorans may be formalized makes use of the concept of *allophonic split* (Barrios et al. 2016b; Eckman et al. 2001). This formalization does not depend on interlingual perceptual assimilations between phonetic categories but makes use of the concept of mapping between surface and underlying representations. Surface allophones that find themselves in either free variation or complementary distribution are mapped onto a single underlying segment, a phoneme. A Sonoran learner of L2 English would need to establish new mappings between familiar phones (Barrios et al. 2016b). The phonological competence of Sonoran Spanish speakers includes both [ʃ] and [tʃ], but, since the two sounds are mapped onto the same phoneme, learning the English /ʃ/–/tʃ/ contrast would require creating a new underlying representation and mapping the two surface allophones to separate phonemes. Eckman et al. (2001) hypothesize that this scenario is particularly challenging for L2 learners, but Barrios et al. (2016b) did not find evidence to support this claim. Insofar as we can conceptualize the learning scenario of Sonoran learners as a case of allophonic split, our data are fully in line with Eckman's hypothesis. Interestingly, our data suggest that cases of allophonic split (Sonora) are more challenging to overcome than cases of new category formation (Querétaro).

Since Sonoran learners possess a phonological system that includes both [ʃ] and [tʃ], one could have hypothesized that acquiring the English /ʃ/–/tʃ/ contrast would be particularly easy in their case. Obviously, this is not what our data suggest. Our data suggest that, in addition to existing phonetic categories, lexico-distributional patterns (i.e., the patterns of lexical distribution of phones,

which determine contrastivity, among other things) determine, to some extent, the significance of acquisitional obstacles.

Recall that our two participant samples differed, not only in their region of origin, but also in the sizes of their English vocabularies. We found that, on average, our Queretaroans had slightly larger vocabularies in English than our Sonorans. One could attribute the difference between the two participant groups in terms of their perceptual behavior with respect to the /ʃ/–/tʃ/ contrast to their overall proficiency in English. The Queretaroans may have been more accurate than the Sonorans in their discrimination of the /ʃ/–/tʃ/ contrast because they may be generally more proficient in English than the Sonorans. While this is certainly possible, we would like to highlight the following. Firstly, along with our main data set, we compared the behavior of Sonorans and Queretaroans in a controlled subset. For this subset, we selected 10 Sonorans and 10 Queretaroans so that the Sonorans had larger vocabularies than the Queretaroans. In this subset, participants' perceptual behaviors with respect to the /ʃ/–/tʃ/ contrast were found to be indistinguishable. Sonorans with larger vocabularies discriminate the /ʃ/–/tʃ/ contrast just as poorly as Queretaroans with smaller English vocabularies—thus, vocabulary size is not the only determinant of /ʃ/–/tʃ/ discrimination. Secondly, the control contrast condition, /i/–/ɪ/, led to comparable behavior across the two groups—whereas vocabulary sizes may differ between groups, their overall state of phonological development may not. Thirdly, Queretaroans, but not Sonorans, seem to become "better" at discriminating the /ʃ/–/tʃ/ contrast as their vocabulary increases. These observations suggest that the difference between Sonorans and Queretaroans reported in the present study is indicative of a larger issue, such as a difference in their native phonologies, not simply a difference in overall English competency. Nevertheless, only future research can resolve this conundrum.

5. Conclusions

The present study investigated the significance of two types of acquisitional obstacles in L2 phonology. The study reported on the identification and discrimination of the English /ʃ/–/tʃ/ (*sheet–cheat*) contrast by two groups of learners whose L1 is Spanish. The first group was recruited in central Mexico, where the local dialect uses [tʃ] but not [ʃ]. To learn the English /ʃ/–/tʃ/ contrast, speakers of this Spanish variety must learn a new phonetic category, [ʃ]. The second group was recruited in northwestern Mexico, where the local dialect uses both [tʃ] and [ʃ] in free variation. Since both obstruents are variants of the same phoneme in this variety, to learn the English /ʃ/–/tʃ/ contrast, speakers of this dialect must develop new mapping between familiar phonetic categories and underlying (contrastive) representations. The study found that the acquisitional obstacle encountered by speakers of the northwestern Mexican variety of Spanish is of a larger magnitude than the one encountered by speakers from central Mexico. Native dialect phonology is a powerful determinant of L2 phonology learning paths.

References

Alessi Molina, María T., and Ana Luisa Torres Díaz. 1994. Aspectos fonéticos del habla sonorense. In *Estudios de Lingüística y Sociolingüística*. Edited by Gerardo López Cruz and José Luis Moctezuma Zamarrón. Hermosillo: Universidad de Sonora, pp. 285–92.

Amastae, Jon. 1996. *Variación y Cambio en el Español de Ciudad Juárez, México*. Ciudad Juárez: Universidad Autónoma de Ciudad Juárez.

Amengual, Mark. 2016. The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics* 37: 1221–51. [CrossRef]

Baker, Wendy, and Laura Catharine Smith. 2010. The impact of L2 dialect on learning French vowels: Native English speakers learning Quebecois and European French. *Canadian Modern Language Review* 66: 711–38. [CrossRef]

Barrios, Shannon L., Anna M. Namyst, Ellen F. Lau, Naomi H. Feldman, and William J. Idsardi. 2016a. Establishing new mappings between familiar phones: Neural and behavioral evidence for early automatic processing of nonnative contrasts. *Frontiers in Psychology* 7: 1–16. [CrossRef] [PubMed]

Barrios, Shannon L., Nan Jiang, and William J. Idsardi. 2016b. Similarity in L2 Phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research* 32: 367–95. [CrossRef]

Best, Catherine T., and Michael D. Tyler. 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 Ocke-Schwen Bohn and Murray M. Munro. Amsterdam: John Benjamins, pp. 13–34.

Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345–60. [CrossRef] [PubMed]

Best, Catherine T., Gerald W. McRoberts, and Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109: 775–94. [CrossRef]

Bohn, Ocke-Schwen. 2017. Cross-language and second language speech perception. In *The Handbook of Psycholinguistics*. Edited by Eva M. Fernández and Helen Smit Cairns. Hoboken: John Wiley & Sons, pp. 213–39.

Broselow, Ellen, and Yoonjung Kang. 2013. Phonology and speech. In *The Cambridge Handbook of Second Language Acquisition*. Edited by Julia Herschensohn and Martha Young-Scholten. Cambridge: Cambridge University Press, pp. 529–54.

Brown, Dolores. 1989. El habla juvenil de Sonora, México: La fonética de 32 jóvenes. *Nueva Revista de Filología Hispánica* 37: 43–82. [CrossRef]

Carreón Serna, Yadira Aimé. 2007. La Lenición en el Español del Norte de México y Otras Lenguas Romances: Una Aproximación a la Variación Sociolingüística y un Modelo Formal Basado en la Teoría de la Optimidad. Ph.D. dissertation, University of California, Santa Barbara, CA, USA, September.

Casillas, Joseph V. 2015. Production and perception of the /i/-/ɪ/ vowel contrast: The case of L2 dominant early learners of English. *Phonetica* 72: 182–205. [CrossRef]

Casillas, Joseph V., and Miquel Simonet. 2016. Production and perception of the English /æ/-/ɑ/ contrast in switched-dominance speakers. *Second Language Research* 32: 171–95. [CrossRef]

Chang, Charles B. 2019. The phonetics of second language learning and bilingualism. In *The Routledge Handbook of Phonetics*. Edited by William F. Katz and Peter F. Assmann. New York: Routledge, pp. 427–47.

Chládková, Kateřina, and Václav Jonáš Podlipský. 2011. Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America* 130: EL186–EL192. [CrossRef] [PubMed]

Colantoni, Laura, Jeffrey Steele, and Paola Escudero. 2015. *Second Language Speech: Theory and Practice*. Cambridge: Cambridge University Press. [CrossRef]

Davidson, Lisa. 2017. Cross-language speech perception and production. In *Oxford Bibliographies in Linguistics*. Edited by Mark Aronoff. Oxford: Oxford University Press. [CrossRef]

Díaz, Begoña, Holger Mitterer, Mirjam Broersma, and Núria Sebastián-Gallés. 2012. Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences* 22: 680–89. [CrossRef]

Eckman, Fred R. 2012. Second language phonology. In *The Routledge Handbook of Second Language Acquisition*. Edited by Susan M. Gass and Alison Mackey. New York: Routledge, pp. 91–105.

Eckman, Fred R., Abdullah Elreyes, and Gregory K. Iverson. 2001. Allophonic splits in L2 phonology: The question of learnability. *International Journal of English Studies* 1: 21–51.

Escudero, Paola. 2005. *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization.* Ph.D. dissertation, Utrecht University, Utrecht, The Netherlands. LOT Dissertation Series 113; Amsterdam: Landelijke Onderzoekschool Taalwetenschap.

Escudero, Paola, and Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language* 26: 551–85. [CrossRef]

Escudero, Paola, and Kateřina Chládková. 2010. Spanish listeners' perception of American and Southern British English vowels. *Journal of the Acoustical Society of America* 128: EL254–EL259. [CrossRef]

Escudero, Paola, and Daniel Williams. 2014. Distributional learning has immediate and long-lasting effects. *Cognition* 133: 408–13. [CrossRef]

Escudero, Paola, Ellen Simon, and Holger Mitterer. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics* 40: 280–88. [CrossRef]

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. York: York Press, pp. 229–73.

Flege, James E., and Ocke-Schwen Bohn. 1989. An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition* 11: 35–62. [CrossRef]

Flege, James E., Murray J. Munro, and Robert Allen Fox. 1994. Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America* 95: 3623–41. [CrossRef]

Flege, James E., Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25: 437–70. [CrossRef]

Gertken, Libby M., Mark Amengual, and David Birdsong. 2014. Assessing language dominance with the Bilingual Language Profile. In *Measuring L2 Proficiency: Perspectives from SLA*. Edited by Pascale Leclerq, Amanda Edmonds and Heather Hilton. Bristol: Multilingual Matters, pp. 208–25.

Hualde, José Ignacio. 2005. *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Jaramillo, June A., and Garland D. Bills. 1982. The phoneme /ch/ in the Spanish of Tomé, New Mexico. In *Bilingualism and Language Contact: Spanish, English, and Native American Languages*. Edited by Florence Barkin, Elizabeth A. Brandt and Jacob Ornstein-Galicia. New York: Teachers College Press, Columbia University, pp. 154–65.

Johnson, Keith, and Molly Babel. 2010. On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* 38: 127–36. [CrossRef]

Kondaurova, Maria V., and Alexander L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124: 3959–71. [CrossRef]

Lado, Robert. 1957. *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: University of Michigan Press.

Lemhöfer, Kristin, and Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44: 325–43. [CrossRef]

Lope Blanch, Juan M. 1990. *Atlas Lingüístico de México*. 6 vols. México City: El Colegio de México-UNAM-FCE.

López Velarde, Mariela, and Miquel Simonet. 2019. Spoken word recognition and shesheo in Northwestern Mexico. In *Recent Advances in the Study of Spanish Sociophonetic Perception*. Edited by Whitney Chappell. Amsterdam: John Benjamins, pp. 266–85.

Martín Butragueño, Pedro. 2009. *Regularidad y Excepcionalidad del Cambio Lingüístico: El Caso de* ⁽ᵗ⁾ *en la Geografía Fónica de México*. México City: El Colegio de México.

Martín Butragueño, Pedro. 2011. Dialectología de nuevos mundos: Una lectura variacionista del Atlas Lingüístico de México. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*. Edited by Luis A. Ortiz-López. Somerville: Cascadilla Proceedings Project, pp. 17–32.

Martín Butragueño, Pedro. 2014. La división dialectal del español mexicano. In *Historia Sociolingüística de México*. Edited by Pedro Martín. México, City: El Colegio de México, pp. 1353–408.

Mayr, Robert, and Paola Escudero. 2010. Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition* 13: 279–97. [CrossRef]

Méndez, Luis Alberto. 2017. The variant [ʃ] in the Spanish of Ciudad Juárez. *Borealis—An International Journal of Hispanic Linguistics* 6: 243–60. [CrossRef]

Moreno de Alba, José G. 1994. *La Pronunciación del Español en México*. México City: El Colegio de México.

Morrison, Geoffrey Stewart. 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. [CrossRef]

Morrison, Geoffrey Stewart. 2009. L1-Spanish speakers' acquisition of the English/i/—/ɪ/ contrast II: Perception of vowel inherent spectral change. *Language and Speech* 52: 437–62. [CrossRef] [PubMed]

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51: 195–203. [CrossRef] [PubMed]

Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [CrossRef]

Sebastián-Gallés, Núria, and Begoña Díaz. 2012. First and second language speech perception: Graded learning. *Language Learning* 62: 131–47. [CrossRef]

Serrano Morales, Julio C. 2000. Contacto dialectal (¿y cambio lingüístico?) en español: El caso de la /tʃ/ sonorense. In *Estructuras en Contexto: Estudios de Variación Lingüística*. Edited by Pedro Martín Butragueño. México City: El Colegio de México, pp. 45–59.

Serrano Morales, Julio C. 2009. Existe el noroeste mexicano como zona dialectal. In *Lengua, Literatura y Región*. Edited by Everardo Mendoza Guerrero, Maritza López Berríos and Ilda Elizabeth Moreno Rojas. Culiacán: Universidad Autónoma de Sinaloa.

Simonet, Miquel. 2016. The phonetics and phonology of bilingualism. In *Oxford Handbooks Online*. Oxford: Oxford University Press. [CrossRef]

Singman, Henrik, Ben Bolker, Jake Westfall, and Frederik Aust. 2018. Afex: Analysis of Factorial Experiments (Version 0.22-1). Computer Software. Available online: https://CRAN.R-project.org/package=afex (accessed on 4 June 2020).

Torchiano, Marco. 2018. Effsize: Efficient Effect Size Computation (Version 0.7.4). Computer Software. Available online: https://CRAN.R-project.org/package=effsize (accessed on 4 June 2020).

Van Leussen, Jan-Willem, and Paola Escudero. 2015. Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology* 6: 1–12. [CrossRef] [PubMed]

Wanrooij, Karin, Paola Escudero, and Maartje E.J. Raijmakers. 2013. What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics* 41: 307–19. [CrossRef]

Wickham, Hadley. 2017. Tidyverse: Easily Install and Load the "Tidyverse" (Version 1.2.1). Computer Software. Available online: https://CRAN.R-project.org/package=tidyverse (accessed on 4 June 2020).