Individual differences in the acquisition of language-specific and dialect-specific allophones of intervocalic /d/ by L2 and heritage Spanish speakers studying abroad in Sevilla

Brendan Regan*
Texas Tech University, Lubbock, TX, USA
*Corresponding author. E-mail: brendan.regan@ttu.edu

Abstract
This study examines the role of language proficiency and other individual factors (attitudes, input) in the acquisition of language-specific [ð] and dialect-specific [∅] allophones of Spanish intervocalic /d/ in the /ado/ context by L2 and heritage Spanish speakers during a short-term study abroad in Sevilla, Spain. Twenty L2-intermediate, 10 L2-advanced, and 10 Heritage-advanced Spanish speakers completed a reading task at the beginning and the end of the program. Based on an acoustic analysis, a mixed-effects linear regression model found that only L2-advanced and Heritage-advanced groups demonstrated more approximant-like [ð] over time. However, proficiency level interacted with attitudes and input. There were a few [∅] realizations, mostly produced by an L2-advanced speaker who also demonstrated metalinguistic awareness of the dialect-specific allophone. The findings imply that advanced (L2 and heritage) speakers with favorable attitudes toward the local variety are most likely to demonstrate gradient language-specific allophonic changes during a short-term SA program.

Introduction
There has been a growing interest in examining the effect of study abroad on second language acquisition (Kinginger, 2009; Pérez-Vidal, 2014; Regan et al., 2009; Sanz & Morales-Front, 2018). The fundamental question underlying this research is if the study abroad (henceforth SA) immersion context provides greater benefits to acquiring a second language than traditional classroom learning at home (henceforth AH) (Díaz-Campos, 2004; Lafford & Collentine, 2006; Segalowitz & Freed, 2004). While studies have investigated the acquisition of morphology, syntax, lexicon, pragmatics, and fluency, there is a relative paucity of studies analyzing phonetic and phonological acquisition in the SA context (Bongiovanni et al., 2015, p. 244; Solon & Long, 2018, p. 71). Recently there has been an increase in studies of phonetic acquisition during SA, but the role of linguistic proficiency remains understudied (Solon & Long, 2018, p. 75).
Furthermore, there lack studies that include heritage speakers in the SA context (Pozzi et al., 2021).

Building on previous studies, this study addresses these gaps utilizing the experimental methods of laboratory phonology and the statistical practices of variationist sociolinguistics. The aims were twofold: (1) to examine the effect of a short-term SA in Sevilla, Spain by L2 and heritage Spanish speakers on acquiring language-specific [ð] and/or dialect-specific [∅] for intervocalic /d/; and, (2) to analyze how this varies by proficiency level and other individual factors.

**Background**

**Acquisition of language-specific phonemes/allophones during SA**

Overall studies have shown benefits for L2 speakers’ production of language-specific phones in the SA context. For example, several studies have demonstrated more targetlike production of vowels during SA by L2 German (L1 English) speakers (O’Brien, 2003) and L2 Spanish (L1 English) speakers (Stevens, 2011), although Avello and Lara (2014) found no gains among L2 English (L1 Spanish-Catalan) speakers. Studies on rhotics among L2 Spanish (L1 English) speakers reported more gains for SA than AH students in the production of /r/ (Detrixhe, 2015) and /ɾ/ (Bongiovanni et al., 2015).

Several studies, however, have demonstrated contrasting findings, such as in the production of intervocalic /b, d, g/, realized as stop [b, d, g] in English and approximant [β, δ, χ] in Spanish, by L2 Spanish (L1 English) speakers. Specifically, Lord (2010) found more approximant realizations for /b, d, g/, Bongiovanni et al. (2015) only found gains for /d/, and Díaz-Campos (2004) found no gains over time for any segment. For word-initial /p, t, k/, realized as [p, t, k] in Spanish and [ph, th, kh] in English, among L2 Spanish (L1 English) speakers, Díaz-Campos (2004) found that speakers reduced their voice onset time (VOT)\(^1\), while Bongiovanni et al. (2015) only found gains for /p/ and /k/. Neither study found significant differences between the SA and AH groups. However, Llanes et al. (2017) found that L2 English (L1 Spanish-Catalan) speakers in the SA context increased their VOT (more English-like [ph, th, kh]) than AH speakers. Contrasting findings may be due to methodological differences between studies (i.e., number of participants, tasks, number of tokens, acoustic vs. auditory analysis). These differences notwithstanding, studies generally show more benefits of learning an L2 phone in SA than AH contexts or overall benefits of learning an L2 phone in SA contexts (without AH comparisons).

**Acquisition of dialect-specific phonemes/allophones during SA**

There has been an emergent body of literature interested in the acquisition of dialect-specific phones\(^2\) during SA. Although it depends on a variety of individual factors, in general studies have found that some dialect features are more acquired than others. For example, studies have shown low production of the Castilian Spanish voiceless interdental fricative /θ/ (George, 2014; Knouse, 2012; Ringer-Hilfinger, 2012) and the

---

\(^1\)VOT is an acoustic parameter to distinguish stops based on the duration (ms) between the release of the stop consonant and the onset of voicing (i.e. the vibration of the vocal cords), distinguishing between voice lead, short voice lag, and long voice lag (Lisker & Abramson, 1964).

\(^2\)See Henriksen et al. (2010) and Trimble (2013) for acquisition of regional intonation patterns.
Castilian Spanish voiceless uvular fricative \( \chi \) (George, 2014) by L2 Spanish (L1 English) speakers in central Spain. Another variable that demonstrates relatively low rates of acquisition is French /l/ deletion in third-person subject clitic pronouns, specifically “il” (i.e., il fa\’it beau ‘it’s beautiful outside’ [il.f.e.bo] vs. [i.f.e.bo]), produced by L2 French (L1 English) speakers in France (Howard et al., 2006; Kennedy Terry, 2017). However, studies reveal high production of the Argentine Spanish voiced and voiceless palatal fricatives [ʒ]/[ʃ] by L2 Spanish (L1 English) speakers in Buenos Aires (Hoffman-González, 2015; Pozzi & Bayley, 2020; Schmidt, 2020). Similarly, L2 Arabic (L1 English) speakers in Cairo produced high rates of the Cairene/Egyptian Arabic variant [g] for /j/, which is realized as [ʒ], [dʒ], or [j] in Modern Standard Arabic (MSA) (Raish, 2015). Regarding individual differences, studies have found that L2 speakers of higher proficiency3 (Geeslin & Gudmestad, 2011; George, 2014), those with positive attitudes toward the local dialect (George, 2014; Schmidt, 2020), and those with stronger social networks with locals (Kennedy Terry, 2017) were more likely to acquire dialect-specific phones.

The differences in acquisition rate between variables is noteworthy. Pozzi and Bayley (2020, p. 18) attribute the high acquisition of L2 Spanish [ʒ]/[ʃ] to the salience of the variants. Given its salience, proficiency and social network were not significant in their study. Whereas French /l/ elision, a less salient feature, demonstrated a significant effect of social network (Kennedy Terry, 2017, p. 566). In addition to salience, it is important to note grapheme-to-phone correlations, in which a one-to-one correspondence results in increased acquisition. For example, L2 Spanish speakers in Buenos Aires have a one-to-one substitution of [ʒ]/[ʃ] for [j] for <y> and <ll> (Pozzi & Bayley, 2020, p. 16) and L2 Arabic speakers in Cairo have a one-to-one substitution of [g] for [ʒ]/[dʒ]/[j] for the Arabic character <ﺝ>. However, acquiring Castilian Spanish /θ/ requires that L2 speakers, who generally produce /s/ for <s, z, ci, ce>, learn a phonemic split to acquire the distinction between /s/ for <s> and /θ/ for <ς, ci, ce> (Knouse, 2012, p. 531). French /l/ elision (Kennedy Terry, 2017) and Spanish coda /s/ elision/aspiration (Geeslin & Gudmestad, 2011; Linford et al., 2021; Schmidt, 2020) also provide challenges for L2 speakers as these allophonic realizations present a grapheme-to-phone mismatch and subsequently are only favored by those with more dense and multiplex social networks with locals (Kennedy Terry, 2017) or with higher proficiency (Geeslin & Gudmestad, 2011).

**Acquisition of heritage language phonemes/allophones in SA settings**

A heritage Spanish speaker in the United States is an individual who grew up speaking or hearing Spanish at home and received their schooling in English (Valdés, 2000). Until recently there has lacked studies examining the sound system of heritage speakers (Rao & Amengual, 2021), and even fewer in the SA context. In general, studies have found that heritage speakers maintain phonological contrasts in their two languages in which they produce sounds closer to monolingual speaker norms than L2 speakers (Amengual, 2018, 2019; Chang et al., 2011), but also exhibit phonological transfer from the community language (Ronquest, 2012; Shea, 2019). However, due to heterogenous experiences with the heritage language, there exists intragroup variation among heritage speakers.

---

3Knouse (2012) found that beginner/intermediate speakers were more likely to produce regional phones while Schmidt (2020) found no proficiency effects. George (2014) found no proficiency effects for /θ/, but that those with higher proficiency favored [χ].
The few previous studies of heritage speakers in the SA context indicate that one’s heritage language variety may play a significant role in the acquisition of dialect-specific phones. For example, Raish (2015) found that two Egyptian–Arabic heritage speakers favored the local Cairene allophone more than one Levantine and one Palestinian heritage speaker who favored supralocal MSA allophones. Similarly, Escalante (2018), examining the acquisition of coda /s/ aspiration in coastal Ecuador (where coda /s/ aspiration is the norm) found that one Ecuadorian Spanish heritage speaker acquired the aspirated variant while one Colombian Spanish heritage speaker and one Mexican Spanish heritage speaker did not. However, George and Hoffman-González (2019) indicate that heritage speakers may adopt features from outside of their heritage variety. Specifically, they found that two Mexican Spanish heritage speakers in Argentina frequently produced [ʃ] and/or [ʒ] and one of two Mexican Spanish heritage speakers in Spain frequently produced [θ]. Thus, to better understand the effect of one’s heritage language variety on the production of dialect-specific phones, it is essential to include heritage speakers in the SA context.

The phonetic variable: Intervocalic /d/

The target feature of the present study is Spanish intervocalic /d/, which provides difficulties to L2 Spanish (L1 English) speakers as it presents an allophone that is not associated with this phoneme in their L1. In monolingual Spanish, the phonemes /b, d, g/ in syllable-initial position vary between voiced stops [b, d, g] and voiced approximants [β, ð, χ] depending upon the phonetic context. Specifically, for /d/, voiced dental stop [d] tends to be realized after a pause (‘día ‘day’ ['di.a]), a nasal consonant (Andalucía [an.ˈda.lu.ˈsi.a]/[an.ˈda.lu.ˈθi.a]), or a lateral (aldea ‘small village’ [a.ˈθe.a]) (Hualde, 2005, p. 139). In all other positions, /d/ undergoes spirantization and tends to be realized as approximant [ð] such as in intervocalic position (Granada [ˈgra.ˈna.ða]) or after nonnasal and nonlateral consonants (Córdoba ['ko.ɾo.βa]) (Hualde, 2005, p. 139). Several varieties also present an elided allophone [∅] in intervocalic position (Hualde, 2005, pp. 21–29).

In monolingual English, syllable-initial /d/ in word-initial position is realized as either voiced [d] or devoiced [d ̥] alveolar stop (Hualde, 2005, p. 149). However, in American English, /d/ (and /t/) in syllable-initial intervocalic position may also be realized as a flap, similar to Spanish [ɾ] (Hualde, 2005, p. 150; Ladefoged & Johnson, 2014, p. 68). Different from Spanish, in English /ð/ exists as a phoneme associated with <th> so that /d/ and /ð/ form minimal pairs (i.e., dough–though) (Ladefoged & Johnson, 2014, p. 70). In addition, [ð] is an approximant in Spanish, whereas it is a fricative in English (Hualde, 2005; Martínez Celdrán, 1991). Thus, although English speakers have [ð] in their phonological inventory, it is not associated with intervocalic /d/ in their L1. Consequently, L2 Spanish speakers must learn to acquire the context-specific variable realizations of [d] and [ð] for /d/.

Intervocalic /d/ among L2 and heritage Spanish speakers

Overall, studies examining L2 Spanish (L1 English) speakers’ acquisition of intervocalic /d/ have found that speakers with higher L2 proficiency produce more [ð] than those

---

4/b, d, g/ variation between stop and approximant allophones is considered complementary distribution (Hualde, 2005, p. 138), but as this varies greatly, and as differences are gradient, here it’s referred to as phonetic trends instead of absolutes.
with lower proficiency who produce more [d] (Alvord & Christiansen, 2012; Face & Menke, 2009; González-Bueno, 1995; Rogers & Alvord, 2014; Zampini, 1994), particularly highly proficient speakers who have lived abroad for extended stays (Alvord & Christiansen, 2012; Rogers & Alvord, 2014). In comparing advanced L2 Spanish speakers to L1 Spanish speakers, Solon et al. (2018) found that advanced L2 speakers produced relatively high levels of spirantization and elision, but less than L1 speakers. Additionally, [Ø] realizations were most influenced by lexical frequency, while [ð] realizations were most influenced by phonetic context and stress, indicating that the predictive factors for L2 speakers were different than those for L1 speakers.

In the SA context, Díaz-Campos (2004) found no increase in [ð] realizations over a 10-week period, neither among SA students in Alicante nor among AH students. However, approximants were favored in informal conversations as compared to a reading task (Díaz-Campos, 2006), but only for the AH students. In an 8-week SA program in Mexico, Lord (2010) found that all students increased [ð] realizations, with greater gains among those with previous phonetic instruction. Finally, Bongiovanni et al. (2015) found more gains in [Ø] realizations over 4 weeks among SA students in the Dominican Republic than AH students.

Regarding intervocalic /d/ realizations by heritage Spanish speakers, Rao (2015) found that heritage speakers who reported more regular use of Spanish produced more approximants than those who reported less language use. Additionally, Amengual (2019) found that both sequential and simultaneous bilingual Spanish heritage speakers produced more lenited [ð] than L2 Spanish speakers.

**Intervocalic /d/ in Andalusian Spanish**

Although intervocalic /d/ elision is not unique to Andalusian Spanish nor a recent development (Moreno-Fernández, 2004, p. 999), it is frequent in Andalusian Spanish (Narbona et al., 1998, p. 93). Sociolinguists have analyzed intervocalic /d/ throughout Andalucía finding overall elision rates between 23–36% in Córdoba, Granada, and Málaga (Urururu, 1996; Moya-Corral & García-Wiedemann, 2009; Villena-Ponsoda, 2012; Villena-Ponsoda & Moya-Corral, 2016) with rates as high as 68% in Jerez (Harjus, 2018). These studies found social patterns in which younger generations, men, those with less educational attainment, and more informal speech styles most favored elision. Several linguistic factors also govern this realization in which atonic syllables, posttonic syllables, and priming effects of prior /d/ elisions favor [Ø]. The strongest linguistic predictor, however, is grammatical category in which the past participle morpheme /ado/ (exagerado ‘exaggerated’) most favors elision, followed by other morphemes such as /ada/ (relejada ‘relaxed’) or /ido/ (salido ‘left’/‘exited’), while root morphemes (cada ‘each’) are more likely to favor retention (Moya-Corral & García-Wiedemann, 2009; Villena-Ponsoda, 2012; Villena-Ponsoda & Moya-Corral, 2016). When only considering the /ado/ context, elision rates rise to 86.3% in Granada and 85.7% in Málaga (Villena-Ponsoda & Moya-Corral, 2016). Thus, although intervocalic /d/ elision is common in Andalusian Spanish, the /ado/ context provides an even higher frequency of elision. When /d/ is elided in the /ado/ context, it is phonetically realized as [ˈa.o] or [ˈau] (Moreno-Fernández, 2004, p. 999; Penny, 2000, p. 133).

---

5A few exceptional root morphemes like todo ‘all’ and nada ‘nothing’ have high rates of elision (Moya-Corral & García-Wiedemann, 2009, p. 111).
Not all varieties of Spanish demonstrate elided [∅] for /d/. For example, the elision of intervocalic /d/ generally does not occur in Mexican Spanish (Moreno-Fernández, 2004, p. 999; Penny, 2000, p. 133). Thus, for L2 and heritage Spanish speakers who are exposed to Texas and/or Mexican Spanish, the elision of /d/ likely presents a new allophone. Another dialectal difference is that the present perfect in Peninsular Spanish has become the “default exponent of past perfective tense/aspect” meaning that the present perfect is being used in contexts previously reserved for the preterit whereas Mexican Spanish continues with the default of preterit (Schwenter & Torres Cacoullos, 2008, p. 33). This indicates that intervocalic /d/, due to high use of the present perfect with past participle /ado/, is more frequent in Peninsular Spanish. Therefore, intervocalic /d/ in Andalusian Spanish provides the opportunity for exposure to language-specific [ð] and dialect-specific [∅].

Although previous literature has examined the acquisition of language-specific /d/ by L2 Spanish speakers in the SA context, these studies have not accounted for the role of proficiency in the SA context. Furthermore, there are no previous studies of heritage Spanish speakers in the SA context focusing on language-specific or dialect-specific allophonic variation of /d/.

**Research questions**

Based on the SA literature, this study sought to answer four research questions.

**RQ1:** What is the effect of short-term SA on the production of language-specific allophone [ð] for intervocalic /d/?

**H1:** Based on Bongiovanni et al. (2015) and Lord (2010), it was hypothesized that there would be an overall increase of approximant-like [ð] at the end of the SA program.

**RQ2:** How does the production of language-specific allophone [ð] vary by proficiency level (L2-intermediate, L2-advanced, Heritage-advanced)?

**H2:** Based on previous L2 intervocalic /d/ studies (Alvord & Christiansen, 2012; Face & Menke, 2009; González-Bueno, 1995; Rogers & Alvord, 2014; Solon et al., 2018; Zampini, 1994), it was hypothesized that L2-advanced speakers would begin with more approximant-like [ð] and also show more gains over time than L2-intermediate speakers. Based on Amengual (2019), it was hypothesized that Heritage-advanced speakers would produce more approximant-like [ð] at Time 1 compared to L2-advanced speakers, but as there are no previous studies of heritage Spanish speakers for intervocalic /d/ in the SA context, there was no formal hypothesis for Heritage-advanced speakers over time.

**RQ3:** What is the effect of proficiency level on the production of dialect-specific allophone [∅] for intervocalic /d/?

**H3:** Based on previous findings of more proficient L2 speakers being more likely to acquire regional phones (Geeslin & Gudmestad, 2011) and of more proficient L2 speakers being more likely to produce [∅] for intervocalic /d/ (Face & Menke, 2009; Rogers & Alvord, 2014; Solon et al., 2018), it was hypothesized that L2-advanced speakers would produce more [∅] than L2-intermediate speakers. Based on the findings that heritage speakers tend to acquire dialect-specific phones from their own heritage language variety (Escalante, 2018; Raish, 2015), it was hypothesized that the Heritage-advanced speakers would not produce the elided variant as none were Andalusian Spanish heritage speakers.
**RQ4:** How does the production of language-specific and dialect-specific allophones vary by other individual factors (attitudes, input)?

**H4:** Based on Elliot (1995), it was hypothesized that students who wanted to sound more native/Andalusian-like, would produce more approximant-like [ð] and/or elision. Based on Kennedy Terry’s (2017) social network finding, it was hypothesized that students who reported more hours speaking Spanish with sevillano/as outside of the classroom would demonstrate more approximant-like [ð] and/or elision over time. Based upon George (2014) and Schmidt (2020), it was hypothesized that those with positive attitudes toward Sevilla Spanish would produce more [∅].

**Methods**

**Participants**

Forty undergraduate students\(^6\) (31 women, 9 men) from a large southwestern public university, ranging in age from 19–22 years (\(M: 20.45; SD: 0.99\)) participated in a 5.5-week SA program in Sevilla, Spain during summer 2018. These students included 30 L2 and 10 heritage Spanish speakers. Heritage Spanish speakers were identified as students who reported speaking Spanish with family members (siblings, parents, grandparents, caretakers). Twenty-one students (20 L2, 1 heritage) enrolled in six-credits of a fourth semester intermediate Spanish language course and 19 students (10 L2, 9 heritage) enrolled in six-credits of advanced senior seminar culture and grammar courses. The intermediate-level course was taught by a native Spanish instructor who speaks northern Mexican Spanish and the advanced courses were taught by a near-native Spanish instructor whose speech closely resembles Andalusian\(^7\) Spanish. Students stayed in pairs with host families. Courses met Monday through Friday for 4 hours per day. Twice a week for 45 minutes, Andalusian university students entered each class and spoke with the students in small groups about cultural topics. One evening a week for 2 hours, small groups of 2–3 students met with Andalusian university students for informal cultural conversations while exploring the city.

For proficiency level (Table 1), students were grouped according to their course level (intermediate, advanced), with a distinction made among the advanced level\(^8\) between

<table>
<thead>
<tr>
<th>Table 1. Proficiency level group demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L2-intermediate</strong></td>
</tr>
<tr>
<td>(n = 20)</td>
</tr>
<tr>
<td><strong>M (SD)</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td><strong>Self-rated Spanish level (1–8)</strong></td>
</tr>
<tr>
<td><strong>Language Questionnaire score</strong></td>
</tr>
<tr>
<td><strong>Hours per week speaking with sevillano/as</strong></td>
</tr>
</tbody>
</table>

\(^6\)An additional student participated but produced /d/ with higher intensity (dB) than the adjacent vowels, indicating a voiced affricate and therefore was excluded from the analysis.

\(^7\)This is a possible limitation as students in the advanced courses may have been exposed to /d/ realizations with less IntDiff than those in the intermediate course.

\(^8\)To clarify, this should not be interpreted that “heritage” is considered a proficiency level. That is, both L2-advanced and Heritage-advanced speakers are of the advanced proficiency level, but with a distinction here based on type of bilingualism following the findings of previous studies.
L2-advanced and Heritage-advanced as previous studies found differences between advanced L2 and heritage speakers for /d/ (Amengual, 2019). Students were asked to self-rate\(^9\) their Spanish level in a preprogram questionnaire (Schmidt, 2020): beginner (1), lower-intermediate (2), intermediate (3), high-intermediate (4), advanced (5), high-advanced (6), near-native (7), and native (8). The one heritage speaker in the intermediate course self-rated her Spanish in line with the other heritage speakers and was therefore placed into the Heritage-advanced group. A one-way ANOVA, \(F(2,37) = 79.93, p < 0.0001\), indicated a main effect of group on self-rated Spanish level. A Tukey Post hoc indicated that the Heritage-advanced group had a higher self-rating than the L2-intermediate \((p < 0.0001)\) and the L2-advanced groups \((p < 0.0001)\), and that the L2-advanced group had a higher self-rating than the L2-intermediate group \((p < 0.01)\).

Additionally, a one-way ANOVA, \(F(2,37) = 4.66, p < 0.05\), demonstrated a main effect of group for the language questionnaire. This questionnaire, discussed in the following section, indicates that those with higher scores wanted to sound more native/Andalusian-like. A Tukey Post hoc indicated that the L2-intermediate group had a higher score than the Heritage-advanced group \((p < 0.05)\). No other comparisons were significant. Finally, there were no significant differences between groups for average hours spoken per week with sevillano/as, \(F(2,37) = 0.74, p = 0.49\).

Although the heritage speakers were placed into one group, it is a fairly heterogenous group. Following Silva-Corvalán’s (1994) sociolinguistic generation (G) categorization, the heritage group was comprised of one G0.5 speaker, one G1.5 speaker, six G2 speakers, and two G3 speakers. The G0.5 speaker arrived in the United States at age 4 and the G1.5 speaker at age 12 (her father is a L1 English speaker). The G2 speakers were born in the United States with at least one parent born in a Spanish-speaking country (meaning countries where Spanish is the majority language) while the G3 speakers and their parents were born in the United States with their grandparents born in a Spanish-speaking country. There were eight heritage speakers of Mexican Spanish, one of Salvadoran Spanish, and one of Ecuadorian Spanish.

**Materials and procedure**

Two days after arriving to Sevilla (the day before classes began), students met individually with the author for the first session of the study. Students completed six tasks (in this order): 5-minute semidirected conversation about studying in Sevilla, passage reading, carrier phrase reading, preprogram questionnaire, and language questionnaire. Only the carrier phrases and the questionnaires are analyzed here.

The carrier phrases consisted of “Yo puedo decir **TARGETWORD quillo**” “I can say **TARGETWORD** dude/bro/bruh.” The rationale for the inclusion of **quillo** after the targetword was twofold. First, following Gerfen’s (2002, p. 249) use of **tío** ‘dude’ in a carrier phrase, it was designed to be informal and put the participants in Andalusian Spanish mode. Prior to the reading, the author explained what **quillo/a** meant and how it is used, providing students with colloquial terminology to use with sevillano/as their age. For reference, **quillo** in Andalusian Spanish is a common way to refer to a male friend as it is the shortened version of **chiquillo**. Secondly, as L1 English speakers produce more creak in phrase-final position, the inclusion of an additional word was to avoid creak during the targetword, which would prevent an accurate acoustic measurement.

---

\(^{9}\)While self-rating may be a limited measure for proficiency, the author decided it was preferred to a written grammar exam as the study examined phonetic (not morphosyntactic) gains.
There were 21 past participle /ado/ words (i.e., *cansado* ‘tired’) included (Appendix A), controlling for syllabic stress (atonic) and phonetic context (preceded by /a/, followed by /o/) as previous studies found adjacent vowel effects for spirantization (Colantoni & Marinescu, 2010; Simonet et al., 2012; Solon et al., 2018). Each word was repeated twice for a total of 42 /ado/ realizations per session. There were also 92 distractors, resulting in 134 sentences per session. All words were presented in the carrier phrase in randomized order. Each participant, sitting comfortably at a desk in a quiet office, read the carrier phrases from a timed Microsoft PowerPoint on a laptop. The carrier phrases were timed at 5 seconds per slide to encourage a similar interspeaker speech rate. This allowed speakers to focus on the screen without having to press continue and thus discouraged them from reading each slide too fast. Speakers were recorded with a Marantz PMD660 solid-state digital recorder wearing a Shure WH20XLR Headworn Dynamic Microphone with a sampling rate of 44.1kHz (16-bit digitization).

The study solely examined the /ado/ context as it is highly frequent in Spanish, particularly in Peninsular Spanish where the present perfect has expanded into the domain of the preterit (Schwenter & Torres Cacoullos, 2008). From a usage-based (Bybee, 2001) and exemplar theory framework (Pierrehumbert, 2001), the /ado/ context is ideal to examine the acquisition of allophonic variation due to its high frequency with many exemplars in the immersion context, providing speakers the opportunity to demonstrate subtle gains during a short-term SA. This does not allow us to overgeneralize the acquisition of intervocalic /d/ in all contexts, but at least in a frequent morphophonological context.

The preprogram questionnaire was based on Schmidt (2020) (also Bedinghaus, 2015; Linford, 2016) and revised for the Sevilla context. The questionnaire consisted of three sections: demographic information, linguistic history, and varieties of Spanish. The language questionnaire was a modified version of Elliot’s (1995) Pronunciation Attitude Inventory (online Appendix B). Modifications were made so that instead of asking about speaking general Spanish, it was specifically framed about one’s attitudes toward sounding Andalusian. Scores ranged from 12 to 60, with higher numbers indicating more of a preference to sound native/Andalusian-like.

Students filled out weekly questionnaires to document how many hours of Spanish they were speaking with sevillano/as outside of class each week. The weekly questionnaires were averaged to provide one score per student.

On the last day of class, 4.5 weeks after the initial recording, students met individually with the author for the second session. They completed the same three verbal-recording tasks from the first session and a postprogram questionnaire. The postprogram questionnaire, based on Schmidt (2020), was modified for the Sevilla context and consisted of SA experience, dialects of Spanish, and motivations for studying Spanish.

**Acoustic measure**

A textgrid was created for each recording by a research assistant, segmenting each word in Tier 2 and each /d/ and following /o/ in Tier 1 in Praat (Boersma & Weenink, 2019). Each segmented /d/ and /o/ were then manually checked by the author. Upon completion of the textgrids, two acoustic measures were taken for each /ado/ token. First, the minimum intensity (dB) of /d/ was obtained by highlighting the /d/ segment and clicking **intensity > get minimum intensity**. Second, the maximum intensity (dB) of the following /o/ was obtained by highlighting the /o/ segment and clicking **intensity >**
get maximum intensity. These measures were used to create the dependent measure: intensity difference (henceforth IntDiff). IntDiff “is the difference between the intensity minimum during the consonant and the intensity maximum in the following tautosyllabic vowel” (Hualde et al., 2011, p. 309). IntDiff provides a relative degree of spirantization or lenition (Colantoni & Marinescu, 2010; Hualde et al., 2011). A larger IntDiff indicates a more stop-like realization (Figure 1) while a smaller IntDiff indicates a more approximant-like realization (Figure 2). Thus, IntDiff provides a gradient measure, different from a segmental analysis ([d], [ð], [∅]). IntDiff has been shown to be a useful parameter for /d/ realizations for L2 and heritage Spanish speakers (Amengual, 2019; Bongiovanni et al., 2015; Rogers & Alvord, 2014; Solon et al., 2018). In the case of elided tokens where there was no consonant to be segmented (Figure 3), following Rogers and Alvord (2014, p. 412), these tokens were classified as having an IntDiff of 0 dB as elision is “one extreme of the gradient scale of spirantization just as an occlusion is the opposite extreme.” Finally, tokens identified as flap [ɾ] (Figure 4) were removed from statistical analyses of the acoustic data as these are not part of the spirantization continuum.

Following the experimental design there were 3,360 possible tokens (21 words × 2 repetitions × 40 speakers × 2 sessions). However, 192 tokens were discarded from the

Figure 1. Waveform, spectrogram with intensity line (dB), and textgrid of stop [d] in comprado ‘purchased’ produced by a L2-intermediate speaker.

Figure 2. Waveform, spectrogram with intensity line (dB), and textgrid of approximant [ð] in reventado ‘burst’/‘exhausted’ produced by a Heritage-advanced speaker.
analysis due to speakers producing a different word from the one presented or the presence of creak that prevented an accurate measure of intensity (dB).

**Independent variables**

There were eight independent variables included in the analysis: time (Time 1 [T1], Time 2 [T2]); language questionnaire score; proficiency level (L2-intermediate, L2-advanced, Heritage-advanced); average hours per week outside of class speaking in Spanish with **sevillano/as**; preferred variety to emulate (Mexican/Texas Spanish, Peninsular Spanish, other); gender (male, female); knowledge of a third-language (yes, no); and whether one likes Sevilla Spanish (yes, neutral/disagree). Preferred variety to emulate was an open-ended question in the preprogram questionnaire that was coded into the three categories mentioned previously. The postprogram questionnaire had several attitude questions including “I like how Spanish is spoken in Sevilla” with the answers agree, neutral, or disagree. This was a coded as a binary agree versus neutral/disagree. Speaker and word were included as random factors.

![Figure 3](image1.png)

**Figure 3.** Waveform, spectrogram with intensity line (dB), and textgrid of elision [∅] in cortado ‘cut’/‘macchiato coffee’ produced by a L2-advanced speaker.

![Figure 4](image2.png)

**Figure 4.** Waveform, spectrogram with intensity line (dB), and textgrid of flap [ɾ] in pensado ‘thought’/‘considered’ produced by a L2-intermediate speaker.
Statistical analysis

For the segmental analysis, chi-square tests and a mixed-effects logistic regression were conducted in R (R Core Team, 2020). Then for the acoustic analysis, following Tagliamonte and Baayen (2012), a random forest was calculated using the cforest function in the party package (Hothorn et al., 2020) in R to determine the importance of each variable. The random forest results are essential prior to regression modeling as it determines the order of the independent variables from most to least important. These variables were placed in this order in regression modeling as the order can affect the output of the models. Mixed-effects linear regression models using treatment contrast were fitted using the lmer function (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2014) in R with the eight independent variables and the random factors of speaker and word. Time was put in interaction with all other factors and several three-way interactions were tested. Nonsignificant main effects and interactions were then discarded from each subsequent model. Models were compared using ANOVA testing and the model with the lowest AIC was chosen for the final analysis. Post-hoc analyses were computed using estimated marginal means (Lenth et al., 2018) for categorical predictors with more than two levels. Additionally, paired Welch two sample t-tests were used to determine statistical differences between Time 1 and Time 2 for each individual speaker. Figures were created with ggplot2 (Wickham, 2013).

Results

Segmental analysis

The 3,168 /d/ tokens were segmentally coded by examining the spectrogram and the waveform of each textgrided segment, resulting in 925 [d], 1,535 [ð], 18 [∅], and 690 [ɾ] realizations. These allophonic realizations varied per proficiency level and time as seen in Table 2 and Figure 5.

A chi-square test found that the distribution of segmental allophonic variants for /d/ was significantly different between proficiency groups, $\chi^2(6) = 842.24, p < 0.0001$. Neither the L2-advanced group ($\chi^2(3) = 6.35, p = 0.10$) nor the Heritage-advanced group ($\chi^2(2) = 3.86, p = 0.15$) demonstrated significant changes in allophonic variants over time. The L2-intermediate group, however, demonstrated significant changes in variants over time ($\chi^2(2) = 15.57, p < 0.001$); specifically, a reduction in [d] ($\chi^2(1) = 5.0, p < 0.05$), an increase in [ð] ($\chi^2(1) = 10.47, p < 0.01$), and no change in [ɾ] realizations ($\chi^2(1) = 0.50, p = 0.48$). To solely compare [d] and [ð] realizations, a mixed-effects logistic regression using treatment contrast was conducted with time and proficiency level as fixed factors and speaker and word as random factors. The model ($R^2_m: 0.35; R^2_c: 0.75$) found a main effect for proficiency in which L2-advanced ($\beta = 2.46, SE = 0.95, z = 2.58, p < 0.01$) and Heritage-advanced ($\beta = 5.21, SE = 0.98, z = 5.33, p < 0.001$)

<table>
<thead>
<tr>
<th>Group</th>
<th>Time</th>
<th>[ɾ]</th>
<th>[d]</th>
<th>[ð]</th>
<th>[∅]</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2-intermediate</td>
<td>Time 1</td>
<td>250</td>
<td>354</td>
<td>175</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>266</td>
<td>297</td>
<td>241</td>
<td>0</td>
</tr>
<tr>
<td>L2-advanced</td>
<td>Time 1</td>
<td>79</td>
<td>99</td>
<td>226</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>95</td>
<td>80</td>
<td>219</td>
<td>7</td>
</tr>
<tr>
<td>Heritage-advanced</td>
<td>Time 1</td>
<td>0</td>
<td>52</td>
<td>328</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Time 2</td>
<td>0</td>
<td>43</td>
<td>346</td>
<td>7</td>
</tr>
</tbody>
</table>
produced more [ð] than L2-intermediate speakers. A main effect for time ($\beta = 0.71$, $SE = 0.17$, $z = 4.2$, $p < 0.001$) indicated that for all speakers together Time 2 had more [ð] than Time 1. Finally, there was a significant interaction between time and proficiency level. Post-hoc pairwise comparisons indicated that only the L2-intermediate group demonstrated a significant increase in [ð] realizations ($p < 0.001$). For Time 1, Heritage-advanced ($p < 0.001$) and L2-intermediate ($p < 0.05$) had significantly more [ð] realizations than L2-intermediate. Heritage-advanced ($p < 0.001$) also had significantly more [ð] realizations than L2-intermediate. For Time 2, Heritage-advanced had significantly more [ð] realizations than L2-intermediate ($p < 0.001$) and L2-advanced ($p < 0.05$). No other comparisons were significant.

**Acoustic analysis**

There was a relatively high occurrence of flaps [ɾ] for intervocalic /d/, particularly among the L2-intermediate group. Following previous acoustic studies (Bongiovanni et al. 2015; Solon et al., 2018), all [ɾ] realizations were removed from further analysis. Additionally, two L2-intermediate speakers (B10: 92.2% [ɾ]; B20: 90% [ɾ]) and one L2-advanced speaker (A10: 92% [ɾ]) who demonstrated nearly categorical flaps were removed from the analysis. Finally, there were 58 [ð] tokens with negative IntDiff, meaning that although the approximant was present, the following /o/ was significantly reduced (i.e., atonic schwa-like vowel). As this value gives the false impression of an elision, these tokens were discarded from further analyses. After removing these tokens, the remaining data consisted of 2,400 tokens.

The random forest (Figure 6) indicated that the most important predictor of IntDiff is speaker, followed by language questionnaire, proficiency level, whether one likes Sevilla Spanish, word, time, average hours per week speaking Spanish with sevillano/as, gender, preferred variety to emulate, and L3. Given language questionnaire and Like-SevillaSpanish demonstrated high degrees of collinearity, they were not placed into the same regression model, but rather substituted for one another in different models. As the models with language questionnaire demonstrated a better fit for the data with a lower AIC score as well as higher variable importance in the random forest, the final regression model opted for language questionnaire.
The best-fit mixed-effects linear regression is seen in Table 3. The model failed to converge with word as a random factor,\(^\text{10}\) and thus the final model only included speaker as a random factor. The table presents the estimate, standard error (SE), \(t\)-value, and \(p\)-value. Larger estimates in either direction from zero indicate a stronger main effect or interaction. Positive estimates indicate that the listed level has a higher \textit{IntDiff} than the reference level and vice versa for lower estimates. Marginal R-squared (\(R^2_m\)) and conditional R-squared (\(R^2_c\)) values are provided to show the goodness-of-fit of the variation (Nakagawa & Schielzeth,\(^\text{2013}\)).

The main effect for time indicated that Time 1 (\textit{EMM}: 10.84dB, SE: 0.63) had a significantly higher \textit{IntDiff} than Time 2 (\textit{EMM}: 9.44dB, SE: 0.63) \((p < 0.001)\). The interaction between language questionnaire and time indicates that those with higher questionnaire scores lowered their \textit{IntDiff} from Time 1 to Time 2 (Figure 7). As one observes, however, those with less desire to sound native/Andalusian (lower scores) started with a low \textit{IntDiff} and did not demonstrate change over time while those with more desire to sound native/Andalusian (higher scores) generally started in Time 1 with a high \textit{IntDiff}, but lowered \textit{IntDiff} by Time 2. The time by proficiency level interaction and post-hoc pairwise comparisons\(^\text{11}\) revealed that L2-advanced had a higher \textit{IntDiff} in Time 1 (\textit{EMM}: 10.68 dB, SE: 1.24) than in Time 2 (\textit{EMM}: 9.52 dB, SE: 1.24) \((p < 0.01)\), Heritage-advanced had a higher \textit{IntDiff} in Time 1 (\textit{EMM}: 9.54 dB, SE: 1.23) than in Time 2 (\textit{EMM}: 6.91 dB, SE: 1.23) \((p < 0.001)\), but that L2-intermediate did not demonstrate a significant difference between Time 1 (\textit{EMM}: 12.3 dB, SE: 0.94) and Time 2 (\textit{EMM}: 11.88 dB, SE: 0.94) \((p = 0.21)\) (Figure 7). The interaction indicates that Heritage-advanced had the largest reduction in \textit{IntDiff} over time.

\(^{10}\)A follow-up mixed-effects linear regression examined word-specific phonetics (Pierrehumbert,\(^\text{2001}\)) with a time by word interaction with speaker as a random factor. The only significant interaction was for \textit{preguntado} in which \textit{IntDiff} increased over time.

\(^{11}\)True means and standard deviations (\(M[SD]\)): L2-intermediate (T1: 12.9 dB [6.4]; T2: 12.0 dB [6.4]), L2-advanced (T1: 10.9 dB [6.04]; T2: 9.69 dB [5.86]), Heritage-advanced (T1: 7.81 dB [6.79]; T2: 5.89 dB [5.59]).
The three-way interaction between time, proficiency level, and average hours speaking with *sevillano/as* indicates that only the Heritage-advanced group was significantly different from the L2-intermediate group reference level (Figure 8). That is, both the L2-intermediate and L2-advanced groups followed the same pattern in which those with more hours reduced IntDiff from Time 1 to Time 2. On the contrary, for the Heritage-advanced group, those with fewer hours with *sevillano/as* demonstrated IntDiff reduction over time while those with more hours demonstrated little to no IntDiff reduction.

To further explore the relationship between the independent variables in the regression model, a conditional inference tree using the *cforest* function in the *party* package.

**Table 3.** Summary of mixed-effects linear regression model for IntDiff, speaker as random factor, n = 2,400 (R²m: 0.23, R²c: 0.49), note: “Ref” = reference level

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>6.07</td>
<td>2.09</td>
<td>2.91</td>
<td>**</td>
</tr>
<tr>
<td>LanguageQ</td>
<td>0.48</td>
<td>0.14</td>
<td>3.44</td>
<td>**</td>
</tr>
<tr>
<td>Time (Ref = Time1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Time2</td>
<td>2.31</td>
<td>0.71</td>
<td>3.25</td>
<td>**</td>
</tr>
<tr>
<td>Level (Ref = L2-Inter)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>L2-Adv</td>
<td>−0.96</td>
<td>2.39</td>
<td>−0.40</td>
<td>0.69</td>
</tr>
<tr>
<td>Heritage-Adv</td>
<td>1.67</td>
<td>2.54</td>
<td>0.66</td>
<td>0.52</td>
</tr>
<tr>
<td>AverWeekSpanish</td>
<td>0.21</td>
<td>0.14</td>
<td>1.49</td>
<td>0.15</td>
</tr>
<tr>
<td>LanguageQ:Time</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>LanguageQ:Time2</td>
<td>−0.22</td>
<td>0.05</td>
<td>−4.48</td>
<td>***</td>
</tr>
<tr>
<td>Time:Level</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Time2:L2-Adv</td>
<td>−0.96</td>
<td>0.79</td>
<td>−1.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Time2:Heritage-Adv</td>
<td>−4.29</td>
<td>0.86</td>
<td>−5.01</td>
<td>***</td>
</tr>
<tr>
<td>Level:AverWeekSpanish</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>L2-Adv:AverWeekSpanish</td>
<td>−0.08</td>
<td>0.20</td>
<td>−0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>Heritage-Adv:AverWeekSpanish</td>
<td>−0.52</td>
<td>0.23</td>
<td>−2.26</td>
<td>*</td>
</tr>
<tr>
<td>Time:AverWeekSpanish</td>
<td>−0.09</td>
<td>0.05</td>
<td>−1.85</td>
<td>0.06</td>
</tr>
<tr>
<td>Time:Level:AverWeekSpanish</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Time2:L2-Adv:AverWeekSpanish</td>
<td>0.03</td>
<td>0.07</td>
<td>0.35</td>
<td>0.73</td>
</tr>
<tr>
<td>Time2:Heritage:AverWeekSpanish</td>
<td>0.24</td>
<td>0.08</td>
<td>3.14</td>
<td>**</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; *** p < 0.001.

**Figure 7.** Left: Language questionnaire by time interaction for IntDiff; Right: Time by proficiency level interaction for IntDiff.

The three-way interaction between time, proficiency level, and average hours speaking with *sevillano/as* indicates that only the Heritage-advanced group was significantly different from the L2-intermediate group reference level (Figure 8). That is, both the L2-intermediate and L2-advanced groups followed the same pattern in which those with more hours reduced IntDiff from Time 1 to Time 2. On the contrary, for the Heritage-advanced group, those with fewer hours with *sevillano/as* demonstrated IntDiff reduction over time while those with more hours demonstrated little to no IntDiff reduction.

To further explore the relationship between the independent variables in the regression model, a conditional inference tree using the *cforest* function in the *party* package.
(Hothorn et al., 2020) was conducted. In plotting a conditional inference tree with language questionnaire, it split into many trees each with few nodes. Thus, as there was a high degree of collinearity between language questionnaire and LikeSevillaSpanish, the conditional inference tree opted for LikeSevillaSpanish. The conditional inference tree in Figure 9 indicates that the most important predictor is proficiency level, separating L2 and heritage speakers. For both L2-advanced and Heritage-advanced speakers, those who liked Sevilla Spanish reduced their IntDiff over time, whereas those who did not like Sevilla Spanish did not reduce their IntDiff. For L2-intermediate speakers, attitudes toward Sevilla Spanish had no effect on IntDiff over time.

Given the linguistic diversity of the Heritage-advanced group, a few follow-up analyses were conducted to examine intragroup variation. Of interest was that G2

Figure 8. Time by proficiency level by average hours speaking with sevillano/as interaction for IntDiff.

Figure 9. Conditional inference tree of proficiency level, time, and LikeSevillaSpanish for IntDiff.

As Gries (2020, p. 620) indicates, the “binary-split nature” of conditional inference trees “can make the interpretation of the effects of numeric predictors quite cumbersome” and thus some continuous predictors are better fit for a regression.
and G3 had overall lower IntDiff values than G0.5 and G1.5. However, as there was only one G0.5 and one G1.5 speaker, this should be taken with caution. Pearson correlations were conducted to assess the relationship between IntDiff change ($\mu_{\text{Time1}} - \mu_{\text{Time2}}$) and language questionnaire ($n = 10$, $df = 8$, $r = -0.49$, $R^2 = 0.24$, $p = 0.147$), and between IntDiff change ($\mu_{\text{Time1}} - \mu_{\text{Time2}}$) and hours per week spoken with sevillano/as ($n = 10$, $df = 8$, $r = -0.12$, $R^2 = 0.01$, $p = 0.743$), but none were significant. A two-way (time, LikeSevillaSpanish) repeated measures ANOVA found a significant main effect for time ($F(1,720) = 36.6$, $p < 0.0001$) and a significant time by LikeSevillaSpanish interaction ($F(1,720) = 6.85$, $p < 0.01$). This indicates that heritage speakers who liked Sevilla Spanish ($T1 M: 9.4$ dB, $SD: 6.1$; $T2 M: 6.8$ dB, $SD: 5.5$) had a significantly greater IntDiff reduction over time compared to those who felt neutral/disliked Sevilla Spanish ($T1 M: 6.2$ dB, $SD: 7.1$; $T2 M: 5.0$ dB, $SD: 5.6$) (Figure 10). Furthermore, heritage speakers who liked Sevilla Spanish spent less hours per week speaking with sevillano/as ($M: 4.8$ hrs, $SD: 1.8$) than those who felt neutral/disliked Sevilla Spanish ($M: 13.8$ hrs, $SD: 7.0$), $t(4.53) = -2.78$, $p < 0.05$ (Figure 10).

**Individual speaker analysis**

Paired Welch two-sample t-tests were conducted to analyze each individual speaker’s IntDiff over time. Twelve of the 40 participants demonstrated statistically significant reduction in IntDiff over time. Specifically, 3 of 20 L2-intermediate speakers (online Appendix C: B5, B6, B12), 4 of 10 L2-advanced speakers (online Appendix D: A5, A9, A14, A16), and 5 of 10 Heritage-advanced speakers (online Appendix E: A1, A6, A12, A21, B7) demonstrated statistically significant reduction in IntDiff.

Regarding the production of dialect-specific [∅], overall there were 18 elisions: 0 L2-intermediate, 9 L2-advanced, 9 Heritage-advanced. Two individuals accounted for most of these tokens: Luis13 (A20) a Heritage-advanced speaker who produced five elisions (T1: 2, T2: 3), and Nancy (A14), a L2-advanced speaker who produced seven.

![Figure 10. Left: LikeSevillaSpanish by time interaction for IntDiff for Heritage-advanced speakers; Right: LikeSevillaSpanish by average hours per week speaking with sevillano/as for Heritage-advanced speakers.](https://www.cambridge.org/core/terms).

---

13All names are pseudonyms.
elisions (T1: 0, T2: 7). Luis, who grew up speaking Spanish at home in El Paso, had a low IntDiff at Time 1 and did not change over time. Of note, for preferred variety to emulate, Luis wrote “I don’t prefer to sound like a specific dialect.” During the semidirected conversation in Time 1, Luis said he was studying abroad in Sevilla because he likes “to explore different cultures and travel” (my translation) and to improve his Spanish. Of note, however, is that Luis is a highly proficient native speaker of Spanish (and English). Luis mentioned already feeling comfortable in Sevilla two days after arrival due to cultural similarities stating, “La cultura mexicana y la cultural española son casi igual. La gente son muy amable y muy este- respetuosa” “The Mexican culture and the Spanish culture are almost the same. The people are very friendly and very eh- respectful.” Similarly, in the postprogram open-comment Luis wrote, “It was a great experience and I love the culture here a lot. I hope to come and visit and maybe come and live in Spain one day.” Although he did not lower his IntDiff values over time as he already produced [ð] upon arrival, from a social psychology perspective (Giles et al., 1973), the few elided tokens could be linguistic accommodation as he was interested in learning new vocabulary and felt welcomed by sevillano/as.

Nancy, a L2-advanced speaker from the Dallas/Fort Worth area, significantly reduced her IntDiff as her Time 1 IntDiff was quite high. At the end of the program, she exhibited metalinguistic awareness of the elided variant as pertaining to the speech of Sevilla. She produced four [∅] tokens in spontaneous speech in addition to her seven tokens from the carrier phrases. Nancy’s acquisition, however, cannot be described as [d] to [ð] to [∅] as she began the program alternating between [d] and [ɾ] and finished the program alternating between [ɾ] and [∅]. During the semidirected conversation in Time 2, the author (A) was conversing with Nancy (N) about her SA experience. She explained how she and her roommate Laura, a Heritage-advanced speaker in the program, had the best host family of all the students: two young parents and two daughters. In this exchange (Excerpt 1), Nancy stated that her 37-year-old host mother’s “cumple” was last week (line 1). As cumple is a colloquial shortening for cumpleaños “birthday,” the author was impressed and asked her what else she had learned in Sevilla (lines 2 and 4).

(1)
1 N: Treinta y siete. Su cumple fue la semana pasada.
3 N: (Risa)
4 A: Has aprendido, has aprendido algo, ¡eh! (entonación de broma)
5 N: Sí, Laura me dijo que yo he aprendido mucho de español aquí. Y ella no porque ya habla español
de toda su vida. Pero yo, he aprendido mucho de aquí. No sé.
6 A: ¿Y qué has aprendido tú de aquí?
7 N: Mucho. Cumple, anoche cuando estábamos estudiando, yo dije [pa.'sau.], no [pa.'sa.ro], [pa.'say] (risa).
8 A: ¿Sí?
9 N: (risa), por accidente. No estaba pensando en eso, pero yo dije [pen.'say], [pa.'say].
10 A: ¿Sí?
11 N: Sí (risa), por accidente. No estaba pensando en eso, pero yo dije [pen.'say], [pa.'say].
12 A: Muy bien. Entonces, ¿Dónde aprendiste todo eso?
13 N: En mi familia, en mi casa.
14 A: Estás presentando mucha atención, ¿no?
15 N: Sí, sí, porque hablamos mucho con la familia. Cenamos juntos, todos juntos y almorzamos todos juntos.
16 Entonces, hablamos mucho con la familia.
17 A: Así se aprende.
18 N: Sí.
19 A: En la clase podemos aprender muchas cosas, pero=
20 N: = en la casa, y también con las niñas, es más fácil para practicar porque no tengo vergüenza con las niñas.
21 A: ¡Verdad!
22 N: Sí, es menos vergüenza con las niñas. Y la niña de seis años es muy mandona y (risa)
23 ella siempre quiere jugar con yo y Laura y siempre está hablando con nosotros,
y “chicas, sientate aquí y escúchate” (imita la voz de la niña) (risa).
In line 5 she mentions that her roommate Laura had commented that Nancy has picked up a great deal of local forms of speech. In line 8 Nancy explains that last night when they were studying she said *pasado* as \([pa.ˈsaɾ]\) (Figure 11), then contrasts that with *pasado* \([pa.ˈsaɾo]\), followed by another elided example. Of note here is that she produces the diphthong \([aʊ]\) common in Andalusian Spanish when the /d/ in /ado/ is elided (Moreno-Fernández, 2004, p. 999). Nancy explains that she was not intentionally thinking about it, but rather it simply came out. She then produces two more elided examples (line 10). When asked where she learned this, Nancy states that she learned that pronunciation from her host family in her house (line 12), expressing that she and Laura spend a great deal of time, particularly during meals, speaking with their host family (lines 14 and 15). In addition to the direct input from her host parents, Nancy indicates that she was also able to practice her Spanish with the two daughters without worrying about mistakes (lines 19–23). The connection to her host family was quite strong as she later stated, “And I want to return to visit my family” (my translation) and in the open-comment section she wrote, “My experience would not have been near as amazing without my host family. They helped me learn and practice, and treated as one of their own! I love the culture of Sevilla.” Notably, for preferred variety to emulate in the preprogram questionnaire she wrote “informal Spanish,” while in in the postprogram questionnaire she wrote, “During my time I started to emulate a Sevilla accent, not because I wanted to, but just because of so much exposure.” Thus, this could also be a case of linguistic accommodation (Giles et al., 1973), due to favorable opinions of her host family. Consequently, due to a combination of input and positive attitudes toward the local variety, Nancy used local colloquial lexicon and pronunciation, as well as demonstrated metalinguistic awareness of dialect-specific [Ø] in /ado/.

**Discussion**

**Revisiting the research questions**

**RQ1:** What is the effect of short-term SA on the production of language-specific allophone [ð] for intervocalic /d/? The findings support Bongiovanni et al. (2015) and Lord (2010) in which as an entire group, students produced more approximant-like [ð] over time in
the segmental and acoustic analyses. Additionally, 12 individual speakers demonstrated significant reductions in IntDiff over time.

RQ2: How does the production of language-specific allophone [ð] vary by proficiency level? The findings that only L2-advanced and Heritage-advanced groups reduced their IntDiff over time, supports previous studies that spirantization is difficult for beginner/intermediate L2 Spanish speakers and appears to be acquired in more advanced proficiency levels (Alvord & Christiansen, 2012; Face & Menke, 2009; González-Bueno, 1995; Rogers & Alvord, 2014; Solon et al. 2018; Zampini, 1994). Furthermore, Heritage-advanced speakers had the overall lowest IntDiff and demonstrated more IntDiff reduction than the L2-advanced speakers, supporting Amengual (2019). This does not imply that the L2-advanced students have not acquired [ð], but rather that Heritage-advanced speakers demonstrate a more lenited approximant, similar to Amengual (2019). These findings do not suggest that L2-intermediate speakers do not exhibit gains over time as there were three individuals who demonstrated a reduction in IntDiff. Moreover, for the segmental analysis, only the L2-advanced group increased [ø] realizations over time. The differences between the segmental and acoustic analyses indicate that L2-advanced and Heritage-advanced speakers may have already acquired the categorical difference between [ð] and [d] and thus demonstrated more gradient changes over time. However, the L2-intermediate group only demonstrated categorical gains in more [ð] production over time, but the acoustic analysis suggests that their [ð] is less approximant-like than the advanced groups.

RQ3: What is the effect of proficiency level on the production of dialect-specific allophone [∅] for intervocalic /d/? All 18 [∅] tokens were produced by advanced (L2 and heritage) speakers, supporting previous findings that more proficient speakers are more likely than beginner/intermediate speakers to produce dialect-specific phones (Geeslin & Gudmestad, 2011). Although the [∅] realizations were few in number, the results demonstrate that even in a short-term SA program speakers are capable of producing dialect-specific phones, supporting previous studies (Knouse, 2012; Schmidt, 2020). Although one Heritage-advanced speaker, Luis, produced five [∅] tokens, it cannot be definitively attributed to his immersion in Sevilla as he produced two tokens in Time 1. These five tokens notwithstanding, overall, the heritage group demonstrated a lack of production of a dialect-phone from outside of their heritage language variety, supporting previous studies (Escalante, 2018; Raish, 2015).
One L2-advanced student, Nancy, produced dialect-specific [∅] and displayed metalinguistic awareness of the feature supporting previous studies that /d/ elision is more common among L2-advanced than L2-intermediate speakers (Face & Menke, 2009; Rogers & Alvord, 2014; Solon et al. 2018). However, given her few [∅] tokens, this appears to be more of an acquisition of highly frequent words that have been lexicalized with the elision (i.e., pasado, pagado, cortado), than a full acquisition of /d/ elision in all /ado/ contexts. Nancy attributed her use of this feature to her host family input. It could be argued that if she heard these words with the [au] diphthong from her host family, they were stored as exemplars (Pierrehumbert, 2001). This would support Solon et al.’s (2018) findings on L2 speakers acquiring intervocalic /d/ elision in highly frequent words. Additionally, as Nancy demonstrated a change from [d]/[ɾ] to [∅] as opposed to [d]/[ɾ] to [ð], this also supports Solon et al.’s (2018) claim that the acquisition of spirantization and elision may involve different processes for L2 speakers.

RQ4: How does the production of language-specific and dialect-specific allophones vary by other individual factors (attitudes, input)? The findings suggest that both attitudes and input impact speakers’ IntDiff over time. In terms of input, the average hours per week spoken with sevillano/as varied by proficiency level in which L2-intermediate and L2-advanced speakers with more hours demonstrated a reduced IntDiff over time, loosely supporting Kennedy Terry’s (2017) findings that those with more input from locals were better able to acquire local features. The Heritage-advanced speakers, however, demonstrated the opposite pattern, in which those with fewer hours speaking with sevillano/as reduced their IntDiff whereas those with more hours did not. Notably, Heritage-advanced speakers with more hours started with low IntDiff, indicating that these speakers had already acquired language-specific [ð] in the /ado/ context.

Regarding attitudes, speakers with more desire to sound native/Andalusian demonstrated more IntDiff reduction than those with less desire to sound native/Andalusian, supporting Elliot (1995). However, those with less desire to sound native/Andalusian already had quite low IntDiff at Time 1 while those with more desire to sound native/Andalusian began with high IntDiff. Thus, speakers with lower language questionnaire scores may have already acquired the language-specific allophone and thus did not feel a desire to sound native/Andalusian. Additionally, liking Sevilla Spanish only demonstrated reduction in IntDiff for L2-advanced and Heritage-advanced groups, whereas there were no differences in IntDiff over time for L2-intermediate speakers based on attitudes toward Sevilla Spanish. As noted previously, Heritage speakers who felt neutral/disliked Sevilla Spanish spent more hours speaking with sevillano/as than those who reported liking Sevilla Spanish. The interpretation that we posit here, is that given heritage speakers who felt neutral/disliked Sevilla Spanish had very low IntDiff values in Time 1, they may have had a higher proficiency and thus already produced a more approximant-like [ð] as compared to their heritage peers who spent fewer hours speaking with sevillano/as. Thus, attitudes toward a particular variety (Dörnyei et al., 2006; Elliot, 1995; Gardner & Lambert, 1972; Schmidt, 2020) influence acquisition.

**Proficiency**

The role of proficiency in the acquisition of language-specific or dialect-specific phones has been understudied. In large part this is due to practical difficulties of the
SA context. Principally, the number of speakers is generally limited in any given proficiency group, which reduces the statistical power of group comparisons.

The results on whether higher proficiency levels provide the opportunity for more gains during SA demonstrate mixed reviews (Issa & Zalbidea, 2018). Some studies analyzing oral proficiency (Davidson, 2010; Leonard & Shea, 2017), pragmatics (Li, 2014), and grammatical gender (Faretta-Stutenberg & Morgan-Short, 2018) have found that more proficient L2 learners demonstrated more gains than beginner or intermediate L2 learners over time. The current findings suggest that L2-advanced and Heritage-advanced speakers are better positioned to have more language-specific allophonic gains during a short-term SA program than L2-intermediate speakers. Thus, the current study in tandem with previous studies lend support to Lafford and Collentine’s (2006) claim that L2 speakers with higher linguistic proficiency, may have more cognitive resources to notice linguistic variation as compared to lower proficiency L2 speakers whose cognitive energies are focused on communication (see also Shea, 2021). However, the current study demonstrates that proficiency level interacts with individual input and attitudes.

Dialect-specific phones

The low production of intervocalic /d/ elision furthers the notion that certain dialect-specific phones are more acquirable than others. Intervocalic /d/ elision is similar to French /l/ elision in third person singular pronouns (Howard et al., 2006; Kennedy Terry, 2017) and Spanish coda /s/ aspiration/elision (Linford et al., 2021; Schmidt, 2020) in which there is a grapheme-to-phone mismatch, thus less acquirable than a one-to-one grapheme-to-phone correspondence such as [ʒ]/ʃ] for <y> and <ll> in Argentine Spanish (Hoffman-González, 2015; Pozzi & Bayley, 2020; Schmidt, 2020) or Cairene Arabic [g] for the Arabic character ݕ (Raish, 2015). Moreover, the elision of intervocalic /d/ may be less salient than other dialectal features. However, higher salience does not guarantee a higher acquisition. For example, /θ/ is quite salient for L2 Spanish speakers (Ringer-Hilfinger, 2012), but demonstrates little overall acquisition as it requires speakers to learn a phonemic split (George, 2014; Knouse, 2012; Ringer-Hilfinger, 2012).

Only one Heritage-advanced student produced more than one elided variant, although only five tokens. It could be that Luis, not unlike a few heritage speakers in George and Hoffman-González (2019), was producing a dialect-specific phone from outside of his heritage language variety, but with so few tokens, it remains unknown. As one’s identity is tied to one’s heritage language variety, it is perhaps not surprising that these speakers did not adopt the elided variant from outside of their heritage variety, similar to previous studies (Escalante, 2018; Raish, 2015). None of the heritage speakers stated Peninsular Spanish as their preferred variety to emulate, but rather seven reported Mexican/Texas Spanish, two speakers reported no preference, and one Central American/U.S. Spanish (online Appendix E). Future studies should examine heritage language proficiency as higher proficient heritage speakers may demonstrate stronger affiliation to their heritage language variety and thus be less likely to produce dialect-specific features from outside of their variety.

The notion of identity and language variety could also be explored with L2 speakers. Within the U.S. context, L2 speakers who live in bilingual regions such as Texas and are exposed to Texas and Mexican Spanish, may have a stronger connection to Mexican Spanish than L2 speakers from other regions such as the Midwest who may be more
open to emulating other varieties of Spanish. For example, of the 10 L2-advanced speakers, 5 reported Mexican/Texas Spanish as their preferred variety to emulate, 1 Latin American Spanish, 3 none, and 1 Peninsular Spanish (online Appendix D). However, one’s L2 dialectal preference may vary by proficiency as among the 20 L2-intermediate speakers 13 stated no preferred variety to emulate, 5 Mexican/Texas Spanish, and 2 Peninsular Spanish (online Appendix C). This may vary by phonetic features as well. For example, the dialect-specific [ø] for intervocalic /d/ can allow a L2 Spanish speaker to sound sevillano/a without sounding completely Peninsular by producing a variant such as [θ], which is likely much more linked to Peninsular Spanish. However, as Geeslin and Schmidt (2018) indicate, we know quite little about L2 and heritage Spanish speakers’ attitudes toward dialect-specific phones, thus future social perception studies can provide insights into why some variables are more acquired than others.

Consequently, it is suggested that the acquisition of dialect-specific phones is dependent upon a combination of interacting factors: relative sociolinguistic salience (Pozzi & Bayley, 2020); grapheme-to-phone correspondences (Knouse, 2012); social evaluations (Geeslin & Schmidt, 2018); and in the case of heritage speakers, one’s heritage language variety (Escalante, 2018; Raish, 2015). These factors, however, also interact with proficiency level.

**Methodological considerations**

It is possible that contrasting findings between studies are due to different lengths of stay and/or methodological differences. Thus, in short-term SA programs, the role of frequency, must be considered. The current study intentionally focused only on the /ado/ context as it is a highly frequent context in Peninsular Spanish as the present perfect has expanded into the domain of the preterit (Schwenter & Torres Cacoullos, 2008). Given short-term SA programs range from 4 to 6 weeks, researchers should think carefully not only about which feature to study but also the specific phonetic context to provide students the best opportunity to demonstrate subtle gains. Experiments that only examine large gains in language-specific or dialect-specific phones lend to the unhelpful dichotomy of native versus nonnative speaker (Birdsong & Gertken, 2013), going against the “bilingual turn” in SLA (Ortega, 2009). Thus, focusing on features in the SA context that students will receive frequent input would follow Ortega’s (2014) proposal for a usage-based approach to SLA.

Additionally, some studies have utilized a segmental analysis (Díaz-Campos, 2004; Lord, 2010) while others an acoustic analysis (Bongiovanni et al., 2015) for intervocalic /d/ in the SA context. As demonstrated in the current study, an acoustic analysis versus a segmental analysis may yield different results and insights into the data. Thus, in light of these findings, and similar to those of Solon et al. (2018), a combination of segmental and acoustic analyses is warranted to examine both categorical and gradient phonetic changes over time.

**Limitations and future studies**

Proficiency, as measured by a self-reported 8-point Likert-scale, is a limitation in the study. Future studies may consider having students take an ACTFL OPI. Additionally, given the diverse range of linguistic input Spanish heritage speakers receive, it would be advisable to have students complete the Bilingual Language Profile
(Birdsong et al., 2012) to analyze heritage, as well as L2, speakers on a continuous bilingual measure. Also, although IntDiff is useful for demonstrating degrees of spirantization, it does not account for the possibility that some L1 English speakers may produce fricative [ð] instead of approximant [ð] (Ladefoged & Johnson, 2014; Martínez-Celdrán, 1991), which could potentially result in a higher IntDiff. Future studies should examine the overall acquisition of intervocalic /d/ in all intervocalic contexts. If speakers were to demonstrate more gains with /ado/ than other contexts, this would provide support for usage-based (Bybee, 2001) and exemplary theory (Pierrehumbert, 2001) evidence for L2 acquisition. Future studies should also examine intervocalic /d/ over a longer semester to see if more speakers demonstrate more IntDiff reduction as well as produce more dialect-specific [∅]. Following Díaz-Campos (2006), studies should compare spontaneous and read speech to explore stylistic differences. Finally, more work is needed to further examine the role of one’s heritage language variety on the acquisition of dialect-specific phones.

Conclusion

This study examined the role of language proficiency and other individual factors (attitudes, input) in the acquisition of language-specific [ð] and dialect-specific [∅] allophones of Spanish intervocalic /d/ in the /ado/ context by L2 and heritage Spanish speakers during a short-term study abroad in Sevilla, Spain. Overall, there was little production of the dialect-specific allophone, with a few exceptions such as an L2-advanced speaker producing [∅] with a high degree of metalinguistic awareness. The low acquisition of the elided variant is likely due to its low salience and its grapheme-to-phone mismatch, and in the case of heritage speakers, that it is not present in their heritage language variety. Regarding language-specific [ð], the findings imply that advanced (L2 and heritage) speakers with favorable attitudes toward the local variety are most likely to demonstrate gradient allophonic changes during a short-term SA program.

Acknowledgement. I am truly indebted to research assistant Lily Lemus in the Texas Tech Sociolinguistics & Bilingualism Research Lab for outstanding textgrid phonetic segmentation, to Mark Amengual for insights into the experimental design and acoustic analysis, to Lauren Schmidt for sharing materials for preprogram and postprogram questionnaires, to Sara Pink, Myriam Rubio Hernández, Lorum Stratton, and Isabel Ojeda at the Texas Tech University Center in Sevilla for their support, to audience members at LSRL 50 for insightful feedback, and to two anonymous reviewers for brilliant suggestions. Finally, thanks to the wonderful students for participating. All errors remain my own.

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S027226312200002X.

Data Availability Statement. The experiment in this article earned Open Materials and Open Data badges for transparent practices. The materials and data are available at https://osf.io/7tgrj/.

References


Appendix A

Stimuli list

<table>
<thead>
<tr>
<th>apagado</th>
<th>enamorado</th>
<th>pensado</th>
</tr>
</thead>
<tbody>
<tr>
<td>cansado</td>
<td>escuchado</td>
<td>pescado</td>
</tr>
<tr>
<td>cerrado</td>
<td>gustado</td>
<td>preguntado</td>
</tr>
<tr>
<td>comprado</td>
<td>hablado</td>
<td>preocupado</td>
</tr>
<tr>
<td>cortado</td>
<td>ocupado</td>
<td>quemado</td>
</tr>
<tr>
<td>encantado</td>
<td>pagado</td>
<td>reventado</td>
</tr>
<tr>
<td>enojado</td>
<td>pasado</td>
<td>trabajado</td>
</tr>
</tbody>
</table>