The Routledge Handbook of Spanish as a Heritage Language

Kim Potowski, Javier Muñoz-Basols

Heritage Spanish Phonetics and Phonology

Publication details
Rebecca Ronquest, Rajiv Rao
Published online on: 08 May 2018

How to cite :- Rebecca Ronquest, Rajiv Rao. 08 May 2018, Heritage Spanish Phonetics and Phonology from: The Routledge Handbook of Spanish as a Heritage Language Routledge
Accessed on: 13 Sep 2023

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
HERITAGE SPANISH PHONETICS AND PHONOLOGY

Rebecca Ronquest
NORTH CAROLINA STATE UNIVERSITY, USA

Rajiv Rao
UNIVERSITY OF WISCONSIN-MADISON, USA

Introduction

Research in the area of bilingual speech production and perception has a long history dating back decades, in which researchers have established that bilingual speakers’ past experiences with their languages, including age of acquisition of one or both languages, age of arrival to the country of immigration, and types of exposure to and usage of the first (L1) and second (L2) language, have a significant impact on shaping their phonetic and phonological systems (see Flege 1995; Flege & Eefting 1988, among many others). Until recently, however, very little work has directly examined the phonetic and phonological systems of heritage speakers, and more specifically, heritage speakers of Spanish (hereafter HSS) as they have been linguistically and culturally defined (see Potowski, this volume). The goal of the present chapter is to review investigations that examine the phonological production and perceptual patterns of HSS who were either born in the U.S. or moved here prior to the age of 6 and have experienced and continue to experience contact between Spanish and English.

In the past several decades, considerable literature in the field of linguistics has examined different components of the linguistic system of HSS. Morphosyntax and pedagogy have received considerable attention, most likely because they are associated with academic forms of language, particularly writing, which HSS typically do not develop as children. However, studies on the heritage Spanish phonetic and phonological system are limited in number, perhaps due to impressionistic observations that HSS often “sound” like native speakers of Spanish (NSS). Researchers such as Benmamoun, Montrul, and Polinsky (2010), however, have suggested the existence of a “heritage accent,” although at present it is not entirely clear what that might entail.

As a point of departure, some valid questions one may ask about the heritage Spanish sound system are: How are HSS similar to and different from NSS and age-matched L2 learners who acquired Spanish later in life? How might such similarities and differences contribute to HSS “sounding” like NSS? Or, conversely, how might they contribute to the perception of a “heritage accent?” This chapter is dedicated to addressing these and other related questions by
reviewing studies on the production of segmental (i.e., individual consonant and vowel sounds) and prosodic (i.e., intonation, rhythm, and stress) features in heritage Spanish, often in comparison to native Spanish and English trends. Where relevant, investigations of speech perception are also incorporated. Finally, the chapter concludes by discussing the theoretical importance of the body of work covered, the utility of laboratory methods, and specific challenges posed by research on HSS.

Consonants

Much of the initial work on the heritage Spanish consonantal system has focused on the voiceless and voiced stop consonants /ptk/ and /bdg/, respectively, as these groups of sounds have key phonetic and phonological differences in English and Spanish. The voiceless stops /ptk/, which are produced with a complete blockage of airflow and an absence of vocal cord vibration, differ with respect to voice onset time (VOT): the duration between the separation of articulatory organs and the initiation of vocal cord vibration associated with the following segment (Lisker & Abramson 1964). The time lapse between these two moments is typically longer in English than in Spanish, especially voiceless stops in word-initial position, resulting in English speakers realizing aspirated allophones [phθkʰ] with VOT > 50 milliseconds and a noticeable expulsion of air from the mouth upon production, and unaspirated allophones [ptk] typically with VOT < 50 milliseconds in other positions (Lisker & Abramson 1964). Spanish, however, is characterized by the presence of unaspirated allophones, and consequently shorter VOT with values ranging from 0–35 milliseconds in all positions.

Regarding the voiced stops /bdg/, Spanish and English differ in their realization of these sounds in two central ways. In English, vocal fold vibration usually begins at the time of the release of the stop, while Spanish voiced stops are often referred to as prevoiced because voicing is initiated before the release of the stop (Lisker & Abramson 1964). A second important difference concerning /bdg/ relates to the presence of weakened variants in most varieties of Spanish. NSS produce these as weakened variants in some positions, but as true stops in others. While some traditional literature (e.g., Quilis 1993) refers to these variants as “fricatives” (i.e., sounds produced with muscular tension that produces audible friction), more acoustically grounded research (e.g., Eddington 2011) argues that they should be classified as “approximants” (i.e., where organs approach one another with less tension and friction). The approximants, represented by the symbols [βðγ], typically occur between vowels, both within and across word boundaries. The stop allophones [bdg], in contrast, are manifested after pauses, nasals (i.e., /m/ and /n/) and /l/ (only in the case of /d/). In English, such weakening between vowels does not take place and stop allophones typically occur across the board. Of additional importance, the phoneme /b/ in Spanish typically corresponds with two written forms, <b> and <v> (i.e., botar “to bounce” and votar “to vote” have different meanings but are both phonemically /botar/), while in English /b/ and /v/ are separate phonemes (i.e., “ban” and “van” are pronounced differently and have different meanings).

The differences highlighted above between English and Spanish related to VOT and the weakening of /bdg/ have inspired what has been discovered about HSS’ consonants to date. Some of the earliest impressionistic (i.e., not acoustic) studies to compare HSS to L2 learners of Spanish revealed that HSS’ realizations of the voiced and voiceless stops more strongly resembled native speech and were judged as sounding more native-like than those produced by L2 learners (Au et al. 2002), and that /b/ and /d/ were judged as sounding more native-like than /g/ (Knightly et al. 2003). A related study (Au et al. 2008) expanded upon previous work by
distinguishing HSS as active (i.e., *childhood speakers*) or receptive (i.e., *childhood overhearers*) users of Spanish as children, but the overall results still showed that HSS performed more native-like than L2 learners.

More recently, studies on stop consonants have emphasized acoustic analysis utilizing software such as Praat (Boersma & Weenink 2017). Such techniques have allowed for a closer analysis of sounds, revealing additional similarities and differences between HSS and NSS as well as evidence of how language dominance may influence fine-grained phonetic detail. Kim (2011), for example, noted that while HSS and native speaker controls patterned similarly with respect to their perception of natural and manipulated VOT stimuli, their production of the voiced and voiceless stops differed. When compared to NSS, the HSS, who were English-dominant, produced the voiceless stops with significantly longer VOT and the voiced stops with significantly less prevoicing (both of which are more English-like). Conversely, their production of these same sounds in English did not differ significantly from the native English control group’s productions. In conjunction, these results not only demonstrate a clear influence of language dominance on speech production, but also indicate a mismatch between production and perception for this particular group of HSS. Furthermore, Amengual’s (2012) analysis of the VOT of /t/ showed that the primary VOT distinction between HSS and NSS occurred when /t/ appeared in words possessing a corresponding cognate in English, such as *teléfono* or *total*. That is, the HSS’ VOT increased and thus trended toward English-like aspiration in words that were more similar to English.

Additional studies focusing on voiced stops have revealed that linguistic and extralinguistic factors influence how HSS produce these sounds. Rao (2014) examined the production of Spanish and English intervocalic /b/ in reading and picture description tasks. The analysis was based on articulatory tension-based distinctions (Martínez Celdrán 1985, among others), which were used to phonetically classify productions as pure approximants (i.e., the most weakened), tense approximants (i.e., more organ tension than the pure form, resulting in a sound closer to the fricative [v]), and stops. The statistical analysis demonstrated that a lesser degree of past and present use of Spanish with family and in social networks, stressed syllables, reading tasks, and word boundaries all significantly decreased pure approximant realizations. These factors increased stop productions, while in the case of /v/, they raised the rate of tense approximants. Subsequently, Rao (2015) expanded upon Rao (2014) by analyzing all three voiced stops (/bdg/). The statistical outcomes of this second study conveyed that /b/ is unique when compared to /dg/: it showed distinctly higher rates of less-weakened forms. In sum, the overarching implications of Rao’s two studies are that visual access to orthography, English phonological processes/orthographic conventions, and prominent phonological contexts promoting articulatory strengthening (i.e., stressed syllables and word boundaries) all combined to reduce pure approximants, though increased past and present exposure to Spanish clearly minimized non-native-like realizations.

Although much of the research examining the pronunciation of consonants by HSS has focused on stop consonants, rhotics (i.e., “r-like” sounds) are also of interest given the different phonemic inventories and realizations of these sounds in Spanish and English. The tap rhotic /ɾ/, which is characterized by one vibration resulting from the tongue tip’s brief contact with the alveolar ridge, is a phoneme in Spanish that contrasts only between vowels with the trill rhotic /r/, which occurs through a series of moments of brief contact between the tongue tip and the alveolar ridge (Hualde 2005). The trill is subject to considerable dialectal variation in Spanish, which, in some regions, is realized as a velarized variant (see Lipski 1994 for a review). The tap is produced in English (e.g., in “better”), but is not a phoneme, and the trill does not exist in English.
Studies analyzing the production of rhotics in heritage Spanish have revealed a large degree of individual variation (Henriksen 2015) as well as a potential influence of social and dialectal factors (O’Rourke & Potowski 2016). Henriksen’s (2015) analysis of first-generation bilingual immigrants and second-generation Chicago area HSS of Mexican descent showed a vast array of rhotic realizations across speakers. While the rhotic contrast was generally maintained, it was done via modifications to segment duration rather than number of vibrations. The author suggests that this finding could be attributed to the large degree of phonetic variability associated with the production of the sounds in question, and not necessarily the result of English influence. O’Rourke and Potowski (2016) also assessed the production of rhotics in Chicago as produced by different generations of heritage speakers of Mexican, Puerto Rican, and MexiRican (i.e., mixed parentage) descent. They reported that, as a group, speakers of Puerto Rican and MexiRican heritage produced fewer velarized trills – a variant found almost exclusively in Puerto Rican Spanish – when communicating with a speaker of Mexican descent, but more frequently produced this variant in conversations with another Puerto Rican speaker. The authors concluded that this was further evidence of the stigmatized value of the velarized variant and, as such, it was utilized as a means to signal in-group membership.

While the majority of the investigations described have focused on how HSS produce consonants, understanding how they perceive the sounds of the heritage and/or majority language is equally important, given that the way in which sounds are perceived may influence how they are produced. Recall that Kim (2011) noted differences only in production, but not perception, of the voiced and voiceless stops when comparing HSS and native controls. Boomershine (2014) and Mazzaro, Cuza, and Colantoni (2016), however, suggest that while HSS often pattern closely with NSS in terms of their perception/discrimination of some consonantal and vocalic contrasts, they may not pattern identically. In an analysis of the perception of [d], [θ] and [r], Boomershine (2014) found that HSS, native Spanish bilinguals (i.e., L1 Spanish – late L2 learners of English) and monolingual Spanish controls all rated [d]/[r] and [r]/[θ] as sounding “more different.” HSS’ perception of [d] and [θ], however, differed from the monolingual controls: monolingual Spanish speakers found these sounds to be very similar, likely because they are separate phonemes in English. Mazzaro, Cuza, and Colantoni’s (2016) comparison of HSS’, long-term immigrants’, and recent arrivals’ (i.e., control group) discrimination of the voiced and voiceless stops and front and back vowels revealed that HSS were just as accurate as the control group in discriminating /p/ and /b/, /t/ and /d/, and vocalic contrasts, but patterned with the long-term immigrants in being less accurate with respect to /k/ and /ɡ/. Perception research on HSS therefore supports the notion that early exposure to the heritage language facilitates the formation of phonetic categories, although factors such as place of articulation and the phonological role of sounds in the heritage or majority language are clearly influential. As described by Kim (2011), there is also the potential for mismatches between perception and production; that is, even though contrasts might be perceived accurately, sounds may not necessarily be realized faithfully.

To recap, previous research on the production of consonants by HSS has demonstrated that variables such as the presence/absence of orthographic representation in tasks, cognate status, and strength of ties to English/Spanish affect the production of consonants, as do linguistic variables such as syllable stress and position relative to word boundaries. In particular, the presence of orthography, cognates, stressed syllables, and word boundaries have all led to the production of less native-like stops, especially in English-dominant HSS. Despite these non-native productions, comparing the overall body of work in this section to previous L2 studies on the same
topic (e.g., Díaz-Campos 2004; Zampini 1994) illustrates that HSS’ consonants trend more strongly toward those of NSS. This idea further supports the previously mentioned positive role that exposure to Spanish during childhood plays in more closely approximating a native-like phonology (Knightly et al. 2003). However, while the study of consonantal production has drastically increased in recent years, there is still plenty of room for further exploration. Additional perception-based studies are needed in order to better understand the potential contribution of consonants to a “heritage accent.” Finally, while the discussion of consonants to this point has been enlightening, further work is needed on classes beyond stops and rhotics, including focusing on /s/ voicing or the production of /l/.

### Vowels

Research on the heritage Spanish phonetic/phonological system has also included studies of how HSS produce and perceive Spanish vowels. The Spanish vowel system consists of five contrastive vowels /ieaou/, which are organized fairly symmetrically in the acoustic space (Martínez Celdrán 1995). In contrast, the English vowel inventory is fairly large, consisting of between 11 and 15 phonemic vowels that are characterized by considerable dialectal and social variation (Ladefoged 2005; Maddieson 1984). While Quilis and Esgueva (1983) noted some cross-dialectal differences in vowel production, most researchers agree that Spanish vowel systems do not vary as widely across geographic regions as do those of English (Hualde 2005; Ladefoged 2005; but see Lope Blanch 1972 and Delforge 2008 for dialectal variation). In addition, unstressed vowels in English are typically reduced and produced as a central vowel known as schwa (Delattre 1969; Ladefoged 2005). A representation of the English and Spanish vowel systems, adapted from Bradlow (1995), is presented in Figure 11.1. The approximate location of schwa is also included, although these values represent a synthetic schwa based on a male vocal tract (Johnson 2003) and not an actual speaker’s production.

Investigations of vowel production have reported a number of similarities and differences between the systems of HSS and NSS, most notably an asymmetrical acoustic distribution, considerable reduction of unstressed vowels, and significant effects of speech style on production. In comparison to the monolingual Spanish vowel system, HSS’ Spanish vowels have been found to exhibit a distinct organization within the acoustic space. Willis’s (2005) findings of a lowered /o/ and fronted /u/ in New Mexican Spanish were also reported in Boomershine (2012) and Ronquest (2012) for HSS in North Carolina and Chicago, respectively. The latter two studies also described a greater dispersion of the front vowel space relative to the back vowel space, as well as a backed and raised /e/. Such distributional differences suggest that the Spanish vowel system may not be as triangular or symmetrical as traditionally described for some HSS. There is also some potential that the more anterior position of Spanish /u/ in heritage vowel systems is the result of English contact, given that English /u/ is typically fronted relative to its Spanish counterpart (cf. Ladefoged 2005 for English; Quilis & Esgueva 1983 for Spanish).

One of the most consistent findings across studies of heritage Spanish vowels, and a similarity they share with L2 learners’ productions (see Menke & Face 2010), is the greater tendency for unstressed vowels (such as otrás and encontrar) to centralize than what has been described for monolingual varieties of Spanish. Alvord and Rogers (2014), Boomershine (2012), and Ronquest (2012, 2013) all reported that HSS’ unstressed vowels occupied a more centralized portion of the acoustic space when compared to their stressed counterparts. Additionally, Ronquest (2013) found that atonic vowels were significantly shorter than tonic vowels, but that the observed differences in vowel quality (i.e., centralization) were not the result of the shorter duration of atonic vowels. The prevalence of centralization, or movement of unstressed...
vowels toward the center of the vowel space, is particularly interesting in the context of bilingual speakers. Many traditional descriptions of Spanish vowels argue that stressed and unstressed vowels are similar with respect to their quality and duration (Delattre 1969; Hualde 2005). English unstressed vowels, in contrast, are frequently shortened and reduced to schwa (Delattre 1969; Ladefoged 2005). Although the reduction observed in these studies could be attributed, in part, to extensive English contact, Alvord and Rogers’s (2014) first-generation speakers, who were the most Spanish-dominant of the three groups analyzed, also reduced unstressed vowels. Their findings suggest that phonetic reduction may not be unique to the L2 and heritage vowel systems, but rather can emerge due to even very minimal contact with English and, therefore, merits further investigation.

Despite their differences, however, the vowel production of HSS and NSS seems to be similarly affected by speech style. Vowels produced in informal tasks (e.g., sociolinguistic interviews, narratives) were more condensed within the acoustic space than those produced in formal, controlled word lists and carrier phrases in both Alvord and Rogers (2014) and Ronquest (2012). These findings are consistent with research on monolingual Spanish varieties, which exhibit similar tendencies (Harmegnies & Poch-Olivé 1992; Poch, Harmegnies, & Martín Butragueño 2008). That is, vowel space expansion in controlled speech relative to more naturalistic/less controlled styles has been documented in both monolingual and bilingual varieties, even though the exact distribution and location of specific vowels within the acoustic space is not necessarily identical across groups. Of additional importance, style-induced variability seems to be fairly consistent across distinct groups of HSS, with Alvord and Rogers’s (2014) Miami Cuban bilinguals exhibiting similar patterns to those reported by Ronquest (2012) for Mexican and Puerto Rican HSS in Chicago.
In addition to examining how HSS produce Spanish vowels, two recent investigations have assessed the perception and discrimination of Spanish and English vowel contrasts. Mazzaro, Cuza, and Colantoni (2016) observed that heritage speakers, long-term immigrants, and recent immigrants did not differ significantly with regard to their discrimination of the front and back vowels (/i/ and /e/ and /o/ and /u/, respectively). Boomershine (2013), however, reported differences between monolingual speakers of Spanish and bilinguals (i.e., HSS and advanced, late L2 learners of English and Spanish) concerning their perception of English front vowel contrasts. Similarity ratings of pairs such as *bait*/bit and *beet*/bit revealed that HSS and the two bilingual speaker groups rated these pairs as “more different,” whereas the monolingual Spanish speakers perceived them as “more similar.” Combined, these two studies confirm that language exposure and dominance shape the perceptual system, resulting in sensitivity to vocalic contrasts in both languages.

In summary, research on the heritage Spanish vowel system, although still in its early stages, suggests that HSS resemble NSS in their production of Spanish vowels in some respects but differ in others. Stylistic variation, for example, appears to be consistent across HSS from different U.S. geographic regions and monolinguals alike. While the distributional differences and the presence of unstressed vowel reduction may suggest influence from English, direct comparisons across studies and to NSS should be done with caution. Research on Spanish vowels is limited compared to studies on consonants, with many investigations only examining Peninsular speakers’ production of stressed vowels in open syllables. Thus, the differences described may not be the result of contact with English, but rather representative of purely phonetic, dialectal, or methodological differences. In addition, as HSS are exposed to a highly variable input variety that might already contain some contact-induced features, direct comparisons with NSS may be somewhat misleading. Assessing similarities between HSS and their caretakers or, as pointed out by Menke and Face (2010), to other groups of bilingual speakers, may offer more useful insight into bilingual vowel production. Moving forward, careful selection of comparison groups and the addition of perceptual measures, along with those of production, may help determine which aspects of vowel production and perception are characteristic of HSS in general.

**Prosody**

Prosody is perhaps the least explored area of the phonetic and phonological systems of HSS. It is divided into three main areas: intonation, which refers to modifications to pitch that influence the interpretation of meaning, for example, in statements versus questions and when conveying different attitudes and emotions (Hualde 2005); stress, which deals with the cuing and perception of relative prominence in speech (Hualde 2005); and rhythm, which generally addresses the distribution of stressed and unstressed moments in speech across time (Arvaniti 2012). Regarding rhythm, it has historically been associated with at least a two-way division based on duration: stress-timed languages (e.g., English) have relatively similar durations between stressed syllables, while syllable-timed languages (e.g., Spanish) tend to exhibit relatively similar durations of stressed syllables themselves (Pike 1945; see Arvaniti 2009, 2012 for an in-depth discussion of the topic of rhythm/timing and the problematic aspects of it).

Specifically with regard to intonation, the few existing phonological studies on HSS are couched in the Autosegmental-Metrical (AM) model of intonational phonology (Ladd 2008; Pierrehumbert 1980). This model proposes that mental representations of intonation, or phonological targets above the level of individual segments, are comprised of a single high/low tone associated with a pitch peak or valley, respectively, or sequences of these tones. These tones or tonal sequences, occurring at the word-level and associating with stressed syllables, are deemed
pitch accents. Such monotonal or bitonal targets can also be boundary tones tied to the edges of phrases, which are ways in which prosody chunks portions of discourse into units. The pitch accent and boundary tone sequences at the ends of utterances are particularly important for pragmatic purposes; for example, pitch rises, falls, or rise-fall movements corresponding with the end of an utterance in Spanish can be used to distinguish a question from a statement or a neutral statement from a more emphatic statement.

Previous studies comparing the intonational systems of HSS to those of bilingual and monolingual native speakers have uncovered significant influences of factors such as utterance type, contact between varieties of Spanish and English, experience with and attitude toward Spanish, and task type. In Alvord’s (2006, 2010a, 2010b) examination of the intonation of three generations of Miami Cuban bilinguals, the third of which was comprised of HSS, the presence of both final rises and final falls across the data was the most noteworthy finding. A statistical look at question patterns revealed that first- and third-generation speakers preferred final falls, while those of the second generation clearly favored rises. Regarding the third generation in particular, analyses suggested that both the affirmation of Cuban identity and potential influence from English motivated the prevalence of final falls. Along similar lines, Henriksen’s (2012) work on intonation in Chicago area HSS and NSS also reported more key intergroup differences in questions as compared to statements, where HSS produced a wider range of configurations. The author posited that increased contact between English and Spanish could have yielded an innovative inventory of intonational strategies.

Robles-Puente’s (2014) study on prosody in Los Angeles incorporated two groups of HSS that differed in age: adults either born in Los Angeles or who emigrated from Mexico as very young children, and adolescents and children born in Los Angeles to immigrant parents from Mexico. HSS’ trends in English and Spanish did not differ significantly; the same tonal configurations occurred in both of their languages. In terms of intergroup differences, while both groups of HSS exhibited phonological targets used by the NSS control group, the younger group produced intonational trends more closely approximating those of the controls. Rao’s (2016) study is smaller in scale but similar to that of Robles-Puente (2014); however, Rao’s study focused more on details concerning participants’ past and present use of Spanish, as well as demographic information about both the participants and their parents. Like Alvord (2006, 2010a) and Henriksen (2012), Rao also noted increased variation in question intonation and suggested that exposure to unique types of input at the individual speaker level plays a role in the observed intonational differences across the data (see also Pascual y Cabo & Rothman 2012).

Finally, another recent investigation on HSS intonation, which is the first to emphasize task type, is Colantoni, Cuza, and Mazzaro (2016), who examined the statement intonation of HSS and older immigrants in reading and narrative tasks. They found significant differences in the pitch accents of both groups only in the reading task, which supports the orthography-related findings described in the previous section. The authors suggest that the lack of formal education, and thus of literacy skills in Spanish, led to the production of read samples that were less natural. As such, it was recommended that researchers beware when implementing reading tasks with HSS. Lastly, the authors extrapolated the results to the pedagogical level by providing some strategies for implementing prosody practice in the heritage Spanish classroom.

Moving onto stress, two notable studies of HSS have revealed perception versus production discrepancies and the need to look at prosody’s interaction with other areas of linguistic competence. Kim (2015) compared the perception and production of lexical stress by HSS, NSS, and 4th semester L2 learners using verb minimal pairs (e.g., pasó/paso; stressed vowels are bolded). The HSS patterned with the NSS in their accurate perception of penultimate (e.g., pasó) and final (i.e., pasó) stress, whereas the L2 learners were not able to distinguish between the
two stress patterns because they were biased toward the penultimate pattern. In the production experiments, however, the HSS and L2 learners often lengthened unstressed vowels in \textit{paso}-type words, effectively producing a \textit{pasó}-type word (i.e., final stress) instead. The latter tendency was not observed in the NSS’ productions. In a different approach to stress, Hoot (2012) revealed that HSS and Mexican NSS shifted stress leftward in a similar fashion when communicating presentational focus/new information (e.g., \textit{Mi mamá compró un carro} rather than \textit{Compró un carro mi mamá} was judged as acceptable in response to ¿Quién compró el carro?), suggesting that, at least for the sample of speakers in question, prosodic prominence in Spanish is flexible and does not necessarily have to fall in rightmost position. This analytical approach is unique in that it discussed the specific interface of prosody and syntax, extended the discussion of interfaces between linguistic areas to a general level, and argued for the inclusion of more interfaces in future examinations of heritage grammars.

Regarding differences related to vowel duration, which is a common metric used to quantify speech rhythm, Robles-Puente (2014) conducted an analysis to see if HSS’ speech was more stress-timed or syllable-timed. It was shown that adult HSS produced English trends when speaking either language, whereas younger HSS born in LA accommodated their measures to the favored pattern of the language being spoken. In sum, when taking into account the differences between the findings for issues related to timing and those tied to intonation, Robles-Puente concluded that components classified within the general area of prosody seem to exhibit distinctions in bilingual grammars.

After reviewing the relatively small body of studies on the prosodic system of HSS, it is clear that HSS pattern similarly to NSS in some respects, but the majority of studies report considerable differences, namely regarding question intonation, stress production, and rhythm. The cumulative findings have shed light on the contribution of prosodic factors to differences between heritage and native grammars above the level of individual sounds. Such differences could be prosodic features contributing to a “heritage accent.” However, once again, there is a lack of work from the perspective of perception, which, as seen at the segmental level, is begging to be addressed for prosody as well. Additionally, since intonation in particular exhibits a tremendous degree of dialectal variation (see Prieto & Roseano 2010), examining this specific part of prosody in HSS descending from a wider range of Spanish-speaking countries beyond Mexico and Cuba is a future research direction deserving of attention. Finally, stress is an important prosodic concept to continue investigating because it has pedagogical implications; that is, conscious awareness of stress in speech could assist many HSS with writing accent marks in Spanish.

\textbf{Future research considerations}

The investigations summarized in this chapter have revealed a number of important similarities and differences between HSS and NSS (and, to a lesser degree, adult L2 learners of Spanish) with regard to their production and perception of Spanish segmental and prosodic features. This section expounds upon the theoretical importance of this work, speaking specifically to the merit of studies of speech perception and the utility of perceptual models in understanding the heritage Spanish phonetic/phonological system. In addition, it comments on the usefulness of laboratory approaches in examining pronunciation and makes suggestions on how researchers might tailor their methodologies based on participants’ experience with formal, written Spanish. Finally, the section concludes by discussing the challenge of addressing individual variation in heritage populations and revisiting the questions posed in the introduction.

The majority of the works summarized in this chapter have focused on investigations of speech production, which have served to characterize and provide acoustic-phonetic details
Heritage Spanish phonetics and phonology

about HSS’ pronunciation of certain individual sounds, syllables, phrases, and utterances. Studies of speech perception, however, are equally important in augmenting our understanding of the underlying phonological system(s) of HSS and offering insight into why the production of certain sounds may prove difficult and/or differ from previously established norms. Models of speech perception (e.g., Speech Learning Model (SLM), Flege 1995; Perceptual Assimilation Model (PAM), Best 1995; Native Language Magnet (NLM), Kuhl 1993) have been widely employed in research examining cross-language perception and L2 phonological acquisition. Although the specific tenets of these models differ and will not be described in detail here, in very general terms, each posits that the way in which L2 and/or non-native speech sounds will be perceived, and in the case of the SLM, subsequently produced, depends on how they are processed with reference to L1 phones, articulatory gestures, or exemplar prototypes. Drawing upon one or more of these models to explore the sound systems of HSS, and including experiments of speech perception as well as production (as done by Kim 2011, 2015), will allow researchers to: (1) understand how HSS’ unique experience with the heritage and majority languages has shaped their perceptual system, and (2) move from purely descriptive characterizations of speech sounds to predicting how and why a particular sound or group of sounds may vary based on their perceptual (dis)similarity with competing sounds of the majority language.

In terms of methodological approaches, the use of “natural” or “spontaneous” speech versus “laboratory” speech samples in the field of phonetics has received much attention in recent years. Although laboratory speech has been criticized as “unnatural” and not representative of how speakers may use language in informal settings, Xu (2010) argues in favor of these approaches, describing some of the advantages of analyzing speech “sampled under experimental control” (p. 330). Speech obtained in naturalistic settings, for example, may be either too poor in quality to analyze acoustically or may be lacking in a sufficient number examples of the segment under investigation. Employing controlled speech tasks therefore allows researchers to isolate the features under study, effectively removing many potentially confounding factors that perhaps could not be eliminated from data obtained under less-controlled conditions. One of the first steps in understanding what might constitute a “heritage accent” is isolating and characterizing individual segments or suprasegmental features – an undertaking that can very likely be achieved more easily by a carefully designed data elicitation instrument than in a sociolinguistic interview. Once detailed descriptions have been obtained, we may consider extending analyses to include a wider range of speech styles.

While the analysis of laboratory speech samples in phonetics research has been successfully implemented in the studies reviewed here, and hopefully will continue to be used, it is necessary to emphasize the importance of carefully designing data elicitation instruments through a detailed consideration of participants’ experience with the heritage language. Some controlled tasks, especially those that involve reading, can be particularly problematic for some HSS who have limited experience with Spanish in its written form (Colantoni, Cuza, & Mazzaro 2016), and may result in orthographically influenced pronunciations (e.g., [v] for <v>). Additionally, since many HSS are not formally educated in Spanish, and thus are not accustomed to exercising their literacy skills in this language, they may find reading tasks particularly alien, which may lead to disfluencies that can impede the analysis of target features.

Given the heterogeneity of heritage speaker populations, one of the biggest challenges in linguistic research on HSS concerns how to contend with individual differences. Extralinguistic factors such as age of onset of bilingualism, length of time in contact with the majority language, and self/formal assessment of proficiency and use of the majority and minority languages have been useful in describing some of the observed variation in speech production and perception.
Grouping HSS based on sociolinguistic generation (i.e., G2, G3), as done by Henriksen (2015) and O’Rourke and Potowski (2016), and conducting longitudinal and/or cross-generational studies of heritage speakers (Nagy & Kochetov 2013), may also prove insightful, allowing for a closer examination of how phonetic categories or mental representations of sounds vary within groups of HSS and how they may change over time. Second language acquisition research has also addressed the impact of cognitive factors such as motivation (see Dörnyei & Skehan 2003), which have, thus far, not been applied to HSS. Collectively, these approaches may permit additional commentary on questions such as incomplete acquisition, attrition, and the role of input, which have not been addressed with specific reference to heritage speakers’ phonetic/phonological systems as widely as they have been for other components of their grammars.

Another potential source of individual variation that is particularly relevant to phonetics research concerns the inherent variability of the speech signal. That is, the inter- and intra-group differences observed in pronunciation may stem, in part, from the fact that the segments themselves exist along a continuum of possible realizations. As noted earlier, Henriksen (2015) argues that some of the variation observed in the production of rhotics may not be caused by English influence, but rather by the fact that Spanish rhotics encompass a variety of possible articulations. The same logic could be applied to voiced approximants, which have been shown to be continuous in nature (e.g., Eddington 2011), as well as voiceless stops, which are characterized by a range of possible VOT values (Lisker & Abramson 1964). Examining individual patterns of production, as did Henriksen (2015), will allow for a more thorough assessment of phonetic variability, and could potentially reveal important individual differences that might be obscured by looking at aggregate data alone. Ultimately, such an approach may aid in teasing apart individual variation attributable to extralinguistic factors from that which is due to purely phonetic factors.

Conclusions

To revisit the questions posed in our introduction, we have described that while HSS seem to pattern more closely with NSS in some respects (e.g., lenition of /bdg/, style-induced variability in vowel production), they also exhibit characteristics resembling those attested in L2 learner speech (e.g., [v]-like production of <v>, lengthened VOT, unstressed vowel reduction). It is perhaps because of this mix of features that we are faced with two seemingly contradictory observations about heritage Spanish pronunciation: “sounding” like NSS on the one hand but having a “heritage accent” on the other. Although the conjunct of present works does not allow us to provide a definitive answer to these highly complex questions at this time, it is clear that HSS’ production of their heritage language places them on a continuum situated somewhere between L2 learners and monolingual NSS.

Combined, the phenomena discussed in this chapter have laid a strong foundation in the area of heritage Spanish phonetics/phonology, offering valuable insight into how HSS both produce and perceive the sounds of their heritage language. They have also raised further questions that researchers will hopefully strive to answer in the upcoming years in an effort to continue advancing our knowledge of heritage Spanish grammars as a whole.

Note

1 Vowel production may also depend on speakers’ location within the U.S. For example, work on English in the South has shown that some Spanish-speaking immigrants and their descendants acquire southern monophthongization (Wolfram, Carter, & Moriello 2004).
Further reading


An alternative but related examination of the rhythmic properties overviewed here.


An overview of extralinguistic factors and how they influence the L2 sound system.


A test of the predictions made by the SLM and PAM with respect to the perception of stop voicing in Spanish.


An innovative analysis accounting for variation in L2 learners’ weakening of /bdg/.

References


Boomershine, A. (2012, October). *What we know about the sound system(s) of heritage speakers of Spanish: Results of a production study of Spanish and English bilingual and heritage speakers.* Paper presented at the 12th *Hispanic Linguistics Symposium,* Gainesville, FL.


Rebecca Ronquest and Rajiv Rao


Heritage Spanish phonetics and phonology


