Testing for underlying representations: Segments and clusters in Sevillian Spanish

Abstract

The current study investigates the representational status of [Ch] sequences in Sevillian Spanish. Like many Spanish varieties, Sevillian debuccalizes coda /s/ to [h] (/sC/ → [hC]). Unlike other varieties, Sevillian is undergoing a change whereby [hC] sequences are variably realized as [Ch] sequences. I argue that surface (phonetic) [Ch] sequences are metathesized versions of underlying /sC/ clusters, and have not phonologized into a new series of aspirated stops /p^h, t^h, k^h/. Evidence supporting the /sC/ cluster analysis comes from a perception study in which Sevillians reconstruct /s/ on the word preceding [Ch] sequences. Listeners of other dialects do not attribute [h] in [Ch] sequences to a preceding /s/. I present a brief cross-dialectal analysis corresponding to the participant groups' different behavior in the perception experiment, in which the dialects share underlying representations but map them differently to surface forms. I also discuss reasons the aspirated stop analysis may be implausible, from both a learning and analytical perspective. The findings have broader implications for our understanding of the mapping between underlying and surface forms, segments, and the connection between dialect variation and perception.

1 Introduction

There is an ongoing change in <sp, st, sk> sequences in Sevillian Spanish. In these sequences, Sevillian debuccalizes coda /s/ to [h], like many varieties of Spanish (/sC/ → [hC]). The process unique to Sevillian is the ongoing change from these *h-stop* sequences to *stop-h* sequences ([hC] → [Ch]) (Torreira 2006; O'Neill 2010; Parrell 2012; Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016). I refer to this change as metathesis. The examples in (1) illustrate both debuccalization of coda /s/ and metathesis.

(1) Sevillian h-stop and stop-h sequences

		Debuccalization	Metathesis	
a.	/ˈtʃispa/	[ˈtʃihpa]	[ˈtʃipha]	'spark'
b.	/ˈpista/	[ˈpihta]	[ˈpitha]	'track'
c.	/ˈboske/	[ˈbohke]	[ˈbokhe]	'forest'

Sevillian Spanish is part of a broader dialect region called Western Andalusian Spanish. Metathesis has been reported in many varieties in this region. Varieties spoken in the neighboring region of Eastern Andalusia do not have robust metathesis, but may be moving in this direction (Ruch and Harrington 2014). Although there are dialectal differences even within Western Andalusia, I use *Sevillian Spanish* to refer to the variety spoken in Seville

1

and other areas of Western Andalusia where similar patterns hold. I use *Andalusian Spanish* only when necessary to refer to properties that hold of the entire dialect region (both Eastern and Western Andalusian), and when the original sources do not specify region.

The forms of coda /s/ in (1) represent both diachronic evolution (retention → debuccalization → metathesis) and synchronic variation within Sevillian Spanish. Diachronically, coda /s/ has undergone reduction in many varieties of Spanish, resulting in debuccalization to [h] and/or full deletion (e.g., Mason 1994). The change from h-stop to stop-h sequences is unique to several dialects in Southern Spain, including Seville. The metathesis change may not be as recent as typically assumed (Moya Corral and Tejada Giráldez 2020: 205–6), but within- and cross-dialect evidence suggests that coda /s/ reduction (and the accompanying lengthening of the stop following /s/) preceded metathesis (Ruch and Harrington 2014). Synchronically, the realization of /s/ in /sp, st, sk/ sequences is also quite variable in Sevillian. In addition to debuccalization and metathesis (1), other attested forms include (and are not limited to) various combinations of metathesis, /s/ deletion, and gemination of the consonant following /s/ (e.g., [patːha], [patha], [patːa]) (Ruch 2008; Ruch and Harrington 2014).

The Sevillian stop-h sequences that result from metathesis are representationally ambiguous: they could be realizations of /sC/ clusters, or of underlying aspirated stops /Cʰ/.[1] In some languages, an [h]-like sound is analyzed as part of the adjacent stop (e.g., English, [pʰɑt]). In others, [h] is analyzed as a segment in its own right (e.g., Acehnese; Asyik 1987). Sevillian [Ch] sequences have been treated, on one hand, as underlying /sC/ clusters that have undergone debuccalization and metathesis (Parrell 2012; Torreira 2012; Ruch 2013). Under this analysis, the change in pronunciation has not resulted in a change of underlying phonological representation. On the other hand, it is also possible that Sevillian [Ch] sequences may be phonologizing into aspirated stops (/pʰ, tʰ, kʰ/) (O'Neill 2009; Gylfadottir 2015). Although few argue for this position, I consider it seriously because sound changes that lead to changes in underlying representations (and new segmental contrasts) are uncommon, as is the ability to study them while the change is ongoing. If the surface form is phonetically similar to an aspirated stop, why don't learners posit this phonological representation? Neither side of the debate presents convincing experimental evidence, or considers what an aspirated stop analysis would mean for the phonological system, how such a reanalysis might happen, or what might prevent it.

The goal of this paper is to test the representational status of Sevillian [Ch] sequences. Section 2 begins by discussing the representational status of [Ch] sequences cross-linguistically, how Sevillian stop-h sequences fit into this classification, and the acoustic and perceptual properties of Sevillian stop-h sequences. Section 3 presents a forced-choice perception study designed to test the representational status of these sequences in Sevillian. Results show that Sevillian listeners attribute [h] in [Ch] sequences to a morphological marker /s/ on the preceding word. That is, they treat [C] and [h] as separable, and attribute them to different morphemes. Based on these results, I argue that Sevillian [Ch] sequences are rep-

---

[1]In principle, they could also be realizations of an underlying /Ch/ cluster. I do not discuss this representation here for reasons described in later sections—namely, that tautosyllabic /Ch/ clusters are cross-linguistically rare, and that there is surface variation in Sevillian between [.Ch] ∼ [s.C] ∼ [h.C]. The latter two surface forms do not lend themselves to a /Ch/ representation. These reasons are further discussed in Section 2.

resentationally /sC/ clusters (not /C^h/ segments), and I treat the mapping from /sC/ → [Ch] as metathesis (following Ruch 2013). When Sevillian listeners hear [Ch] sequences, they 'undo' the metathesis to arrive at underlying /sC/. In contrast, listeners of other varieties of Spanish (Mexican and Argentinian) do not perform this same mapping when they hear [Ch] forms, essentially ignoring the presence of [h].

The fact that performance on the perception task depends on listeners' native dialect indicates that, although the dialects share an underlying /sC/ representation, their phonological grammars differ. Their grammars map /sC/ to surface forms differently in production, and listeners whose dialects do not map /sC/ to [Ch] in production have difficulty perceiving this form. Section 4 presents a cross-dialectal analysis in which different constraint rankings in each dialect produce different outputs in production. These differences in production are mirrored in the results of the perception experiment. I also suggest that a promising avenue for a more thorough formal analysis is to treat metathesis as gestural overlap and reorganization, rather than as discrete segmental transposition. An analysis of this type could account for several properties of metathesis that are difficult to capture in a more traditional, segment-based, featural analysis.

Finally, Section 5 discusses several factors that might prevent Sevillian [Ch] sequences from coming to be represented as aspirated stops. While many arguments for sequences of sounds being single segments vs. clusters are based on phonotactic and phonetic considerations (e.g., Trubetzkoy 1939; Riehl 2008), I examine how sociolinguistic variation and phonological alternations may also provide evidence that [h] in [Ch] sequences is independent from the stop and that it can belong to a distinct word or morpheme. These properties may make it difficult for listeners to interpret [h] as a contrastive feature of the stop itself. I also consider what an aspirated stop analysis would require. The difficulty of constructing a sensible analysis assuming underlying aspirated stops, as well as the oddness of the resulting phonological system, further suggest that such an analysis—formally developed or posited by children—is implausible. Section 6 concludes.

2 Background

2.1 Segments and clusters

Whether a sequence of sounds is analyzed as a single segment or cluster depends on a variety of factors, including syllable structure, phonotactics, and phonetics (e.g., Trubetzkoy 1939; Riehl 2008; Gouskova and Stanton 2021). Most of the debate about the segmenthood of [h] centers on *preaspiration*, where [h] is realized before the stop, as in Icelandic (Haugen 1958; Thráinsson 1978; Arnason 1986; Suh 2001). I focus on sequences where [h] *follows* a stop, which are the relevant sequences for Sevillian.

Finnish provides a straightforward example of stop-h sequences whose properties point towards a cluster analysis (Suomi, Toivanen, and Ylitalo 2008: 57). Stop-h sequences are heterosyllabic and arise only at morpheme boundaries (*saat-han* 'you do get, don't you'). They do not arise morpheme-internally (e.g., **lathi*) or word-initially. Both of these properties suggest that [h] and [C] belong to different morphemes, and thus form clusters at the underlying level.[2]

---

[2]Cross-linguistically, distinguishing between segments and clusters at the *surface* level is difficult because

Tautosyllabic stop-h sequences are most often analyzed as single-segment stops, but they can also be analyzed as clusters. English tautosyllabic stop-h sequences ([ph, th, kh]) are analyzed as single segments: they have been argued to be allophonically aspirated realizations of underlying voiceless stops specified with the [spread glottis] feature (Iverson and Salmons 1995; Honeybone 2005; see Beckman, Jessen, and Ringen 2013 for overview of alternative proposals). In monomorphemic contexts, these sequences count as simple onsets.[3] In a word like *intelligence*, [th] is tautosyllabic ([ɪnˈ.t̲h̲ɛlɪdʒəns]; *[ɪnt̲.ˈh̲ɛlɪ.dʒəns]). English allows complex onsets, but not with [h] as the second consonant. Thus, [th] is best analyzed as a single segment instead of a cluster. Analyzing [ph, th, kh] as underlying clusters would require an arbitrary exception in which /h/ could form clusters only with voiceless stops.

In contrast, tautosyllabic stop-h sequences have been argued to be clusters in several languages, including Acehnese (Austronesian), Old Khmer (Austroasiatic), and Taba (Austronesian). Acehnese [ph, th, kh] sequences can be split by an infix *-eun-* (2a) (Asyik 1987: 16, 112). Analyzing [Ch] as a cluster of two independent segments allows them to split. If [Ch] were analyzed as a single-segment aspirated stop, it would be difficult to explain why an infix splits [h] from the stop.

(2) Languages with [Ch] sequences analyzed as clusters

    a. meu-k̲-eun-eu-h̲ie   'it tastes a little bit like bad cooking oil'   Acehnese

       (from *k̲h̲ie*)   (Asyik 1987: 16)

    b. [k̲am h̲əŋ]   'anger'   Old Khmer

       (from [k̲h̲əŋ] 'to be angry' + -m-)   (Schiller 1994: 314–5)

    c. [ˈm̲h̲onas]   'sick'   Taba

      [ˈn̲h̲ik]   'bat'   (Bowden and Hajek 1999)

      [ˈk̲h̲an]   'I go'

Old Khmer and Taba show similar patterns, also suggesting that the stop and [h] form clusters. In Old Khmer, like in Acehnese, some infixes split the components of the [Ch] sequence (2b) (Schiller 1994; Minegishi 2006). Furthermore, the language allows many onset consonant clusters, so analyzing [Ch] as a cluster fits with broader phonotactic patterns (Minegishi 2006). In Taba (2c), Bowden and Hajek (1999) analyze stop-h sequences as clusters because [h] forms onset sequences with many consonants, not just voiceless stops. Furthermore, word-initial [Ch] sequences often arise from single-consonant prefixes, so the components come from different morphemes. They take these phonotactic patterns—and the fact that the components come from separate morphemes—as evidence for a cluster analysis. Languages like Acehnese, Old Khmer and Taba, where tautosyllabic stop-h sequences are analyzed as clusters (as in (2)), are rare. This may be because stop-h sequences would be confusable with aspirated stops.

---

phonetic diagnostics do not always correlate with hypothesized segmental status (see Section 5.3).

   [3]English does have stop-h sequences across morpheme boundaries, e.g., *top-hat*. These are heterosyllabic, and are clearly clusters because the two segments come from different morphemes (as in Finnish).

2.2   Spanish segments and clusters

Sevillian [Ch] sequences are ambiguous between a phonological representation as /sC/ clusters, and an aspirated stop representation (/C^h/), the latter of which is closer to their current surface form. There are two related—but distinct—questions related to segmenthood. First, are stop-h sequences representationally clusters (/sC/) or single segments (/C^h/)? Second, if they are representationally single segments, is aspiration contrastive?

First, I argue that stop-h sequences are realizations of underlying /sC/ clusters. Although the surface form has changed, the underlying form has not. Stop-h sequences clearly originated from /sC/ sequences, possibly due to the same pressures that have been acting on coda /s/ for centuries, so I will briefly discuss the conditions under which they are argued to have arisen. Diachronically and orthographically, stop-h sequences are /sp, st, sk/ (<sp, st, sk>) clusters. In these clusters, /s/ is syllabified as a coda and /p, t, k/ is syllabified as the onset of the following syllable. Stop-h sequences arise from coda /s/ reduction and metathesis. Coda /s/ reduction in Spanish is extensive (Hammond 2001), and variants can be seen as steps on a reduction continuum ([s] → [h] → ∅; Mason 1994). Diachronically, coda /s/ has reduced in Spanish since its divergence from vulgar Latin (Labov 1994: 583–585; Mason 1994). Some have argued that this change was due to a general preference for open syllables and a pressure against coda consonants in modern Spanish (Malmberg 1965; Catalán 1971). Synchronically, coda /s/ variation appears to be stable in most dialects; variants include [s.C], [h.C], and deletion (Labov 1994: 584). Dialects differ in which variant is most common, but all dialects have socially and linguistically-conditioned variation (see Erker 2012 for a thorough overview).

Debuccalization, deletion, and metathesis can be seen as solutions to the same problem. If the goal of coda reduction is to remove material from the coda, then metathesis accomplishes this goal by moving /s/ instead of deleting it altogether (Moya Corral 2007; Vida Castro 2016; Moya Corral and Tejada Giráldez 2020). In contrast to most varieties of Spanish—where coda /s/ realizations are in stable variation—the patterns of variation in Sevillian coda /s/ indicate an on ongoing change towards metathesized variants. Ruch and Peters (2016) find that younger speakers produce longer stop releases in /sp, st, sk/ sequences than older speakers, and younger listeners are more sensitive to stop release duration as a cue to /st/ than older listeners. Furthermore, there is an additional ongoing change from [th] to [ts] in /st/ sequences, which has been reported in several varieties of Western Andalusian Spanish (Ruch 2010; Vida Castro 2016).

The results of my experiment suggest that [Ch] sequences still correspond to underlying /sC/ clusters, as has diachronically been the case. Assuming this representation, the phonological grammar maps /sC/ to [Ch]. This is the analysis that I outline in Section 4.

Second, is aspiration underlyingly contrastive? Because I treat [Ch] as a surface representation of an underlying /sC/ cluster, this question is not relevant. However, it is important to consider the alternative that [Ch] is a contrastively aspirated stop /C^h/ at some stage of the derivation, and I return to this debate in Section 5.3. For now, I note that if the aspirated stop analysis were right, it would be an oddity in the development of laryngeal contrasts, both within and outside of Romance languages (no Romance languages have aspirated stops or three-way laryngeal contrasts). The aspirated stop analysis becomes even less plausible when morphological complexity is taken into account. My experiment takes advantage of

this morphological complexity to diagnose the underlying cluster representation instead.

## 2.3 Previous experimental studies on Sevillian stop-h sequences

The goal of my perception experiment is to test the underlying representation of [Ch] sequences in Sevillian. The experiment and analysis build on previous production and perception studies on Sevillian Spanish, which provide the basis for stimuli design, the choice of participant groups, and the predictions.

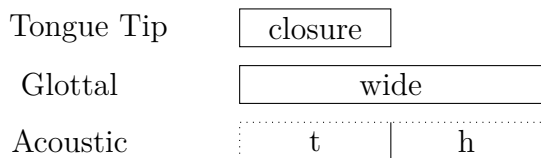### 2.3.1 Production and acoustics of Sevillian stop-h sequences

Most descriptions of Sevillian metathesis are gestural in nature. For example, Torreira (2006), Parrell (2012), and Cronenberg et al. (2020) argue that metathesis occurs by gestural realignment: a closure gesture slides leftward across a wide glottal gesture, resulting in [Ch] instead of [hC] in the acoustics (3a-3b).

(3) Sevillian metathesis occurs by gestural realignment (adapted from Parrell 2012: 38)

    a. Gestural score for [ht] sequence

Tongue Tip		closure		
Glottal		wide		
Acoustic	h		t	h
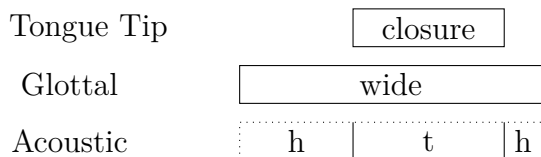
    b. Gestural score for [th] sequence

Tongue Tip	closure		
Glottal		wide	
Acoustic	t		h

    Acoustic studies show that the change from h-stop to stop-h is acoustically real, and that the realization of /sC/ sequences is highly variable. Multiple acoustic cues distinguish /sp, st, sk/ and intervocalic /p, t, k/: pre-posed [h] ([hC]), gemination of the consonant following /s/ ([hC:]), and post-posed [h] ([Ch]) (Torreira 2006; O'Neill 2009; Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016).[4] These cues are variable, can co-occur, and are sometimes found to show durational trading relationships (e.g., Parrell 2012; Torreira 2012; Ruch and Harrington 2014). Some variants of Sevillian /sC/ are shown in (4) (adapted from Ruch 2008: 33–4; Ruch and Harrington 2014). Parentheses indicate that the cue is variably present.

(4) Variants of Sevillian <sC> clusters (/pasta/ 'pasta')

---

[4]Previous studies refer to [h] in [Ch] as VOT and *postaspiration*. I do not, because this term implies that [h] belongs to the stop, an analysis I argue against.

a.	[pasta]	Retention	d.	[pat(ː)ha]	(Gem.) + metathesis
b.	[paht(ː)a]	Debucc. + (Gem.)	e.	[patha]	Metathesis
c.	[pahtha]	Surrounding	f.	[pat(ː)a]	Deletion + (Gem.)

Less investigated acoustic cues to /sp, st, sk/ include breathiness on the preceding vowel (Torreira 2007; O'Neill 2010; Torreira 2012), resistance (or lack thereof) to intervocalic voicing during the stop closure (O'Neill 2010; Galvano and Henriksen 2021), and spectral properties of metathesized /s/ (Ruch 2008; Vida-Castro 2015; Del Saz 2019).[5] While some of these findings are based on acoustic analyses, others appear to be based on impressionistic observations. Because the duration of [h] following the voiceless stop has been established as the main cue to /sC/ sequences (Ruch and Peters 2016: 25), my perception study manipulates only this cue.

The experiment also relies on the observation that metathesis occurs across word and morpheme boundaries (5) (Ruch 2008; Horn 2013). The word boundary context is ideal to test the representational status of [Ch] sequences, because both representational options are possible. Surface [Ch] can map to an underlying /sC/ cluster, where /s/ is part of the preceding word ([tjene phali] → /tjene̱s pali/), or [Ch] can map to /Ch/, where the preceding word ends in a vowel ([tjene phali] → /tjene p̱hali/). Both /tjenes/ and /tjene/ are real words ('have-2SG' vs. 'have-3SG', respectively). By asking listeners about the word preceding [phali], it is possible to distinguish between the representations. The experiment would not work with morpheme internal contexts like [patha] 'pasta', because there is no clear way to tell whether listeners map [th] to /st/ or /th/. The rate and extent of metathesis differ by word boundary context (Ruch 2008; Horn 2013), but what matters is that metathesis occurs regularly even across word boundaries.

(5) Metathesis morpheme-internally and across morpheme and word boundaries

a.	/pasta/	['patha]	'pasta'	Morpheme-internal
b.	/des-tapar/	[detha'par]	'to uncover'	Morpheme boundary
c.	/mas toɾta/	[ma 'thoɾta]	'more cake'	Word boundary

2.3.2 Perception of Sevillian stop-h sequences

In perception, previous studies on Sevillian stop-h sequences establish that: (1) Sevillians perceive a distinction between stop-h ([VC̱hV]) and intervocalic voiceless stops ([VC̱V]); (2) Sevillians are better than other listeners at distinguishing stop-h ([ɡaḵho]) from intervocalic voiceless stops ([ɡaḵo]) and from other coda consonants ([ɡaf̱ko], [ɡaḻko], etc.); (3) the ability to perceive [h] in [Ch] sequences may be related to the ability to perceive [h] in [hC] sequences. I discuss each of these findings in turn.

---

[5]Vowel laxing and harmony in the environment of coda /s/ weakening is commonly reported for Eastern Andalusian Spanish, but not for varieties of Western Andalusian Spanish (e.g., Henriksen 2017; Herrero de Haro 2016; see Herrero de Haro 2017 for overview).

First, Sevillians perceive the acoustic differences between stop-h sequences ([VthV]) and intervocalic stops ([VtV]). Ruch and Peters (2016) tested Sevillian listeners' ability to distinguish the minimal pair *pata-pasta* ('paw' - 'pasta'). Listeners heard words from two continua and chose the corresponding orthographic representation. The continua were built off two base words, [patha] ($\leftarrow$ /pasta/) and [pata] ($\leftarrow$ /pata/). For *pata*, [h] was spliced onto [t]. Then, the duration of [h] in both versions of [patha] was shortened in steps. Sevillian listeners—particularly younger ones—distinguished between *pata-pasta* based solely on the duration of [h] ([pat<u>h</u>a]). Ruch and Peters (2016) reason that if post-posed [h] is the primary cue to the distinction between [VthV] and [VtV], then the target [Ch] is, to a certain degree, phonologized.

While the results in Ruch and Peters (2016) are suggestive, they are not conclusive. The study relied on orthographic responses to monomorphemic words, but orthography cannot be assumed to match underlying representations. [Ch] could be represented as a cluster (/sC/ $\rightarrow$ [Ch]) or an aspirated stop (/C^h/ $\rightarrow$ [Ch]), and these options cannot be distinguished in monomorphemic words. Listeners could map [patha] to underlying /pasta/ or /pat^{h}a/, and yet either one to orthographic <pasta>, since orthography can diverge from phonological structure. My experiment tests representations using the word boundary context, because in this context phonological alternations provide clear evidence that /s/ is underlyingly present.

Second, Sevillian listeners perceive stop-h sequences more accurately than listeners whose dialects lack these sequences. In a word identification task, Bedinghaus (2015) played participants words like [pahtha] ($\leftarrow$ /pasta/) and [pata] ($\leftarrow$ /pata/) and asked them to choose the corresponding orthographic representation. For words like [pahtha] (with [h] on both sides of the stop), Sevillians chose *pasta* more often than listeners of other dialects (75% vs. 26% accuracy). Furthermore, non-Sevillian listeners often mistook [pahtha] forms for *pata* (55%), while Sevillian listeners rarely did (4.4% of the time) (Bedinghaus 2015: 190). Sevillian listeners perceived [h], while listeners of other dialects did not, or identified it as another consonant.

Third, although Sevillians perceive [h] in [Ch] most accurately, listeners of other dialects may also perceive it to an extent. Ruch and Harrington (2014) tested Argentinian listeners' ability to distinguish Sevillian /st/ vs. /t/. Their stimuli were items from two *pata-pasta* continua, created from a single base word /pasta/ produced as [pah$_1$th$_2$a]. In one continuum, a series of [h$_1$] steps were combined with a single-duration long [h$_2$]. In the second continuum, the same set of [h$_1$] steps were combined with a single-duration short [h$_2$]. Argentinian listeners responded <pasta> at higher rates when [h$_2$] was long, regardless of the duration of [h$_1$]. That their responses were conditioned by the duration of post-posed [h] is surprising, since their dialect does not have metathesis in <st> clusters.

Why were Ruch and Harrington's (2014) listeners able to make use of post-posed [h], while Bedinghaus' (2015) were not? The listener groups differ in one crucial way: Bedinghaus' (2015) listeners speak non-debuccalizing varieties that maintain coda /s/ as [s], while Ruch and Harrington's (2014) Argentinian listeners speak a variety that debuccalizes coda /s/ to [h]. Ruch and Harrington (2014) interpret Argentinians' sensitivity to the duration of post-posed [h] as indicating that it is perceptually parsed with pre-posed [h]. If native familiarity with coda [h] increases perception accuracy in [hC] sequences, and if [h$_1$] and [h$_2$] are perceptually linked in [<u>h$_1$</u>C<u>h$_2$</u>], then familiarity with debuccalization might help with perception in both [hC] *and* [Ch] sequences. Speakers of non-debuccalizing varieties—

like Bedinghaus' (2015) participants—are known to have low accuracy in perceiving coda (pre-posed) [h] (Schmidt 2013), which may also put them at a disadvantage in perceiving post-posed [h].

The perceptual findings discussed above motivate three listener groups for my perception experiment: Mexico ([sC], non-debuccalizing), Argentina ([hC], debuccalizing), and Seville ([Ch], debuccalizing and metathesizing).

3 Perception experiment

A binary forced-choice task tests the following question: do listeners map [h] in [Ch] sequences to an underlying preceding /s/? Participants heard short sentences consisting of a form of the verb *tener* ('to have') plus a nonce word. The subject of the sentence was removed (e.g., *** *tiene p(h)ali* '*** has *pali*'), and the duration of [h] on *p(h)ali* was manipulated. Participants were asked to choose the most likely subject of the sentence. The choices for test items were between 3SG and 2SG, whose corresponding verb forms differ only in the presence of final /s/ (/tjene/ vs. /tjene-s/). For listeners who attribute [h] in [p̪ali] to the preceding verb, subject choice should depend on the duration of [h]. To do the task successfully, listeners must:

- Perceive [h] in the [Ch] sequence;
- Connect [h] to the lexical representation of the verb, parsing it as a realization of underlying /s/ in the preceding word.

Listeners who do not perceive [h] cannot interpret it. Listeners who do perceive it must be able to connect it to a lexical representation. Both components are necessary, and changes in [h] duration should not affect responses for listeners lacking either.

3.1 Methods

3.1.1 Materials and procedure

The perception task takes advantage of the difference between Spanish 3SG and 2SG present tense verbs, which differ in the presence of final /s/ (6a vs. 6b).

(6) Spanish 3SG vs. 2SG verb forms

a. Juan /tjene-∅ pali/.	b. Tú /tjene-s pali/.
Juan has-3SG *pali.*	You have-2SG *pali.*
Juan has *pali.*	You have *pali.*

The stimuli sentences consist of combinations of subjects, forms of the verb *tener* ('to have'), and nine disyllabic nonce word nouns with stress on the first syllable. The words begin with all combinations of /p, t, k/ and /a, i, u/: *pali, pina, pumi; tali, tinu, tumi; kali, kina, kuma.* A male native speaker from Seville recorded the nonce words in full present tense paradigms, embedded in sentences like those in (6). Recordings were done in a soundbooth at New York University with a Zoom H4N Pro recorder and an Audio-Technica AT831b lapel

microphone. The speaker is linguistically trained and was instructed to produce metathesis where appropriate. He did so easily. Intensity was normalized to 60dB with a Praat script.

Listeners heard short phrases like (7): the subject is removed, the verb [tjene] has no acoustically present coda [s] or [h], and the duration of [h] *following* the first consonant of the nonce word is manipulated. If listeners choose 'Tú' (2SG, /s/), this indicates they have reconstructed an underlying preceding /s/ on the verb; if they choose 'Juan' (3SG, no /s/), it suggests that they have not. Crucially, listeners always hear [tjene], which is ambiguous for Sevillian listeners between being a realization of /tjene/ (2SG) or /tjenes/ (3SG). No stimuli have a verb with a phonetically-present coda [s] or [h] ([tjenes], [tjeneh]) that would clearly indicate a 2SG subject.

(7) Sample test sentence

  Hear: *** [ˈtjene ˈp̱ali]   **Answer choices:** Juan Tú

The test stimuli sentences were created as illustrated in Table 1, starting from two sentences: one with underlying /s/ and one without (Table 1a). The 3SG (no /s/) sentence is the base sentence (Table 1b). From this sentence, the subject was removed and replaced with a 30ms pure tone (beep), inserted immediately preceding the verb (Table 1c). Then, the nonce word from the 2SG (/s/) sentence was spliced into this modified version at the zero-crossing of the burst of the nonce word (*** ˈtjene | p̱ali$_2$) (Table 1d). This ensures that all other acoustic cues—like closure duration of the stop and the duration and quality of the vowel preceding the nonce word—remain the same as in the naturally-produced 3SG (no /s/) sentence. All test sentences consist of the verb from the 3SG (no /s/) recording and the nonce word from the 2SG (/s/) recording of the sentence pair.

Table 1: Steps in stimuli creation. [p] = Step-0; [p^h] = Step-1; [ph] = Step-2/Naturally long.

a.	Two original sentences	1. Juan tiene pali$_1$	[ˈhuan ˈtjene ˈpali]
		2. Tú tienes pali$_2$	[ˈtu ˈtjene ˈphali]
b.	(1) is base sentence	Juan tiene pali$_1$	[huan ˈtjene ˈpali]
c.	Subject in (1) replaced with beep	* tiene pali$_1$	[* ˈtjene ˈpali]
d.	Nonce word in (1) replaced with (2) (naturally long [h])	* tiene pali$_2$	[* ˈtjene ˈphali]
e.	3 [h] duration steps created		[* ˈtjene ˈphali]
			[* ˈtjene ˈp^{h}ali]
			[* ˈtjene ˈpali]

Next, an H-Step continuum was created by shortening [h] in nonce word-initial [ph, th, kh] (Table 1e). Recall that this [h] was naturally long because it was recorded a 2SG context. Continua were calculated independently for each nonce word sentence pair, as illustrated in Table 2. Step-2 is the longest, and was left as originally produced in 2SG contexts. Step-0 was shortened to approximately the duration of the normal intervocalic VOT.[6] Step-1 is

---

[6]Although metathesized [h] and VOT may not be phonetically identical, I assume that shortening

halfway between Step-2 and Step-0. The [h] durations reported for Step-0 are comparable to the intervocalic VOTs in natural 3sg (no /s/) contexts, and to those reported for other Spanish dialects (Rosner et al. 2000). The Step-0 tokens are not identical to the original intervocalic VOTs (from the 3sg no /s/ recordings), because splicing duration out of [h] required adjustments to hit zero-crossings. For the same reason, the [h] duration steps are roughly equal but not identical.

Table 2: H-Step continua for three nonce words

	VOT/H-duration in original recordings		Manipulated H-duration		
	Intervoc. VOT	/sC/: [Ch]			
UR	/huan tjene.../	/tu tjenes.../			
	[huan tjene p‿ali]	[tu tjene p‿h‿ali]	Step-0	Step-1	Step-2 (orig.)
/p‿ali/	11.0ms	66.7ms	9.41ms	38.9ms	66.7ms
/t‿ali/	17.2ms	82.7ms	16.1ms	49.4ms	82.7ms
/k‿ali/	22.7ms	96.6ms	22.7ms	60.1ms	96.6ms

The duration of post-posed [h] in my stimuli reflects the common cross-linguistic pattern for languages with aspirated stops, whereby velar stops have the longest VOT, followed by alveolars and bilabials (Cho and Ladefoged 1999). For aspirated stops, these patterns are likely due to articulatory factors like the speed of articulators, backness of articulation, and extent of contact between articulators (see Cho and Ladefoged 1999 for overview). That my stimuli—which have post-posed [h] as part of a sequence, *not* as VOT on a stop—also show this cline is notable.

Test sentences for one nonce word are illustrated in Table 3 (see Appendix A for the full list). For each of the 9 test words, there were 3 h-step versions (=27 sentences). These items were presented with the answer choices 'Juan' (3sg) and 'Tú' (2sg).

Table 3: Set of test items for *pali*. ([p] = Step-0; [pʰ] = Step-1; [ph] = Step-2)

Sentence	Listeners hear	Answer choices	
Juan tiene pali[Step-0]	[* ˈtjene ˈpali]	Juan	tú
Juan tiene pali[Step-1]	[* ˈtjene ˈpʰali]	Juan	tú
Juan tiene pali[Step-2]	[* ˈtjene ˈphali]	Juan	tú

There were also filler and control sentences, which were presented with no manipulations other than removing the subject. 18 control sentences consisted of each nonce word naturally produced in 3sg and 2sg contexts, with 2sg and 3sg answer choices. These unmanipulated recordings by the male Sevillian speaker thus contained all original acoustic cues. 50 filler sentences were designed to have answers that would be unambiguous for speakers of all

---

metathesized [h] to the same length as a short-lag Spanish VOT sounds similar enough to naturally-produced short-lag VOT. Indeed, the experiment shows that listeners accept shortened [h] as the VOT of intervocalic /p, t, k/.

varieties of Spanish. They consisted of 5 nonce words (*pali, pumi, kuma, kina, tinu*) with combinations of subject choices whose corresponding verb forms are very different. For example, one filler phrase was '*** [tenemo phali]' ('1PL have *pali*'), and the answer choices were *tú* (2SG) and *nosotros* (1PL). The verb forms corresponding to the answer choices, <tienes> (2SG) and <tenemos> (1PL), are different enough that recognizing the intended subject depends on much more than coda /s/. Controls and fillers distracted listeners from test sentences and made the task bearable for non-Sevillians. Listeners not familiar with [Ch] have difficulty perceiving [h], and unambiguous fillers and controls kept these listeners engaged. Controls and fillers were also used as exclusion criteria.

The experiment was run online with PCIbex (Zehr and Schwarz 2018). Participants saw one practice item with a verb form that unambiguously revealed the subject. The test, control, and filler items were randomized for each participant. There were a total of 95 items, and the task lasted 10-30 minutes. After the experiment, participants completed a demographic form. They were paid for their time.

3.1.2 Participants

Listeners of Sevillian (33), Mexican (30), and Argentinian (24) Spanish participated in the study; 8 were later excluded (see Section 3.1.4). Mexican and Argentinian listeners were recruited on Prolific (`https://www.prolific.co/`). Prolific was set to recruit Mexican listeners currently residing in Mexico. This was not possible for Argentinian listeners, because most Argentinians on Prolific reside outside Argentina. Sevillian participants were recruited through social networks and personal contacts. Nine Sevillians and nearly all of the Argentinians reported having spent a year or more in a different country.

Participants (except those excluded) in the three groups were of similar mean ages (Argentina = 27.3; Mexico = 25.5; Seville = 35.0), and were split by gender (Argentina = 9F/13M; Mexico = 7F/20M/1 no response; Seville = 19F/10M). The majority of participants in all regions had completed some post-secondary education. 15 Sevillian participants had also done graduate studies. Most participants reported knowledge of other languages, including English (intermediate to advanced proficiency), Catalan, Danish, Dutch, French, German, Italian, Portuguese, Russian and Valencian. Three Argentinian listeners reported native proficiency in other languages (1 = English and Danish; 1 = English; 1 = English and Portuguese), and four Mexican listeners reported native proficiency in English.

3.1.3 Predictions

The three different groups of participants test the two components necessary to successfully do the task: (1) perceive [h] in [Ch] sequences; (2) connect [h] to the lexical representation of the verb. Sevillian, Mexican, and Argentinian Spanish have different sets of relevant dialectal characteristics (Table 4). Shaded cells in each column are the same for those dialects.

Table 4: Properties of Argentinian, Mexican and Sevillian Spanish

	Debuccalizes?	2SG		3SG	
Mexico	No	/tu ˈtjenes/	[tu ˈtjenes ˈpali]	/xuan ˈtjene/	[huan ˈtjene ˈpali]
Argentina	Yes	/bos teˈnes/	[boh teˈneh ˈpali]		
Seville		/tu ˈtjenes/	[tu ˈtjene ˈphali]		

All dialects have the same underlying representation for 3SG verbs (Table 4). Argentinian Spanish, however, mostly uses a different 2SG pronoun (/bos/) and verb form than Sevillian and Mexican Spanish. The 2SG /bos/ verb form differs from 3SG present tense verbs by more than final /s/: it also differs by stem vowel and stress (3SG: [huan ˈtjene] vs. 2SG: [boh teˈneh]).[7] When Argentinian listeners hear [ˈtjene ˈphali], confusion could result from conflicting information. The stem vowel and verb stress are similar to Argentinians' 3SG form, while [h] may lead them towards 2SG.

Argentinian listeners were chosen despite this regional difference in 2SG forms because their dialect has extensive coda /s/ debuccalization to [h]. Other potentially accessible listener groups—like Dominican or Puerto Rican speakers—have coda /s/ *deletion*, but little debuccalization. My goal was to isolate debuccalization from metathesis by targeting a group that had one but not the other. Furthermore, even though Argentinian listeners do not *use* the /tu tjenes/ 2SG subject and verb forms, they have exposure to them through their own educational system, interactions with speakers from other regions, and media. Education and media are rather formal, however, and debuccalization may be less frequent in these contexts. The realizations of /tu tjenes/ that Argentinians hear may have lower rates of debuccalization than elsewhere in their dialect.

Under an analysis of [Ch] sequences as underlying /sC/ clusters, predictions for each listener group are schematized in Table 5. The table also shows what the prediction would be if [Ch] sequences were represented as aspirated stops. '≫' means strong preference, '>' means preference, and '∼' means no preference.

Table 5: Predicted response patterns (≫ indicates much greater; > indicates greater)

	Predicted under cluster analysis			Predicted under aspirated stop analysis		
	Step-0	Step-1	Step-2	Step-0	Step-1	Step-2
Seville	3SG ≫ 2SG	3SG ≈ 2SG	3SG ≪ 2SG	3SG ≫ 2SG	3SG ≫ 2SG	3SG ≫ 2SG
Mexico	3SG ≫ 2SG	3SG ≫ 2SG	3SG ≫ 2SG			
Argentina	3SG > 2SG	3SG > 2SG	3SG > 2SG			

Prediction 1: Sevillian listeners should base their responses on the duration of [h] in [Ch]. When [h] is shortest, they should have a strong preference for 3SG (no /s/) ≫ 2SG (/s/). When [h] is longest, they should have a strong preference for 2SG (/s/) ≫ 3SG (no /s/). At the middle [h] step, they should have no preference.

---

[7]Diphthongs occur only in stressed syllables, which is why /tu/ and /bos/ forms differ in both stress and quality of the first vowel (/ˈtu ˈtjenes/ vs. /ˈbos teˈnes/).

Prediction 2: Mexican listeners should not base their choice of subject on the duration of [h] in [Ch]. I expect them to have a stong preference for 3SG (no /s/) ≫ 2SG (/s/) responses at all durations of [h]. Although their dialect shares with Sevillian the lexical representation of the verb (/ˈtjene-s/), Mexican Spanish does not debuccalize coda /s/. Using [h] to choose the subject requires recognizing metathesis and [h] as an allophone of /s/, which—as established by previous studies—is difficult for listeners of non-debuccalizing varieties (see Section 2.3).

Prediction 3: Argentinian listeners' choice of subject may depend somewhat on [h] duration, but I expect them to have a slight preference for 3SG (no /s/) > 2SG (/s/) at all [h] durations. Native familiarity with debuccalization may help them perceive [h] in both [hC] and [Ch] sequences (Section 2.3). However, they may have difficulty connecting [h] to lexical representations, since their most common 2SG lexical representation (/bos teˈnes/) differs from that used in the experiment. If they do respond to [h] duration, the effect should be smaller than for Sevillians.

If Sevillian listeners' choice of subject does not vary by H-duration, that would constitute evidence for the aspirated stop analysis. This result would indicate that they do not use [h] to draw conclusions about the segmental content of the preceding word, and that they do not attribute [h] to its phonological context. If they do not attribute post-posed [h] to the surrounding phonological context, then they must consider it part of the stop itself.[8]

3.1.4   Statistical analysis

The data were analyzed in logistic mixed-effects regressions in R (RCoreTeam 2020), fit with the BOBYQA optimizer. The *lmerTest* package was used to calculate *p*-values (Kuznetsova, Brockhoff, and Christensen 2017). The models were run only on test items, and predict the response (2SG (coded as 1) vs. 3SG (coded as 0)) from H-Duration Step (Step-0, Step-1, Step-2), Region (Seville, Mexico, Argentina), and their interaction. Step-0 (shortest [h] duration) and Argentina are the baselines. A main model containing data from all regions had significant interactions involving Region, so separate models were then run to investigate the effect of H-Duration Step within each region. I present only model results by region, which are easier to interpret (results from the main model are in Appendix B). All models also contained random intercepts of participant and item. Further random intercepts and slopes resulted in model fit errors and were omitted. Post-hoc comparisons were run with *emmeans* (with Tukey adjustment for multiple comparisons) (Lenth 2020).

Eight participants were excluded. Four Sevillian participants were excluded for accuracy below 75% on the control items (50%-66% accuracy). Sevillians should perform well on these natural recordings of their native dialect. Two Mexican and two Argentinian listeners were excluded for accuracy lower than 75% on fillers. Fillers were used for these listeners because the only cues to the subject in control items were [h] and possibly gemination, and Mexican

---

[8]Note that the expected pattern of results under the aspirated stop analysis for Seville is the same as for Mexico under a cluster analysis. In that case, how would we know that the representation for Sevillian Spanish differs from that of Mexican Spanish? My results suggest that Mexican listeners essentially treat post-posed [h] as linguistically irrelevant. For Sevillian listeners, in contrast, post-posed [h] has been found to be one of the strongest cues to the /C/-/sC/ contrast (Ruch and Peters 2016), so it would be unlikely for them not to perceive it. If they perceive it but do not attribute it to phonological context, that suggests that they may consider it part of the stop.

and Argentinian listeners lack familiarity with these cues. Fillers had unambiguous answers.

3.2   Results

Figure 1 shows the results for each listener group. The x-axis represents h-duration step, with 0 being the shortest and 2 being the longest. The y-axis represents the proportion of 3SG responses (black line) and 2SG responses (gray dashes), at each h-duration step. The 2SG and 3SG lines are each others' inverse, since these were the only two answer options (at each h-duration step, the proportion of 2SG and 3SG responses sums to 1). Recall that listeners responded to stimuli like [*** tjene p̱ali], with the duration of [h] manipulated and no evidence of coda /s/ on the verb [tjene].
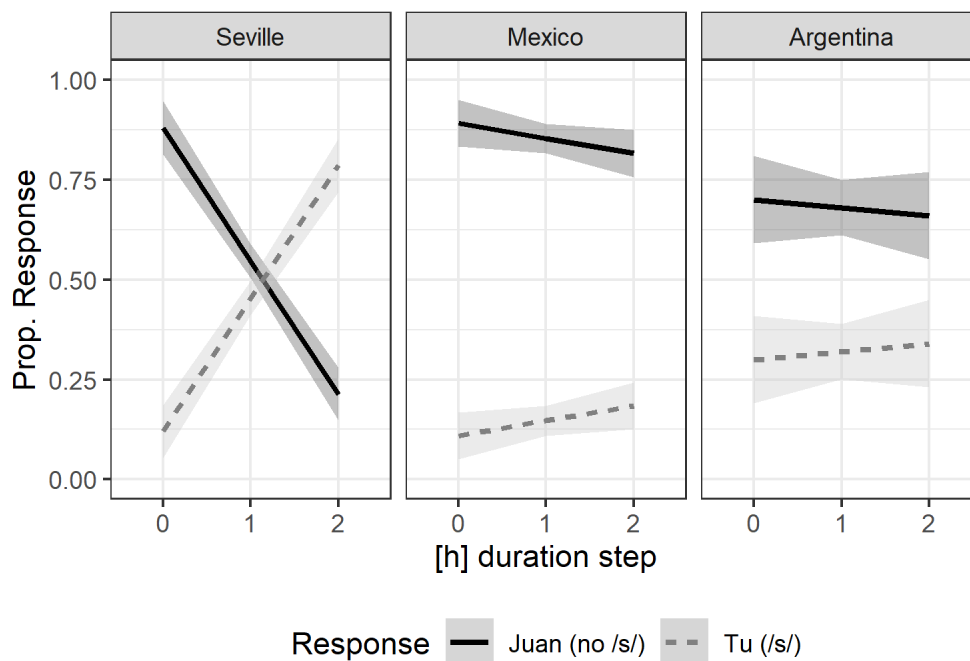

Figure 1: Response rate of 2SG and 3SG for listener groups at each h-duration step. Response rates are plotted for each individual participant. Shading represents 95% confidence intervals.

When Sevillian listeners hear a stimulus like [*** tjene p̱ali], their responses depend on the duration of [h] (Figure 1, left panel). At Step-0 (shortest [h]), Sevillian listeners choose mostly 3SG (no /s/) subjects (89%) and rarely 2SG (/s/) subjects (11%). At Step-1, they are equally likely to respond with 3SG (no /s/) and 2SG (/s/) responses. At Step-2 ([h] is longest), their responses are a mirror image of Step-0, with 78% 2SG (/s/) responses and 22% 3SG (no /s/) responses.[9] Sevillian listeners use the duration of [h] in [Ch] as information

---

[9]There are several reasons that 2SG responses may not be at 100% for the longest [h] step. One, overall accuracy for identifying [h] in [hC] and [Ch] sequences hovers around around 75%-80% in other studies too, even for listeners whose native dialects have debuccalization (Bedinghaus 2015). A more interesting possibility is that the stimuli provide conflicting cues, since the verb comes from a 3SG sentences and the nonce word comes from a 2SG sentence. As reported for some other varieties spoken in Southern Spain

about the preceding word—specifically, about the morphological distinction between 2sG (/s/) and 3sG (no /s/) verb forms.

Mexican listeners' responses (Figure 1, middle panel) differ drastically from Sevillian listeners'. When they hear [*** tjene p̲h̲ali], they respond 3sG (no /s/) at rates higher than 75% regardless of the duration of [h], while 2sG (/s/) responses range from 10%-18%, even at Step-2 where [h] is the longest. There is a slight increase in 2sG (/s/) responses at Step-1 and at Step-2, but these responses are still well below chance. Mexican listeners do not consistently parse long [h] in [Ch] as morphological information about the preceding verb.

Argentinian listeners (Figure 1, right panel) give 2sG (/s/) responses at rates between 30%-34% when they hear [*** tjene p̲h̲ali], regardless of the duration of [h]. These listeners were predicted to fall between Sevillian and Mexican listeners, because their variety has debuccalization ([hC]), lacks metathesis (*[Ch]), and has a different lexical representation of 2sG verbs. Indeed, Argentinian listeners do fall between Mexican and Sevillian listeners. Their results reflect the predictions in Table 5, in that they have a preference for 3sG > 2sG at all h-duration steps, and in that this preference is not as strong as Mexican listeners' preference. However, their results are also different than expected. They are *not* a less extreme version of Seville, with rates of 2sG responses increasing as h-duration increases. Instead, their responses are qualitatively like Mexican listeners', with no meaningful increase in 2sG responses as h-step duration increases, but their 2sG response rate is closer to chance.

The statistical modeling confirms these visual interpretations. The main model including all three regions (Appendix B, Table 14) has a significant interaction between Region and H-Duration Step. Models by region confirm that the effect of H-Duration Step differs by region. Step-0 is the baseline for H-Step Duration, and positive effects indicate higher likelihood of a 2sG response.

For Sevillian listeners (model in Table 6), the effect of H-Duration Step is significant: Sevillians give more 2sG (/s/) responses at Step-1 than at Step-0 ($\beta = 2.44$, $p < .0001$), and more at Step-2 than at Step-1 (*emmeans* post-hoc test: $\beta = $ -1.84, $p < .001$).[10] For Mexican listeners (Table 7), H-Duration Step is also significant, although the estimates are much smaller than for Seville. Mexican listeners give more 2sG (/s/) responses at Step-2 than at Step-0 ($\beta = .76$, $p < .05$). The difference between Step-1 and Step-0 is not significant, nor is the difference between Step-1 and Step-2 (*emmeans* post-hoc tests). In the Argentina-only model (not shown), H-Duration Step is not a significant predictor of 2sG vs. 3sG response, nor are there trends. Argentinians' responses do not depend on the duration of metathesized [h].

(Henriksen 2017), my 2sG verbs showed vowel laxing in the wake of /s/ weakening (/tjenes/ → [tjenɛ]) compared to 3sG verbs. Inserting nonce words with long [h] after 3sG verbs leads to a conflict: the verb-final vowel indicates 3sG, but the long [h] indicates 2sG. This conflict may have lowered the rate of 2sG responses, but the high rate of 2sG responses (78%) suggests that [h] duration carries more weight than preceding vowel quality. Thanks to Aaron Kaplan for bringing this to my attention.

[10]The directionality of the *emmeans* effects is reversed from the effects in the models, where positive estimates indicate higher likelihood of a 2sG response. This is because the baseline level of the comparison is different. The models predict the likelihood of a 2sG response at Step-1 and Step-2 as compared to the baseline Step-0. The positive effects indicate more 2sG responses in those steps as compared to Step-0. In contrast, *emmeans* compares the likelihood of a 2sG response at Step-0 vs. Step-1 and Step-2, and Step-1 vs. Step-2. The baselines are the shorter h-duration steps; negative effects indicate fewer 2sG responses at the shorter steps as compared to the longer ones.

Table 6: Model of H-Duration results for Seville

	Est.	SE	z value	Pr(> |z|)	
(Int.)	-2.602	0.426	-6.115	0.000	***
Step-1	2.444	0.509	4.800	0.000	***
Step-2	4.286	0.534	8.034	0.000	***

Table 7: Model of H-Duration results for Mexico

Est.	SE	z value	Pr(> |z|)	
-2.880	0.396	-7.269	0.000	***
0.586	0.316	1.853	0.064	
0.757	0.313	2.419	0.016	*

In short, Sevillian listeners' 2SG responses increase monotonically as h-duration increases: Step-0 < Step-1 < Step-2. Mexican listeners' 2SG responses increase slightly from Step-0 to Step-2 (Step-0 < Step-2), and there is no hierarchy to Argentinian listeners' 2SG responses.[11]

3.3 Distribution of responses

Neither Mexican nor Argentinian listeners perform like Sevillian listeners. But although their patterns are qualitatively similar, the distribution of responses differs. The large confidence intervals in the Argentinian data (Figure 1, right panel, CIs represented by gray shading) could hide distinct individual patterns, or could reflect uncertainty at the level of the individual.

The density plots in Figure 2 show the distribution of individuals' proportions of 2SG (/s/) responses at each H-Duration step. Sevillian and Mexican listeners have peaked distributions. The Sevillian listeners' sharp 2SG response peak moves neatly by H-Duration Step. Mexican listeners' sharp 2SG response peak is around 0% at all H-Duration Steps. At the individual level, participants in both Seville and Mexico are certain about their answers. At the group level, participants agree, resulting in peaked response distributions. In contrast, individual Argentinian listeners are closer to 50% (more uncertain), and there are more differences within the group. Mexican listeners are certain of a 3SG (no /s/) response at all [h] duration steps, and Argentinian listeners are less certain, regardless of [h] duration.

---

[11]Readers may wonder if place of articulation affected listeners' responses. The patterns are very similar across places of articulation: 2SG responses increased as [h] duration increased. In Seville, listeners had numerically higher 2SG response rates for /st/, followed by /sk/ and /sp/, at all H-Duration steps. Rates of 2SG responses were also higher for /st/ than /sk, sp/ in Mexico and Argentina. However, when place of articulation is included in the main model (three-way interaction between Region*H-Step*POA), the only statistically-significant effect is for /st/ in Seville at Step-2. That is, there is a disproportionate boost in 2SG responses at Step-2 for /st/ in Seville. This tendency to interpret long [h] in [th] as 2SG cannot just be due to [h] duration: in the stimuli, [h] is longer on [k] than on [t] (see Table 2). Instead, this tendency follows if the change started in /st/ sequences (as proposed in e.g., Ruch and Peters 2016), and listeners are most used to the metathesized realization of these sequences. Excluding place of articulation from the main model did not affect direction or significance of other effects, so I omit it to focus on the effect of [h] duration, which holds across places of articulation.

Figure 2: Density plots of 2sg (/s/) responses by H-Duration Step
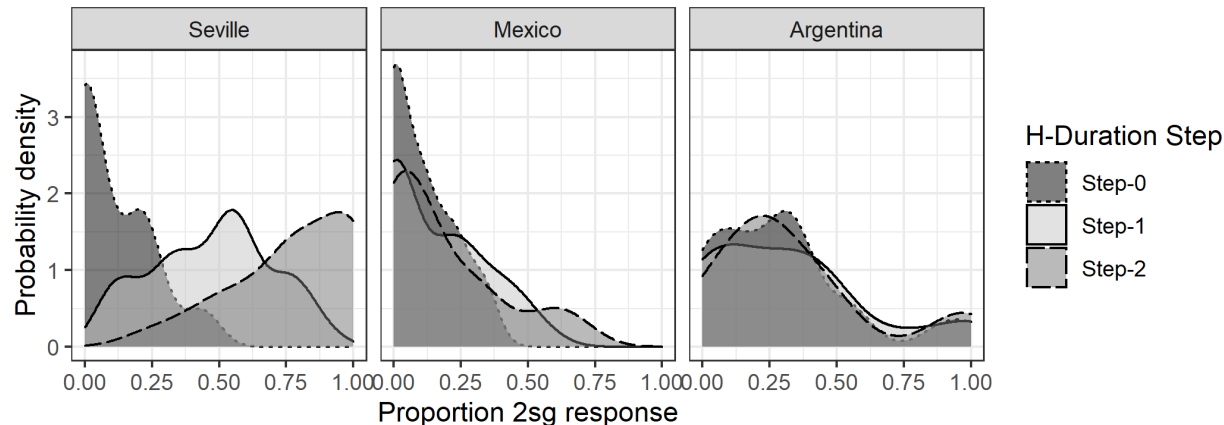
3.4 Results discussion

The results are largely in line with the predictions. Sevillian listeners perceive [h] in [Ch] and attribute it to the preceding word, interpreting it as a morphological distinction. This suggests that surface [Ch] sequences consist of an underlying preceding /s/ followed by a stop. Mexican listeners do not consistently parse [h] in [Ch] sequences as underlying preceding /s/, regardless of the duration of [h]. Argentinian listeners respond even less to [h] duration than Mexican listeners.

Contrary to predictions, however, the Mexico data does show a slight effect of H-Duration step. This effect is driven by four Mexican listeners who had higher rates of 2sg (/s/) responses than 3sg (no/s/) responses at the longest [h] duration step. These four listeners responded 2sg at a rate of 55%-66% at Step-2. For comparison, 20 Sevillian listeners chose 2sg (/s/) at a rate of 75% or higher at Step-2. None of the Mexican participants with this effect reported contact with Southern Spain, or having lived outside Mexico. One possible explanation is that, for some reason, they are able to perceive [h] in [Ch] sequences and map it to an underlying /sC/ cluster. Another possibility is that these listeners perceive [Ch] as non-canonical, and choose 2sg because long [h] is *not* acceptable following 3sg verbs. A final possibility is that these listeners have some contact with Sevillian Spanish (or a similar variety) through sources not captured in the demographic questionnaire.

The lack of effect for Argentinian listeners suggests that familiarity with debuccalization ([hC]) does not straightforwardly facilitate perception of metathesis, *contra* Ruch and Harrington (2014). There are several possible reasons why my results may differ from theirs. Most importantly, my study uses verb forms while theirs uses monomorphemic nouns. Recall that Argentinian listeners' native dialect has different 2sg forms than those used in my stimuli (see Section 3.1.3 and Table 4). My stimuli had subtle phonetic differences between forms (['tjene 'pʰali]-['tjene 'pʰali]), and all of these forms differ substantially from the forms that Argentinian listeners are most used to, in terms of the stress and first vowel of the verb ([te'neh 'pali]). Even if Argentinian listeners' native dialect experience increases their ability to perceive [h] in some contexts, their lexical representations of the 2sg form /tjenes/ may

not be robust enough to allow them to connect this representation to the phonetic forms heard in the experiment. Both perception and lexical representations are necessary. Ruch and Harrington's (2014) study would not have presented this same ambiguity for Argentinian listeners because the words were monomorphemic nouns, clearly connected to two distinct lexical items through orthographically distinct answer choices. A reviewer also wonders if the difference in [h] duration between the two studies may have affected Argentinian listeners' accuracy, since [h] may be shorter across word boundaries (/s#t/) than within words (/st#/). However, this does not explain my Argentinian listeners' difficulty. In my /s#t/ stimuli, the duration of [h] ranges from 16ms-82ms sequences, the upper end of which is substantially *longer* than [h] in Ruch and Harrington (2014)'s monomorphemic /st/ word (29ms).

4 Analysis

My experimental results show that the three listener groups differ in how they perceive [Ch] sequences. In addition, they speak dialects that are known to differ in the production of /sC/ sequences, suggesting that their phonological grammars have different constraint rankings. I suggest that the differences in perception may be linked to differences in their production grammars. Listeners hearing forms not frequent (or present) in their own dialects may have difficulty mapping them to an underlying representation, accounting for why my Mexican and Argentinian participants did not perceive [Ch] as /sC/.

Table 8 summarizes listeners' behavior in the perception experiment, alongside the typical realizations of /sC/ for each dialect in production.[12]

Table 8: Mappings in the perception experiment and in production

Dialect	Perception Experiment		Production	
a. Mexico		→ /ˈtjene pali/		→ [ˈtjenes pali]
b. Argentina	[ˈtjene phali]	→ /ˈtjene pali/	/ˈtjenes pali/	→ [ˈtjeneh pali]
c. Seville		→ /ˈtjenes pali/		→ [ˈtjene phali]

In this section I offer a basic analysis, framed in Optimality Theory (Prince and Smolensky 1993), of three outcomes of /sC/ clusters: retention ([sC]), debuccalization ([hC]), and metathesis ([Ch]). These outcomes are the most common in Mexican, Argentinian, and Sevillian Spanish, respectively. Although all of the varieties also have synchronic variation in coda /s/ production, I focus just on the most common outcome in each variety.

4.1 Analysis of /s/ retention, debuccalization, and metathesis

I frame the analysis in featural terms, following other formal approaches to Spanish coda /s/ reduction (e.g., Goldsmith 1981; Morris 2000; Lloret and Martínez-Paricio 2020): /s/ is

---

[12]Recall that Argentinian Spanish typically uses a different form for 2sg (Table 4). However, when Argentinian Spanish speakers do produce the 2sg form of the other dialects, it is with debuccalization, as in Table 8.

composed of both place and laryngeal features, which can be manipulated independently by the grammar.

I loosely follow McCarthy (2008b) in using the constraints in (8)-(10) to drive coda /s/ reduction.[13]

(8) MAX: Assign a violation for a segment present in the input that has no correspondent in the output.

(9) MAX[PLACE]: Assign a violation for a place feature present in the input that has no correspondent in the output.

(10) CODACOND: Assign a violation for place features not linked to an onset (Itô 1989).

Different outcomes are obtained by the re-rankings schematized in (11)-(12) and are illustrated in the tableaux below.

(11) Retention: MAX, MAX[PLACE] » CODACOND

(12) Debuccalization: MAX, CODACOND » MAX[PLACE]

Coda /s/ retention, which is typical of Mexican Spanish, is obtained when faithfulness constraints protect the segment and its place features at the expense of violating markedness constraints (tableau in (13)). MAX[PLACE] prefers candidate (a), where /s/ maintains its supralaryngeal place feature ([s]), over candidate (b), where debuccalization results in placeless [h]. Candidate (a) wins, even though it violates CODACOND by not having the place features of [s] linked on an onset. A general, high-ranked MAX constraint prevents deletion of the entire /s/ segment (c).

(13) Retention of coda /s/ as [s] (MAX[PLACE] ≫ CODACOND)

	/tjenes pali/	MAX	MAX[PLACE]	CODACOND
☞a.	tjenes pali			*
b.	tjeneh pali		*!	
c.	tjene pali	*!	*	

For debuccalization of /s/ to [h]—the most common outcome in Argentinian Spanish—the ranking must be the opposite: the markedness constraint dispreferring place features in coda

---

[13]My constraints limit codas by restricting place features in coda position, and I assume that these features can be delinked independently of other features. Other approaches force coda reduction with constraints that limit the feature [continuous] (e.g., Morris 2000), force delinking of the entire supralaryngeal node (place and manner) (e.g., Hualde 1989b), or force delinking of both place and manner separately (e.g., Lloret and Martínez-Paricio 2020). My basic analysis could be framed in any of these ways. A closer analysis of some phenomena I do not analyze—total assimilation and gemination of an /sC/ sequences to [C:], debuccalization and metathesis of coda /p, t, k/, and coalescence in /sb, sd, sg/ sequences (e.g., /desde/ → [deθːe])—may favor one approach or another.

position must outrank the faithfulness constraint protecting place features. In the tableau in (14), the ranking CODACOND ≫ MAX[PLACE] prefers candidate (b) (debuccalization) over (a) (retention), because coda [h] in (b) has no supralaryngeal place features. As in the previous tableau, MAX prevents complete deletion (c).

(14)   Debuccalization of coda /s/ to [h] (CODACOND ≫ MAX[PLACE])

	/tjenes pali/	MAX	CODACOND	MAX[PLACE]
a.	tjenes pali		*!	
☞b.	tjeneh pali			*
c.	tjene pali	*!		*

MAX and MAX[PLACE] cannot be ranked in relation to each other based on this tableau because they are in a stringency relationship. Between candidates (c) and (b), though, MAX breaks the tie. While these constraints are sufficient for debuccalization in Argentinian, I note that MAX will also need to outrank another constraint, *CODA-LARYNGEAL, in order for the system to prefer debuccalization over deletion. This constraint is necessary for the analysis of metathesis, presented next.

For metathesis, two more constraints are needed. LINEARITY (15) is violated by segmental orders that differ between the input and output, and *CODA-LARYNGEAL (16) is a positional markedness constraint that is violated by a laryngeal feature in coda position. This constraint is needed to differentiate candidates with [h] in coda position from candidates where [h] has metathesized, since both violate MAX[PLACE] equally.

(15)   LINEARITY: Assign a violation for two elements, where 'the output order [does not] preserve the order of input elements' (paraphrased from McCarthy 2008a: 198; see this work and McCarthy and Prince 1995: 123 for a more formal definition).

(16)   *CODA-LARYNGEAL: Assign a violation for a laryngeal feature in coda position.

The *CODA-LARYNGEAL constraint requires further explanation. I treat Spanish /s/, realized as [s] or [h], as having the laryngeal feature [spread glottis]. This differs from the traditional assumption for Spanish, which is that fricatives are not specified with [spread glottis] because (a) they do not contrast for voicing, and (b) because the voicing contrast in stops argued to be one of a [voice]—not [spread glottis]—feature (Beckman and Ringen 2009). However, some previous work on Spanish also assumes that /s/ is specified as [spread glottis] (Widdison 1995; Vaux 1998, Morris 2000; Gerfen 2002:fn.11). For Sevillian Spanish, several findings support the [spread glottis] specification of /s/: /s/ causes devoicing of surrounding segments (/desde/ → [deθ̪ːe] 'since') (Hualde 1989a; Martinez-Gil 2012; Hualde 2014: 160–1) and resists the voicing assimilation that is common in other dialects, such as Peninsular Spanish (/desde/ → [dez̪ð̪e]; Martinez-Gil 2012; Hualde 2014; Campos-Astorkiza 2016). While other dialects differ in the details of how /s/ affects surrounding segments, I treat them all as specifying /s/ with [spread glottis] for the sake of simplicity.[14]

---

[14]The specific formulation of the constraints is affected by this assumption, but the structure of the analysis does not hinge on it.

Cross-linguistically, restrictions against laryngeal features in coda position are widespread. Patterns reflecting these restrictions include final devoicing in languages like German and Dutch, neutralization of syllable-final laryngeal features in Thai and Korean, and neutralization of *specific* laryngeal features syllable- or word-finally (e.g., the aspiration contrast is lost word-finally in Bengali, while the voicing contrast is maintained) (Lombardi 1991:Ch.3). From a licensing perspective, Lombardi (1991; 2001) argues that laryngeal features are more frequently licensed in onset position (pre-sonorant) than in coda position (preconsonantal).

More specifically, why would the feature [spread glottis] be dispreferred in coda position? While explanations vary, the typological patterns are clear and positional licensing constraints on the feature are common. Cross-linguistically, coda /h/ is much rarer than onset /h/ (Holmberg and Gibson 1979). Davis and Cho (2003) argue that [spread glottis] is only licensed in prominent positions. Although their argument is based on English and Korean, the cross-linguistic asymmetry in which onset /h/ is more common than coda /h/ suggests that it can be extended to other languages. In Spanish, codas are not salient. The avoidance of [spread glottis] and the neutralization of the voicing contrast in coda position that occur in Spanish are both typical of non-salient positions.

The dispreference for [spread glottis] in coda position is also reflected in the cross-linguistic asymmetry between preaspiration and postaspiration. In languages that have stops specified with [spread glottis], postaspiration ([Ch]) is clearly the default; preaspiration ([hC]) is rarer (Silverman 2003). Casserly (2012) uses this asymmetry to motivate a constraint preferring the [spread glottis] gesture to align with the right edge of a stop, resulting in postaspiration. Similarly, Canfield (2015) and Takahashi (2019) use a *GLOTTAL-PLOSIVE constraint to disprefer [h]-stop sequences. Kingston (1990) argues that aspiration is more likely to 'bind' to a stop's release than its onset, for both articulatory and perceptual reasons.

For Sevillian metathesis, there are three crucial rankings. The first, *CODA-LARYNGEAL, CODACOND ≫ MAX[PLACE], prefers candidates where there is no laryngeal feature in coda position, or where [s] has reduced to [h]. The second crucial ranking, *CODA-LARYNGEAL ≫ LINEARITY, prefers metathesizing [h] over leaving it in coda position. The third, MAX ≫ LINEARITY, prefers metathesis over deleting /s/ altogether.

The tableau in (17) derives Sevillian metathesis. Candidates (a) (retention) and (d) (deletion) are ruled out by *CODA-LARNGEAL, CODACOND and MAX, respectively. The crucial comparison is between candidates (b) and (c). In candidate (b), coda /s/ has debuccalized to [h], but it is still in coda position. In candidate (c), [h] has metathesized with the following stop. Candidate (c) removes the laryngeal feature [spread glottis] from coda position and wins, even though metathesis is unfaithful to the order of segments in the input (violating LINEARITY). Metathesis is preferred over leaving material in coda position, or deleting this material altogether.

(17)   Metathesis: /sC/ → [Ch] (*CODA-LARYNGEAL, CODACOND ≫ MAX[PLACE]; *CODA-LARYNGEAL ≫ LINEARITY; MAX ≫ LINEARITY)

	/tjenes pali/	Max	*Coda-Lar	CodaCond	Max[place]	Linearity
a.	tjenes pali		*	*!		
b.	tjeneh pali		*!		*	
☞c.	tjene phali				*	*
d.	tjene pali	*!			*	

Another possible candidate, [tjene psali], has metathesis without /s/ debuccalization. Consideration of this candidate highlights other crucial properties of metathesis, which I discuss further in Section 4.2.2.

4.2 Further complexities: Metathesis and markedness

Analyzing the Sevillian process of /sC/ → [Ch] as metathesis raises several possible objections. Here, I address the goodness of [Ch] as an onset (Section 4.2.1), the connection between metathesis and debuccalization (Section 4.2.2), and why metathesis may be limited to /s/-voiceless stop sequences (Section 4.2.3).

4.2.1 Stop-h as an onset sequence

One possible objection to analyzing [Ch] as an onset sequence is that these sequences are not obviously phonotactically legal in Spanish. However, Spanish dialects are phonotactically different, and [Ch] sequences may be less marked than coda obstruents in Sevillian. Spanish generally disprefers word-medial coda obstruents. In the lexicon, some obstruents are frequent word-medially (e.g., /s/), but are often reduced or deleted in this context. Other obstruents are much less frequent (or non-existent) as word-medial codas (e.g., /p, t, k, b, d, ɡ/) (Gerfen 2002). Where they do occur, they frequently reduce and the voicing contrast is lost (/apto/ → [apto ∼ ab̥to ∼ abto ∼ aβto]) (Campos-Astorkiza 2012: 95; see also Colina 2012). This low tolerance for coda obstruents is widespread in Romance languages (Vennemann 1988).

Varieties of Andalusian Spanish have particularly low tolerance for coda obstruents and are considered 'innovative' or 'radical' (Alvar 1955; Villena-Ponsoda 2008; Hualde 2014: 286). Only coda sonorants surface faithfully; coda obstruents are drastically reduced (Gerfen 2001: 194; Campos-Astorkiza 2003: 1–2). Coda /s/ almost always reduces (Ruch 2008). Coda stops (voiced and voiceless) undergo more extreme reduction than in other Spanish varieties. In addition to the reduction common in other varieties, Andalusian coda voiceless stops can debuccalize to [h] (e.g., /apto/ → [ahːo] 'apt') (Gerfen 2001; Campos-Astorkiza 2003) and even undergo metathesis (e.g., [atːho]) (Del Saz 2019; Gilbert 2022).

Thus, although onset [Ch] sequences may seem odd in Spanish, coda obstruents are so marked in Andalusian Spanish that [Ch] sequences may be the better option. Recall from Section 2.2 that Sevillian metathesis has also been argued to be driven by this intolerance for closed syllables.

One option to avoid onset [Ch] sequences would be to syllabify them heterosyllabically, as [VC.h]. This heterosyllabic parse is bad for reasons akin to those that arise in unmetathesized [Vh.C] forms. The heterosyllabic [VC.h] parse avoids the onset sequence, but leaves an

obstruent in coda position. If metathesis is motivated by a preference for open syllables ([Vh.C] →[V.Ch]), then the metathesized form syllabified as [VC.h] is just as bad as the unmetathesized form [Vh.C]. The heterosyllabic parse not only fails to resolve the coda issue, but can make it worse. If coda restrictions are based on the coda's ability to license place features, then a coda obstruent is worse than a placeless coda [h]. Furthermore, the heterosyllabic parse would create a marked sonority profile. Cross-linguistically, there is a strong preference for sonority to fall across a syllable boundary (Syllable Contact Law; Murray and Vennemann 1983; Clements 1990; Gouskova 2004). The [VC.h] syllabification would create a sonority *rise* from a voiceless stop to [h] across the boundary. Finally, if the parse were heterosyllabic, the coda stop might be expected to debuccalize like other coda obstruents in this variety. It does not (*[Vh.tV] → [Vt.hV] → [Vh.hV]).

Steriade's (1994: 213–5) proposal for complex Mazatec onsets suggests a possible solution to the unexpected [Ch] onsets in Sevillian. She proposes that the surface representation of [Ch] sequences is the same, regardless of whether the underlying representation is an underlying aspirated stop or a cluster. She also argues that the more similar sequences of sounds are to single segments, the less marked they are. Thus, a surface [Ch] sequence is more marked than an unaspirated stop (because it has an extra feature), but less marked than a cluster like [pr], which is not structurally similar to a single segment. In the case of Sevillian, a [Ch] sequence may not be as bad as other onset clusters for this reason. Recall that [ph, th, kh] can be analyzed as onset clusters in other languages as well (e.g., Acehnese, Old Khmer, and Taba; Section 2.1).

4.2.2 Metathesis and debuccalization

In a parallel OT analysis, another possible candidate is one with metathesis but no /s/ debuccalization, [tjene psali] (← /tjenes pali/). My analysis does not distinguish between this candidate and the winner with both debuccalization and metathesis, [tjene phali]. The analysis could be augmented to prefer [Ch] over [Cs] in several ways. One possibility would be to add a markedness constraint preferring [Ch] over [Cs] onsets because [Ch] has a larger sonority rise. Cross-linguistically, larger sonority rises are preferred in onset sequences (Sonority Sequencing Principle; Clements 1990; Blevins 1996). Or, [Ch] might be better than [Cs] because [Ch] is more segment-like (as in Steriade's 1994 proposal; Section 4.2.1). However, deciding between these two candidates based on onset phonotactics misses the substantive point: metathesis is a step in the /s/ reduction continuum, which is about constraints on coda consonants. In a parallel analysis, there is no clear reason why [Ch] should be preferred over [Cs].

However, there are clear reasons to think that debuccalization precedes metathesis. The steps leading to metathesis in Sevillian are visible in cross-dialectal synchronic variation, as well as the diachronic evolution of /s/-stop sequences in Sevillian. Synchronically, coda /s/ is realized variably as [s], [h], or elided in many varieties of Spanish. Diachronically, coda /s/ weakening has been a trend in the development of Romance languages (see Mason 1994 for overview), Spanish varieties (Alonso 1945; Lapesa 1981; Mason 1994), and in Southern Spain specifically (Alvar 1955). Debuccalization preceded the metathesis change in Seville. I consider Sevillian metathesis to result from the same mechanism that has been at work for centuries, in both Spanish and cross-linguistically (Moya Corral and Tejada Giráldez 2020).

The [Cs] candidate is not included in the Seville analysis (tableau in (17)) because I assume that debuccalization is a prerequisite for metathesis. A full analysis will need to capture the fact that coda [s] reduces to [h] before metathesis occurs, and may also be one of the factors that *allows* metathesis to occur. Phonetically, [h] may have 'stretched out' features realized across multiple segments, a situation argued to allow listeners to interpret it in a different position (Blevins and Garrett 2004). Articulatorily, the loss of the supralaryngeal tongue tip gesture in [s] → [h] may permit a looser coupling between the laryngeal gesture and the stop closure gesture, setting the stage for more variable gestural timing and, thus, metathesis.

In short, the connection between debuccalization and metathesis is not captured in a parallel analysis. A complete analysis of Sevillian metathesis will likely be serial (McCarthy and Pater 2016), and involve additional complexities related to gestural timing, features, and variation.

4.2.3   Limiting metathesis to /s/-voiceless stop sequences

Full metathesis in Sevillian occurs only with /s/-voiceless stop sequences. When /s/ precedes a vowel, /s/ often debuccalizes or deletes (Villena-Ponsoda 2008: 149). When /s/ precedes a voiced stop or sonorant, some gestural realignment may occur, but it does not appear to be full metathesis. Voiced stops /b, d, ɡ/ are realized as spirantized continuants [β, ð, ɣ] in most environments in Spanish, including intervocalically and following /s/. In Andalusian Spanish, /sb, sd, sɡ/ sequences have been reported to be devoiced (/desde/ → [deθːe] 'since') and more constricted than intervocalic /b, d, ɡ/ (Hualde 1989a; Romero 1995; Martinez-Gil 2012). /s/-sonorant sequences /sm, sn, sl/ may undergo partial devoicing as well (e.g., /isla/ → [il̥la] 'island'; Martinez-Gil 2012). Both devoicing and higher constriction suggest gestural realignment, but not clearly full metathesis. Since it is unlikely that LINEARITY constraints would be specific to different manners and voicing, what other markedness constraints could prevent metathesis in /s/-voiced stop and /s/-sonorant sequences?

One possibility is, again, sonority:[15] onset sequences tend to rise in sonority (Sonority Sequencing Principle; Clements 1990; Blevins 1996). Stops are generally considered to be the least sonorous, followed by fricatives, sonorants, and vowels (Parker 2002). Within stops and fricatives, voiceless sounds are less sonorous than their voiced counterparts. The relative position of /h/ is not well-established (see Parker 2002 for overview), but it is usually considered to be more sonorous than voiceless stops. Thus, sonority rises Sevillian [ph, th, kh] sequences. In contrast, metathesis in /s/-voiced stop and /s/-sonorant clusters leads to sequences without the necessary sonority rise. In /s/-voiced stop sequences, the sequences resulting from metathesis [βh, ðh, ɣh] would have either a sonority fall or plateau, depending on the relative sonority of voiceless and voiced fricatives. The same applies to /s/-sonorant sequences: metathesized [mh, nh, lh] sequences would have falling sonority. Metathesis in /s/-voiced stop and /s/-sonorant sequences could be ruled out by constraints requiring sonority rises in onsets.

Another possibility is that metathesis does not occur in /s/-voiced stop and /s/-sonorant clusters for articulatory or perceptual reasons. Articulatorily, Kingston (1990) argues that

---

[15]Thanks to Michael Kenstowicz for this suggestion.

stops differ from sonorants and fricatives in that stops are asymmetrical: they have a release and burst at the offset resulting from the closure and sudden opening of the oral cavity. This release attracts and 'binds' glottal articulations to offset of the stop. In contrast, sonorants and fricatives are symmetrical: they have 'mirror-image' onsets and offsets because airflow is continuous. In Sevillian, then, metathesis could be limited to /s/-voiceless stop sequences because they are the only ones in which [h] can bind to a release.[16]

In terms of perception, metathesis in /s/-voiceless stop sequences may be the only context in which it is perceptually beneficial. Preconsonantal [h] has been argued to be difficult to perceive for a host of acoustic and auditory reasons, while [h] docked on a stop has been argued to be easier to perceive because it occurs on a highly-salient release (Bladon 1986; Kingston 1990; Silverman 2003). Metathesizing [h] around a sonorant or spirantized voiced stop might not have the same perceptual benefit. This hypothesis is supported by the cross-linguistic tendency for laryngeal metathesis to operate (almost always, to my knowledge) in the direction of [hC] to [Ch]. Furthermore, in some languages, including Cherokee and Korean, laryngeal metathesis occurs only when metathesis would 'dock' [h] on a voiceless stop release (Flemming 1996; Cho 2012). Otherwise, the laryngeal feature deletes (Korean), or metathesis fails to occur (Cherokee). Both these, and other cases of laryngeal metathesis, have been argued to be motivated by perceptual optimization of [h] (Yoon 2012). These proposals should be taken with caution though, because recent experimental studies find no evidence that [h] is more difficult to perceive preconsonantally than postconsonantally (Clayton 2010; Gilbert 2022).

For /s/-sonorant sequences, there is a final possibility: the features of the sonorant and [h] may be incompatible. Assuming that metathesis involves gestural overlap, as illustrated in (3), features of both segments must be realized simultaneously. In this case, the [spread glottis] gesture of [h] would need to be produced concurrently with the voicing of the sonorant. These are incompatible, since they require the vocal folds to both vibrate and *not* vibrate. Furthermore, Solé (2007) argues that aerodynamic requirements make the combinations of (a) nasalization and frication, and (b) nasality and voicelessness, difficult. Co-occurrence constraints against these combinations are common (e.g., Pulleyblank 1989; Steriade 1994), and are supported by the typology: voiceless or aspirated nasals are rare (Ladefoged and Maddieson 1996; Chirkova, Basset, and Amelot 2019).

4.3 Analysis summary

Speakers of Mexican, Argentinian, and Sevillian Spanish differ in both perceiving and producing /sC/ sequences. The dialects share the underlying /sC/ representation, but differ in how their phonological grammars map this representation to surface forms. In Mexican Spanish, coda /s/ is typically retained as [s] due to high-ranked faithfulness constraints. In Argentinian Spanish, coda /s/ typically debuccalizes to [h], because reducing the coda resolves the issue of having place features in the coda. Similarly, Sevillian Spanish debuccalizes /s/ to [h], but also metathesizes it with the following consonant, removing the coda from the preceding syllable altogether.

The crucial ranking differences are summarized in Table 9. In Mexican and Argentinian

---

[16]Recall that voiced stops in this context are spirantized, with no closure or release.

Spanish, LINEARITY is undominated. The dialects differ mainly in the ranking of faithfulness constraints: Mexican Spanish is highly faithful to underlying representations of /s/, while Argentinian and Sevillian Spanish are progressively less faithful. In the latter two varieties, markedness constraints are higher-ranked, resulting in /s/ reduction and, for Sevillian, metathesis.

Table 9: Crucial ranking differences between dialects

Dialect	Crucial ranking	Outcome (> alternative)	
Mexico	MAX[PL] ≫ CODACOND	Retention	/sC/ → [sC]
Argentina	CODACOND ≫ MAX[PL]	Debuccalization (> retention)	/sC/ → [hC]
Seville	*CODA-LAR, CODACOND ≫ MAX[PL]	Debuccalization (> retention)	
	MAX ≫ LINEARITY	Metathesis (> deletion)	/sC/ → [Ch]
	*CODA-LAR ≫ LINEARITY	Metathesis (> coda [h])	

The analysis presented in the previous section accounts for the outputs of the different dialects. However, the parallel analysis does not capture the observation that metathesis seems to be driven by the same pressure that drives debuccalization: to reduce coda material. This pressure is present in most varieties of Spanish, and a more complete analysis (likely serial) will need to account for this motivation. Furthermore, a complete formal account will need to integrate gestural representations with featural representations, integrate these representations with a theory of moraic timing, and capture variation.

5 Further arguments against analyzing [Ch] as an aspirated stop

The results from the perception experiment suggest that Sevillian listeners treat [h] in word-initial [Ch] sequences as belonging to the preceding word, not as part of an underlyingly aspirated stop. Is there reason to think that learners *could* represent [Ch] sequences as aspirated stops? Note that this is not a question about how learners choose a typical realization for /sC/ sequences. Instead, this is a question about how learners decide to classify the [Ch] forms they hear, possibly without connecting them to /sC/ sequences at all.

Many studies have found evidence that learners build phonetic—and then phonological—categories based on statistical distributions of variable phonetic realizations in the input (Maye, Werker, and Gerken 2002; Anderson, Morgan, and White 2003; Boersma, Escudero, and Hayes 2003; Pierrehumbert 2003). Surface realizations of a particular cue are variable, but cluster around a mean in a probability distribution. Learners build phonological categories around the modes in the learning data.

In Sevillian, what could learners do with [Ch]? Assuming that learners do not know the linguistic history of their language, a [Ch] sequence is most similar to a voiceless stop. However, it seems unlikely that [Ch] sequences will be perceptually categorized as Spanish voiceless stops. Voiceless stops and [Ch] sequences differ in several phonetic cues, but I focus on VOT/[h]-duration. Voiceless stops in Spanish have short-lag VOT (10-30ms). For each stop /p/, /t/, /k/, VOTs cluster around a certain mode, with some higher and lower values.

In perception, values within that distribution are classified as voiceless stops. Sevillian [Ch] sequences have 'VOTs' that are much too long (60-100ms) to fall within the 10-30ms range of VOTs for Spanish /p, t, k/. If a learner hears a cluster of short-lag VOTs, and also a cluster of productions with much longer 'VOT', they could posit [Ch] as the mode of its own separate aspirated stop category.

The frequency of metathesized [Ch] forms in the input affects statistical distributions, which affect the categories learners posit, what they consider 'typical' of that category, and what other pronunciations are acceptable but more marginal. While it is not clear just how frequent a realization must be for a category to be established, [Ch] surface forms are highly frequent in spoken language. Ruch (2008) reports rates of surface variants of /st/ in Seville based on production data from 53 speakers in multiple speech styles (Table 10). She distinguishes a postaffricated variant of /st/, [tˢ], which I group with [Ch] because both have the same (metathesized) linear order.

Table 10: Rates of variants of /st/ in Sevillian Spanish (Ruch 2008). Most frequent ones are in bold.

Variant		Conversation	Reading	Word List
Full [s]	[st]	3.3	17.4	**71.3**
Debucc.	[ht]	3.5	3.2	1.5
Surrounding	[hth], [sth]	11	9.5	5.2
Stop-h	[th]	**49.1**	**43.7**	13.0
PostAffricated	[ts]	22.0	13.7	6.6
Deleted or geminated	[t, tː]	11	12.4	2.4
		100%	100%	100%

In conversational speech—the speech style children are likely to hear most—/st/ is realized as metathesized in over 70% of forms. Even in more careful read speech, metathesized realizations constitute over 55% of forms. While a high proportion of metathesized forms is by no means evidence of a change in representation, [Ch] forms appear to be frequent enough in the input to suspect that learners *could* posit underlying /C^h/ based on this new phonetic mode and distribution. Providing a learning model is beyond the scope of this paper, but the data a learner would be exposed to contain enough metathesized forms to make the aspirated stop analysis a possibility, and to suggest that other factors may prevent this representation.

As already discussed in Section 2.3, stop-h sequences can arise in morphologically-derived environments, where the components belong to different morphemes. If learners know that the components come from different morphemes, it seems unlikely that they would posit them as parts of a single phoneme. In this section, I focus on two factors that may prevent speakers from representing [Ch] sequences as aspirated stops: the presence of phonological alternations across morpheme and word boundaries (Section 5.1) and sociolinguistic variation (Section 5.2). Finally, I consider what an aspirated stop analysis would look like, and argue that the difficulties raised by this analysis are further evidence that the aspirated stop analysis is

off-track (Section 5.3).

5.1 Phonological alternations

Just as phonologists make use of phonological alternations to analyze phonological systems, learners may use this information to learn underlying forms and their allophonic realizations. Sevillian learners must make two decisions about [Ch] sequences: (1) the underlying representation of [C], [h], and [Ch]; (2) the underlying location of /s/, if they posit that [h] derives from /s/.

I suggest that phonological alternations across word and morpheme boundaries provide evidence that /s/ and /C/ are separate consonants, that [h] derives from /s/, and that /s/ precedes /C/ underlyingly, even though the surface form is [Ch]. Prevocalically, word- and morpheme-final /s/ surfaces in a single location corresponding to orthography (Table 11a). Preceding /p, t, k/, word- and morpheme-final /s/ can be realized before or after /p, t, k/ ([s#t], [h#t] vs. [th]) (Table 11b).

Table 11: Phonological alternations across morpheme boundaries in Sevillian

Variants	a. Prevocalic (/s#V/)	b. Preconsonantal (/s# p, t, k/)
	/mas aɰwa/	/mas toɾta/
Full [s]	[ˈmas ˈaɰwa]	[ˈmas ˈtoɾta]
Debuccalized [h]	[ˈmah ˈaɰwa]	[ˈmah ˈtoɾta]
Metathesized [ɕh]	—	[ˈma ˈthoɾta]
Deleted	[ˈma ˈaɰwa]	[ˈma t(ː)oɾta]
	'more water'	'more cake'

I suggest the following reasoning as a plausible path for Sevillian learners. Learners deduce that a word like /mas/ has an underlying final /s/, because it often surfaces as [s] or [h] prevocalically, and sometimes preconsonantally. Now, suppose a learner encounters *más* in a preconsonantal context, like [ˈma ˈthoɾta]. If the learner already knows that *más* ends in /s/, then the [th] that surfaces after this word must derive from /s#t/. Phonologically-conditioned alternations provide evidence that stop-h sequences are clusters where /s/ belongs to a preceding morpheme, which may prevent learners from positing an aspirated stop representation.

Furthermore, although the evidence for the /s#t/ cluster representation comes from multi-morphemic strings, it could be transferred to monomorphemic words (/pasta/ → [patha]). Although monomorphemic words lack alternations to provide clear evidence for /sC/, the evidence from sociolinguistic variation and orthography are also compatible with this order, as I discuss in the next section.

5.2 Sociolinguistic variation

In addition to phonological alternations, sociolinguistic variation may also provide learners with evidence that [h] in [Ch] is a realization of underlying /s/ in an /sC/ cluster, and that

/s/ precedes /C/ in the underlying representation.

First, to learn that [h] derives from /s/, learners need to consider words with multiple coda consonants. Suppose that learners hear coda consonants /s/, /p/ and /t/, as in Table 12.[17] Like /s/, preconsonantal coda voiceless stops /p, t, k/ undergo debuccalization in Andalusian Spanish (Gerfen 2001; Campos-Astorkiza 2003), and also appear to undergo metathesis before /p, t, k/ (Gilbert 2022).[18]

Table 12: Variants of coda /s/, /p/, /t/ that learners encounter

Type	/pasta/	/apto/	/atleta/
a. Faithful	[pasta]	[apto]	[atleta]
b. Debuccalization + (Gemination)	[paht(ː)a]	[aht(ː)o]	[ahl(ː)eta]
c. (Gemination) + metathesis	[pat(ː)ha]	[at(ː)ho]	—
d. Deletion + gemination	[patːa]	[atːo]	[alːeta]
e. Deletion	[pata]	[ato]	[aleta]
	'pasta'	'apt'	'athlete'

Learners are exposed to multiple realizations of /sC/, /pC/, and /tC/, including coda [h] (Table 12b). When learners compare debuccalized forms (Table 12b) to faithful forms of the same word (Table 12a) they can deduce that [h] corresponds to multiple consonants, including [s, p, t]. /s, p, t/ map predictably to [h] in certain environments, so /s, p, t/ must be underlying.

Thus, variation may give learners evidence that [pahta] comes from underlying /pasta/, that [h] is derived from /s/, and that /s/ precedes /C/ underlyingly. The presence of variation shows learners that metathesis, gemination of the following consonant, and deletion of coda /s/ are simply different realizations of the same underlying representation. Differently from phonological alternations, which require morpheme boundaries, this variation also shows learners that [Ch] forms are derived from /sC/ morpheme-internally. The observation that many phonological variables—including coda /s/ in Spanish varieties—show synchronic variation that is stable across generations (Labov 1994: 583–585) is consistent with the hypothesis that alternations and variability may provide enough evidence to prevent representational change.

Several existing approaches to learning posit that variability plays a crucial role in forming underlying representations (e.g., Dell 1981), and that the amount of surface variability can be a cue to differences in representational status. For example, epenthetic vowels in Lebanese Arabic and Brazilian Portuguese are phonetically different from their lexical counterparts (Gouskova and Hall 2009; Kouneli 2016), and Brazilian Portuguese lexical vowels are more consistently present than epenthetic vowels (Kouneli 2016). Rasin and Katzir

---

[17]In Seville, these variants are sociolinguistically conditioned, within and across speakers. Conditioning factors include gender, age, and level of education (Ruch 2008; Horn 2013; Ruch and Peters 2016; Moya Corral and Tejada Giráldez 2020). There is also stylistic variation: some variants are more frequent in certain speech styles (Ruch 2008; see also Table 10).

[18]Coda /k/ preceding onset /p, t, k/—the consonants that could show metathesis—is extremely rare.

(2016) present a computational model whose conclusions about underlying representations are based on comparing forms derived by variation. The presence of variability leads the learner to posit underlying representations that balance economy and restrictiveness, capturing unpredictable information and leaving out predictable (derivable) information. In the case of Sevillian, the presence of /s/ is unpredictable, and thus would be posited in the UR. In contrast, [h] is derivable.

Further modeling and psycholinguistic experimentation would be useful to shed light on how variation and morphological complexity affect the types of representations posited by learners.

5.3 Alternative: Aspirated stop analysis

In addition to phonological alternations and sociolinguistic variation, which may prevent learners from positing an underlying aspirated stop representation for [Ch] sequences, the system of contrasts resulting from an aspirated stop analysis is implausible. Before explaining why, I address the proposals that Sevillian [Ch] sequences are phonologizing into aspirated stops /pʰ, tʰ, kʰ/.

Previous work has made several arguments, both phonetic and phonological, that [Ch] sequences are phonologizing into aspirated stops. First, Gylfadottir (2015) notes that stop-h sequences can surface phrase-initially (/esˈta/ 'is-3sg' as [ˈtha]). Because Spanish disallows onset clusters with /s/ (*sta), she argues that analyzing stop-h sequences as aspirated stops avoids violating this restriction (✔tʰa).

A second argument for aspirated stops is that there is a further change in some metathesizing regions from stop-h to stop-s (*postaffrication*; /ˈpasta/ → [ˈpat̲h̲a] → [ˈpat̲s̲a]) (Moya Corral 2007; Ruch 2010; Vida Castro 2016; Del Saz 2019). Gylfadottir (2015) takes this to be fortition of an underlying aspiration feature. An alternative consistent with the cluster analysis is that affrication is due to articulatory and perceptual factors. Affrication occurs only with /st/ sequences, and previous studies find that [th] releases have more high-frequency energy than [ph] releases (Harrington 2010: 104), and this high-frequency energy on [th] could be reinterpreted as [ts] without positing an underlying aspiration feature.

A third argument comes from duration measurements. O'Neill (2009) argues that each word has two representations: one with an aspirated phoneme (/ˈpatʰa/ and another with an unaspirated phoneme (/ˈpasta/). He argues that [Ch] realizations come from the underlying aspirated phoneme, and that duration measurements support the single-segment representation of [Ch]. In words with stop-h sequences, he finds that the preceding vowel is the same length as when it is in an open syllable ([pa̲tʰa] ∼ [ˈpa̲ta]), which would be unexpected if the syllable were underlyingly closed. Other studies, however, find inconsistent results for the duration of the components (vowels, stop closure, [h]) (Gerfen 2002; Torreira 2006; Torreira 2007; Parrell 2012; Ruch and Harrington 2014; Ruch and Peters 2016). Even if they were consistent, duration is not a clear argument for diagnosing representational status (see Gouskova and Stanton 2021: 183–6 for a critique). Durational arguments do not consistently distinguish whether a sequence functions as a single segment that leaves the preceding syllable open, or a cluster in which the first segment closes the preceding syllable.

A final argument put forth in favor of the aspirated stop analysis is that Sevillian listeners make a categorical distinction between minimal pairs based only on the duration of [h] in

stop-h sequences (Ruch and Peters 2016). Ruch and Peters (2016) interpret this finding as evidence of incipient phonologization. However, a sharp perceptual boundary does not mean that [h] belongs to the stop underlyingly. Listeners could map metathesized [h] to underlying /sC/, and still categorically distinguish surface [C]-[Ch].

When considered in more detail, the aspirated stop analysis presents analytical difficulties. I argue that this analysis is plausible for morpheme-internal sequences: [Ch] sequences could be analyzed as a single segment, and the distinction between e.g., /kapa/ ('cape') - /kaspa/ ('dandruff') could be restructured to /kapa/-/kap^{h}a/. If the change progresses to a point where there is little variation (currently variation is extensive; see Section 5.2), this kind of restructuring is plausible. In contrast, the aspirated stop analysis seems to be untenable for word- and morpheme-initial stop-h sequences. These arise across word and morpheme boundaries, and are derived by the preceding phonological context (Section 5.1).

Analyzing word/morpheme-initial [Ch] as an aspirated stop /C^h/ makes little sense when it arises from word/morpheme concatenation. To illustrate why, I consider nonce words of a hypothetical Sevillian-like system that has unaspirated and aspirated stops /p, t, k/ and /p^h, t^h, k^h/. In this system, aspirated stops would contrast with unaspirated stops (18).

(18) Hypothetical contrast between unaspirated and aspirated stops

    a. /maka/ vs. /mak^{h}a/   Word-medial aspiration contrast

    b. /paka/ vs. /p^{h}aka/   Word-initial aspiration contrast (untenable)

This contrast would be possible word-medially (18a), but is untenable word-initially (18b). Word-initially, the contrast would be neutralized, as illustrated in (19): unaspirated stops /p, t, k/ acquire [h] at their release allophonically, following words ending ending in /s/ (19b), so that they look just like underlying aspirated stops on the surface (19a). A realization that is predictably conditioned by the environment is not contrastive in that environment. The same arguments hold for word-internal aspiration that arises across morpheme boundaries.

(19) Hypothetical aspiration contrast is neutralized word-initially

    a. /tjenes p^{h}aka/ → ['tjene p^{h}aka]

    b. /tjenes paka/ → ['tjene p^{h}aka]

The aspirated stop analysis could be saved by arguing that [Ch] sequences are phonologizing only word-medially, but not across morpheme and word boundaries. According to this analysis, learners would distinguish processes that are similar on the surface and phonologize only one of them. But this seems unlikely, given that [Ch] is phonetically similar across contexts (Ruch 2008; Horn 2013).

6 Conclusion

I have argued that Sevillian [Ch] sequences are realizations of underlying /sC/ clusters. Results from a perception study support this analysis: Sevillian listeners attribute [h] in [Ch] to the preceding word, not to the stop itself. Listeners of other dialects (Mexico,

Argentina) do not interpret [h] in stop-h sequences as linguistically relevant. The analysis ties perception to production. Although listeners of the three dialects share an underlying /sC/ representation, their phonological grammars map /sC/ to different surface forms, a difference that also seems to affect perception. The Sevillian grammar maps /sC/ to [Ch] in production, and Sevillian listeners are able to undo metathesis to get from [Ch] back to /sC/ in perception. The grammars of Mexican and Argentinian Spanish do not produce [Ch] forms from /sC/, making it difficult for these listeners to perceive [Ch] as /sC/ too. In the future, a fuller analysis of Sevillian metathesis will be necessary. A formalization in gestural terms may be the most promising, but will raise challenges in terms of integrating gestural and featural analyses.

Finally, I suggested that Sevillian stop-h sequences may not be phonologizing into aspirated stops due to the presence of phonological alternations and sociolinguistic variation, which give learners evidence for the cluster representation. The odd phonological system resulting from the aspirated stop analysis is also implausible, and would likely be difficult to learn. Further work on this phenomenon has the potential to shed light on our understanding of metathesis, phonological categories and alternations, and the relationship between gestures and features.

References

Alonso, Amado (1945). "Una ley fonológica del español". In: *Hispanic Review* 13.2, 91–101.

Alvar, Manuel (1955). "Las hablas meridionales de España y su interés para la lingüística comparada". In: *Revista de Filología Española* 39, 284–313.

Anderson, Jennifer L., James L. Morgan, and Katherine S. White (2003). "A statistical basis for speech sound discrimination". In: *Language and Speech* 46.2-3, 155–182. DOI: 10.1177/00238309030460020601.

Arnason, Kristján (1986). "The segmental and supersegmental status of preaspiration in modern Icelandic". In: *Nordic Journal of Linguistics* 9.1, 1–23.

Asyik, Abdul Gani (1987). "A contextual grammar of Acehnese sentences". PhD Dissertation. Ann Arbor, Michigan: University of Michigan.

Beckman, Jill, Michael Jessen, and Catherine Ringen (2013). "Empirical evidence for laryngeal features: Aspirating vs. true voice languages". In: *Journal of Linguistics* 49.2, 259–284. DOI: 10.1017/S0022226712000424.

Beckman, Jill and Catherine Ringen (2009). "A typological investigation of evidence for [sg] in fricatives". In: Manchester Phonology Meeting.

Bedinghaus, Robert William (2015). "The effect of exposure to phonological variation on perceptual categorization and lexical access in second language Spanish: The case of /s/-aspiration in Western Andalusian Spanish". PhD Dissertation. Bloomington: Indiana University.

Bladon, Anthony (1986). "Phonetics for hearers". In: *Language for hearers*. Ed. by Graham McGregor. Great Britain: Pergamon Press, 1–24.

Blevins, Juliette (1996). "The syllable in phonological theory". In: *The handbook of phonological theory*. Ed. by John Goldsmith. Oxford, UK: Blackwell Publishing, 206–244.

Blevins, Juliette and Andrew Garrett (2004). "The evolution of metathesis". In: *Phonetically based phonology*. Ed. by Bruce Hayes, Donca Steriade, and Robert Kirchner. New York: Cambridge University Press, 117–156.

Boersma, Paul, Paolo Escudero, and Rachel Hayes (2003). "Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories". In: *Proceedings of ICPhS XV*. Ed. by Maria-Josep Solé, Daniel Recasens, and Joaquín Romero. Barcelona: Universitat Autónoma de Barcelona, 1013–1016.

Bowden, John and John Hajek (1999). "Taba". In: *The handbook of the IPA*. Cambridge, UK: Cambridge University Press, 143–146.

Campos-Astorkiza, Rebeka (2003). "Compensatory lengthening as root number preservation". In: *Proceedings of the XVII International Congress of Linguists*. Ed. by E. Hajicová, A. Kotesovcová, and J. Mirovský. Prague: Matfyzpress, MFF UK.

— (2012). "The phonemes of Spanish". In: *The handbook of Hispanic linguistics*. Ed. by José Ignacio Hualde, Antxon Olarrea, and Erin O'Rourke. 1st edition. Blackwell Handbooks in Linguistics. West Sussex, UK/Malden, MA: Wiley-Blackwell, 89–110. ISBN: 978-1-4051-9882-0.

— (2016). "Procesos fonológicos". In: *Enciclopedia de lingüística hispánica*. Ed. by Javier Gutiérrez-Rexach. Vol. 1. London/New York: Routledge Taylor & Francis Group, 847–858.

Canfield, Tracy A. (2015). "Metathesis is real, and it is a regular relation". PhD Dissertation. Washington, DC: Georgetown University.

Casserly, Elizabeth D. (2012). "Gestures in optimality theory and the laryngeal phonology of Faroese". In: *Lingua* 122.1, 41–65. DOI: `10.1016/j.lingua.2011.11.003`.

Catalán, Diego (1971). "En torno a la estructura silábica del español de ayer y del español de mañana". In: *Sprache und Geschichte, Festschrift Harri Meier*. Munich: Fink-Verlag.

Chirkova, Katia, Patricia Basset, and Angélique Amelot (2019). "Voiceless nasal sounds in three Tibeto-Burman languages". In: *Journal of the International Phonetic Association* 49.1, 1–32. DOI: `10.1017/S0025100317000615`.

Cho, Hyesun (2012). "Laryngeal feature mobility in Cherokee and Korean". In: *Language and Linguistics* 56, 283–306.

Cho, Taehong and Peter Ladefoged (1999). "Variation and universals in VOT: Evidence from 18 languages". In: *Journal of Phonetics* 27, 207–229. DOI: `10.1006/jpho.1999.0094`.

Clayton, Ian (2010). "On the natural history of preaspirated stops". PhD Dissertation. Chapel Hill: University of North Carolina Chapel Hill.

Clements, G. N. (1990). "The role of the sonority cycle in core syllabification". In: *Papers in laboratory phonology I: Between the grammar and physics of physical speech*. Ed. by John Kingston and Mary Beckman. Cambridge: Cambridge University Press, 283–333.

Colina, Sonia (2012). "Syllable structure". In: *The handbook of Hispanic linguistics*. 1st edition. Blackwell Handbooks in Linguistics. West Sussex, UK/Malden, MA: Wiley-Blackwell, 133–151.

Cronenberg, Johanna et al. (2020). "A dynamic model of the change from pre- to post-aspiration in Andalusian Spanish". In: *Journal of Phonetics* 83, 1–22. DOI: `10.1016/j.wocn.2020.101016`.

Davis, Stuart and Mi-Hui Cho (2003). "The distribution of aspirated stops and /h/ in American English and Korean: An alignment approach with typological implications". In: *Linguistics* 41.4, 607–652.

Del Saz, María (2019). "From postaspiration to affrication: New phonetic contexts in Western Andalusian Spanish". In: *Proceedings of ICPhS XIX*. Ed. by Sasha Calhoun et al., 760–764.

Dell, François (1981). "On the learnability of optional phonological rules". In: *Linguistic Inquiry* 12.1, 31–37.

Erker, Daniel (2012). "An acoustically based sociolinguistic analysis of variable coda /s/ production in the Spanish of New York City". PhD Dissertation. New York: New York University.

Flemming, Edward (1996). "Laryngeal metathesis and vowel deletion in Cherokee". In: *Cherokee papers from UCLA*. Vol. 16. UCLA Occasional Papers in Linguistics. Los Angeles: Department of Linguistics, UCLA, 23–44.

Galvano, Amber and Nicholas Henriksen (2021). "Phonemic voiceless stops in Western Andalusian Spanish: The confluence of post-aspiration and intervocalic voicing". In: Presentation at MidPhon25.

Gerfen, Chip (2001). "A critical view of licensing by cue: Codas and obstruents in Eastern Andalusian Spanish". In: *Segmental phonology in Optimality Theory*. Cambridge University Press, 183–205. DOI: `10.1017/cbo9780511570582.007`.

— (2002). "Andalusian codas". In: *Probus* 14, 247–277. DOI: `10.1515/prbs.2002.010`.

Gilbert, Madeline (2022). "An experimental and formal investigation of Sevillian Spanish metathesis". PhD Dissertation. New York: New York University.

Goldsmith, John (1981). "Subsegmentals in Spanish phonology: An autosegmental approach". In: *Selected papers from the 9th Linguistic Symposium on Romance Languages (LSRL IX)*. Ed. by William W. Cressey and Donna Jo Napoli. Washington, D.C.: Georgetown University Press, 1–16.

Gouskova, Maria (2004). "Relational hierarchies in Optimality Theory: The case of syllable contact". In: *Phonology* 21.2, 201–250.

Gouskova, Maria and Nancy Hall (2009). "Acoustics of epenthetic vowels in Lebanese Arabic". In: *Phonological argumentation: Essays on evidence and motivation.* Ed. by Steve Parker. London, UK: Equinox, 203–225. DOI: `10.1558/equinox.29399`.

Gouskova, Maria and Juliet Stanton (2021). "Learning complex segments". In: *Language* 97.1, 151–193. DOI: `10.1353/lan.2021.0011`.

Gylfadottir, Duna (2015). "An investigation of Andalusian stop cluster post-aspiration in naturalistic speech". In: Presentation at the 39th Penn Linguistics Conference.

Hammond, Robert M. (2001). *The sounds of Spanish: Analysis and application (with special reference to American English).* Somerville, MA: Cascadilla Press.

Harrington, Jonathan (2010). "Acoustic phonetics". In: *The handbook of phonetic sciences.* Ed. by William J. Hardcastle, John Laver, and Fiona E. Gibbon. 1st edition. Wiley, 81–129. DOI: `10.1002/9781444317251.ch3`.

Haugen, Einar (1958). "The phonemics of Modern Icelandic". In: *Language* 34, 55–88.

Henriksen, Nicholas (2017). "Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception". In: *Journal of Phonetics* 63, 106–126. DOI: `10.1016/j.wocn.2017.05.001`.

Herrero de Haro, Alfredo (2016). "Context and vowel harmony: Are they essential to identify underlying word-final /s/ in Eastern Andalusian Spanish?" In: *Dialectología* 20, 107–145.

— (2017). "The phonetics and phonology of Eastern Andalusian Spanish: A review of literature from 1881 to 2016". In: *Íkala, Revista de Lenguaje y Cultura* 22.2, 313–357. DOI: `10.17533/udea.ikala.v22n02a09`.

Holmberg, Eva and Alan Gibson (1979). "On the distribution of [h] in the languages of the world". In: *PERILUS I. Stockholm University.*

Honeybone, Patrick (2005). "Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications". In: *The internal organization of phonological segments.* Ed. by Marc van Oostendorp et al. Studies in Generative Grammar 77. Berlin, New York: Mouton de Gruyter, 319–354.

Horn, Meagan E. (2013). "An acoustic investigation of postaspirated stops in Seville Spanish". MA thesis. Columbus, Ohio: The Ohio State University.

Hualde, José Ignacio (1989a). "Delinking processes in Romance". In: *Selected Papers from the 17th Linguistic Symposium on Romance Languages (LSRL XVII), Rutgers University, 27-29 March 1987.* Studies in Romance Linguistics. Amsterdam/Philadelphia: John Benjamins Publishing Company, 177–193.

— (1989b). "Procesos consonánticos y estructuras geométricas en español". In: *Lingüística* 1, 7–44.

— (2014). *Los sonidos del español.* Cambridge: Cambridge University Press.

Itô, Junko (1989). "A prosodic theory of epenthesis". In: *Natural Language & Linguistic Theory* 7.2, 217–259. DOI: 10.1007/BF00138077.

Iverson, Gregory K. and Joseph C. Salmons (1995). "Aspiration and laryngeal representation in Germanic". In: *Phonology* 12.3, 369–396.

Kingston, John (1990). "Articulatory binding". In: *Papers in laboratory phonology 1: Between the grammar and physics of speech.* Ed. by J. Kingston and M. E. Beckman. Cambridge: Cambridge University Press.

Kouneli, Maria (2016). "Acoustics of epenthetic vowels in Brazilian Portuguese". In: *The Journal of the Acoustical Society of America* 140.4, 3109–3109. DOI: 10.1121/1.4969715.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen (2017). "lmerTest package: Tests in linear mixed effects models". In: *Journal of Statistical Software* 82.13, 1–26. DOI: 10.18637/jss.v082.i13.

Labov, William (1994). *Principles of linguistic change: Internal factors.* Cambridge, MA: Blackwell Publishers.

Ladefoged, Peter and Ian Maddieson (1996). *The sounds of the world's languages.* Cambridge, MA: Blackwell Publishers.

Lapesa, Rafael (1981). *Historia de la lengua española.* 9th ed. Madrid: Gredos.

Lenth, Russell (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* URL: https://CRAN.R-project.org/package=emmeans.

Lloret, Maria-Rosa and Violeta Martínez-Paricio (2020). "El Modelo de las Estructuras Paralelas y el debilitamiento de la /s/ implosiva en español". In: *Sintagma* 32, 23–38.

Lombardi, Linda (1991). "Laryngeal features and laryngeal neutralization". PhD Dissertation. Amherst, MA: University of Massachusetts Amherst.

— (2001). "Why Place and Voice are different: Constraint-specific alternations in Optimality Theory". In: *Segmental phonology in Optimality Theory.* Ed. by Linda Lombardi. 1st. Cambridge University Press, 13–45. DOI: 10.1017/CBO9780511570582.002.

Malmberg, Bertil (1965). "La estructura silábica del español". In: *Estudios de fonética hispánica.* Madrid: Consejo Superior de Investigaciones Científicas, 3–28.

Martinez-Gil, Fernando (2012). "Main phonological processes". In: *The handbook of Hispanic linguistics.* Ed. by Jose Ignacio Hualde, Antxon Olarrea, and Erin O'Rourke. 1st edition. Blackwell Handbooks in Linguistics. West Sussex, UK/Malden, MA: Wiley-Blackwell, 111–131.

Mason, Keith William (1994). "Comerse las eses: A selective bibliographic survey of /s/ aspiration and deletion in dialects of Spanish". PhD Dissertation. Ann Arbor, Michigan: University of Michigan.

Maye, Jessica, Janet F Werker, and LouAnn Gerken (2002). "Infant sensitivity to distributional information can affect phonetic discrimination". In: *Cognition* 82.3, B101–B111. DOI: 10.1016/S0010-0277(01)00157-3.

McCarthy, John (2008a). *Doing Optimality Theory: Applying theory to data.* John Wiley & Sons, Inc.

— (2008b). "The gradual path to cluster simplification". In: *Phonology* 25, 271–319. DOI: 10.1017/S0952675708001486.

McCarthy, John and Alan Prince (1995). "Faithfulness and reduplicative identity". In: *University of Massachusetts occasional papers 18: Papers in Optimality Theory.* Ed. by Jill

Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey. Vol. 18. Amherst, MA: GLSA, University of Massachusetts, 249–384.

McCarthy, John J. and Joe Pater, eds. (2016). *Harmonic grammar and harmonic serialism*. Advances in Optimality Theory. Sheffield, UK; Bristol, CT: Equinox Publishing Ltd. ISBN: 978-1-84553-149-2.

Minegishi, M. (2006). "Khmer". In: *Encyclopedia of language & linguistics*. Elsevier, 189–192. DOI: 10.1016/B0-08-044854-2/05056-2.

Morris, Richard E. (2000). "Constraint interaction in Spanish /s/-aspiration: Three peninsular varieties". In: *Hispanic linguistics at the turn of the millennium: Papers from the 3rd Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Press.

Moya Corral, Juan Antonio (2007). "Noticia de un sonido emergente: La africada dental procedente del grupo -st- en Andalucía". In: *Revista de Filología* 25, 457–465.

Moya Corral, Juan Antonio and María de la Sierra Tejada Giráldez (2020). "Patterns of the linguistic change in Andalusia". In: *Spanish in Context* 17.2, 200–220.

Murray, Robert W. and Theo Vennemann (1983). "Sound change and syllable structure in Germanic phonology". In: *Language* 59.3, 514. DOI: 10.2307/413901.

O'Neill, Paul (2009). "S-aspiration and occlusives in Andalusian Spanish: Phonetics or phonology?" In: *Oxford University Working Papers in Linguistics, Philology & Phonetics* 12. Ed. by 'Oiwi Parker Jones and Elinor Payne, 73–85.

— (2010). "Variación y cambio en las consonantes oclusivas del español de Andalucía". In: *Estudios de fonética experimental* XIX, 11–41.

Parker, Stephen (2002). "Qualifying the sonority hierarchy". PhD Dissertation. Amherst, MA: University of Massachusetts Amherst.

Parrell, Benjamin (2012). "The role of gestural phasing in Western Andalusian Spanish aspiration". In: *Journal of Phonetics* 40, 37–45. DOI: 10.1016/j.wocn.2011.08.004.

Pierrehumbert, Janet B. (2003). "Phonetic diversity, statistical learning, and acquisition of phonology". In: *Language and Speech* 46.2-3, 115–154. DOI: 10.1177/00238309030460020501.

Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar*. Malden/Oxford: Wiley Blackwell.

Pulleyblank, Douglas (1989). "Patterns of featural cooccurrence: the case of nasality". In: *Coyote Papers 9*. Ed. by S. Lee Fulmer, M. Ishihara, and Wendy Wiswall. Tucson, Arizona, 98–115.

Rasin, Ezer and Roni Katzir (2016). "On evaluation metrics in Optimality Theory". In: *Linguistic Inquiry* 47.2, 235–282.

RCoreTeam (2020). *R: A language and environment for statistical computing*. Vienna, Austria. URL: https://www.R-project.org.

Riehl, Anastasia (2008). "The phonology and phonetics of nasal obstruent sequences". PhD Dissertation. Ithaca, NY: Cornell University.

Romero, Joaquín (1995). "Gestural organization in Spanish: An experimental study of spirantization and aspiration". PhD Dissertation. Storrs, CT: University of Connecticut.

Rosner, Burton S. et al. (2000). "Voice-onset times for Castilian Spanish initial stops". In: *Journal of Phonetics* 28, 217–224. DOI: 10.1006/jpho.2000.0113.

Ruch, Hanna (2008). "La variante [ts] en el español de la ciudad de Sevilla: Aspectos fonético-fonológicos y sociolingüísticos de un sonido innovador". MA thesis. Zürich, Germany: University of Zürich.

Ruch, Hanna (2010). "Affrication of /st/-clusters in Western Andalusian Spanish: variation and change from a sociophonetic point of view". In: *Proceedings of Sociophonetics, at the crossroads of speech variation, processing and communication, Pisa, December 14th-15th, 2010.* Ed. by Silvia Calamai, Chiara Celata, and Luca Ciucci. Pisa, Italy: Edizioni della Normale, 61–64.

— (2013). "Investigating a gradual metathesis: Phonetic and lexical factors on /s/ aspiration in Andalusian Spanish". In: *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 41* 19.2, 171–180.

Ruch, Hanna and Jonathan Harrington (2014). "Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish". In: *Journal of Phonetics* 45, 12–25. DOI: `10.1016/j.wocn.2014.02.009`.

Ruch, Hanna and Sandra Peters (2016). "On the origin of post-aspirated stops: Production and perception of /s/ + voiceless stop sequences in Andalusian Spanish". In: *Laboratory Phonology* 7.1, 1–36. DOI: `10.5334/labphon.2`.

Schiller, Eric (1994). "Khmer nominalizing and causativizing infixes". In: *Second Annual Meeting of the Southeast Asian Linguistics Society.* Ed. by Karen L. Adams and Thomas John Hudak. Arizona State University, Program for Southeast Asian Studies, 309–326.

Schmidt, Lauren (2013). "Regional variation in the perception of sociophonetic variants of Spanish /s/". In: *Selected proceedings of the 6th Workshop on Spanish Sociolinguistics.* Ed. by Ana Maria Carvalho and Sara Beaudrie. Somerville, MA: Cascadilla Proceedings Project, 189–202.

Silverman, Daniel (2003). "On the rarity of pre-aspirated stops". In: *Journal of Linguistics* 39.3, 575–598. DOI: `10.1017/S002222670300210X`.

Solé, Maria-Josep (2007). "Compatibility of features and phonetic content. The case of nasalization". In: *Proceedings of ICPhS XVI*, 261–266.

Steriade, Donca (1994). "Complex onsets as single segments: the Mazateco pattern". In: *Perspectives in phonology.* Ed. by Jennifer Cole and Charles Kisseberth. Stanford, CA: CSLI Publications, 203–291.

Suh, Chang-Kook (2001). "Aspects of phonetics and phonology of Icelandic preaspiration". In: *The Phonology-Morphology Circle of Korea* 6, 63–83.

Suomi, Kari, Juhani Toivanen, and Riikka Ylitalo (2008). *Finnish sound structure: Phonetics, phonology, phonotactics and prosody.* Studia Humaniora Ouluensia 9. Oulu, Finland: Oulu University Press.

Takahashi, Chikako (2019). "No transposition in Harmonic Serialism". In: *Phonology* 36.4, 695–726. ISSN: 0952-6757, 1469-8188. DOI: `10.1017/S0952675719000344`.

Thráinsson, Höskuldur (1978). "On the phonology of Icelandic pre-aspiration". In: *Nordic Journal of Linguistics* 1, 3–54.

Torreira, Francisco (2006). "Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison". In: *Selected proceedings of the 9th Hispanic Linguistics Symposium.* Ed. by Nuria Sagarra and Almeida Jacqueline Toribio. Somerville, MA: Cascadilla Proceedings Project, 113–120.

— (2007). "Pre- and postaspirated stops in Andalusian Spanish". In: *Segmental and prosodic issues in Romance phonology.* Ed. by Pilar Prieto, Joan Mascaró, and Maria-Josep Solé. Vol. 282. Current Issues in Linguistic Theory, 67–82. DOI: `https://doi.org/10.1075/cilt.282.06tor`.

Torreira, Francisco (2012). "Investigating the nature of aspirated stops in Western Andalusian Spanish". In: *Journal of the International Phonetic Association* 42.1, 49–63. DOI: `10.1017/s0025100311000491`.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie*. Prague: Travaux du cercle linguistique de Prague 7.

Vaux, Bert (1998). "The laryngeal specifications of fricatives". In: *Linguistic Inquiry* 29, 497–511.

Vennemann, Theo (1988). *Preference laws for syllable structure and the explanation of sound change with special reference to German, Germanic, Italian, and Latin*. Berlin: De Gruyter Mouton. ISBN: 978-3-11-084960-8.

Vida-Castro, Matilde (2015). "Características acústicas de la aspiración de /-s/ implosiva en el español hablado en Málaga. Hacia la resilabificación prestigiosa de un segmento subyacente entre los jóvenes universitarios". In: *Lenguas, lenguaje y lingüística. Contribuciones desde la Lingüística General*. Ed. by Adriana Gordejuela Senosiáin et al. Pamplona: Servicio de Publicaciones de la Universidad de Navarra, 495–504.

— (2016). "Correlatos acústicos y factores sociales en la aspiración de /-s/ preoclusiva en la variedad de Málaga (España). Análisis de un cambio fonético en curso". In: *Lingua Americana* 20.38, 15–36.

Villena-Ponsoda, Juan Andrés (2008). "Sociolinguistic patterns of Andalusian Spanish". In: *International Journal of the Sociology of Language* 2008.193/194. DOI: `10.1515/IJSL.2008.052`.

Widdison, Kirk (1995). "Two models of Spanish s-aspiration". In: *Language Sciences* 17.4, 329–434.

Yoon, Tae-Jin (2012). "Laryngeal metathesis as perceptual optimization". In: *Studies in Phonetics, Phonology and Morphology* 18.3, 413–435.

Zehr, Jeremy and Florian Schwarz (2018). *PennController for Internet Based Experiments (IBEX)*. URL: `https://doi.org/10.17605/OSF.IO/MD832`.

7 Appendix A: Stimuli

Table 13: Full illustration of test, control, and filler items for *pali.* [p] = Step-0; [pʰ] = Step-1; [ph] = Step-2/Naturally long.

Type	Phrase recorded in	Word	Answer 1	Answer 2	Listeners hear
Test	Juan tiene pali (3sɢ)	pali[Step-0]	Juan	tú	[* 'tjene 'pali]
(3)	Juan tiene pali (3sɢ)	pali[Step-1]	Juan	tú	[* 'tjene 'pʰali]
	Juan tiene pali (3sɢ)	pali[Step-2]	Juan	tú	[* 'tjene 'phali]
Controls	Juan tiene pali (3sɢ)	pali[orig]	Juan	tú	[* 'tjene 'pali]
(2)	Tú tienes pali (2sɢ)	pali[orig]	Juan	tú	[* 'tjene 'phali]
Fillers	Juan tiene pali (3sɢ)	pali[orig]	Juan	nosotros	[* 'tjene 'pali]
(10)	Juan tiene pali (3sɢ)	pali[orig]	Juan	ellas	[* 'tjene 'pali]
	Tú tienes pali (2sɢ)	pali[orig]	tú	nosotros	[* 'tjene p'hali]
	Tú tienes pali (2sɢ)	pali[orig]	tú	ellas	[* 'tjene 'phali]
	Nosotros tenemos pali (1ᴘʟ)	pali[orig]	nosotros	tú	[* te'nemo 'phali]
	Nosotros tenemos pali (1ᴘʟ)	pali[orig]	nosotros	Juan	[* te'nemo 'phali]
	Nosotros tenemos pali (1ᴘʟ)	pali[orig]	nosotros	ellas	[* te'nemo 'phali]
	Ellas tienen pali (3ᴘʟ)	pali[orig]	ellas	tú	[* 'tjenen 'pali]
	Ellas tienen pali (3ᴘʟ)	pali[orig]	ellas	Juan	[* 'tjenen 'pali]
	Ellas tienen pali (3ᴘʟ)	pali[orig]	ellas	nosotros	[* 'tjenen 'pali]

8 Appendix B: Model results

Table 14: Model results for Mexico, Seville, and Argentina data together

	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	-1.160	0.390	-2.970	0.003	**
Step-1	0.162	0.357	0.452	0.651	
Step-2	0.252	0.357	0.707	0.480	
Mexico	-1.723	0.496	-3.474	0.001	***
Seville	-1.345	0.481	-2.794	0.005	**
Step-1:Mexico	0.408	0.386	1.057	0.291	
Step-2:Mexico	0.477	0.383	1.247	0.213	
Step-1:Seville	2.171	0.363	5.971	0.000	***
Step-2:Seville	3.845	0.386	9.953	0.000	***