An experimental and formal investigation of Sevillian Spanish metathesis

by

Madeline B. Gilbert

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

May, 2022

Maria Gouskova

Acknowledgements

I had plenty of advice from sage people who had done PhDs before starting mine, but nobody warned me that this is, by far, the hardest part of the dissertation to write. I chose to come to NYU because of the people in this department. What I did not know was that that choice was going to make this part so difficult. I cannot do justice to everyone who has helped me get to this point—and get here with so much laughter and so many good memories—but I will do my best.

Maria Gouskova, Lisa Davidson, Gillian Gallagher, Juliet Stanton, and Joseph Casillas have been the sharpest, most thorough, inspiring, and brilliant dissertation committee I can imagine. I have learned a lot from you all about academics, and have also enjoyed your company over the years.

Maria, it has been a privilege to work with you. You have been my de facto advisor since day one, when I was not even a phonologist yet. Your level of engagement, thoroughness, and insistence on clarity have made me a better scholar, and I am grateful for the unbelievable amount of time and effort you have put into teaching me how to think, write, and exist in academia. Thank you for challenging me, for believing in me more than I believed in myself, for your support and patience and hilarious comments on my drafts. I could not have asked for a better guide through this process. Evidence of your guidance is laced throughout this dissertation, and your influence will stay with me long after I leave NYU.

My other committee members have also done a heroic amount of work to help me through this process, and have been wonderful company along the way. Lisa, I have always appreciated

your respect for the data, your ability to look at the bigger picture, and your tendency to question the assumptions we all take for granted. Your intense curiosity, enthusiasm for all things sound-related, and infectious laugh have lit up some of the most difficult days. Thank you for your down-to-earth perspective, encouragement, understanding, 2pm coffee, and babysitting my succulents. Gillian, thank you for constantly drawing me back to the big, broad questions of phonology, and the importance of making a clear connection between experimental work and theory. I am also more grateful than you know for your support during fieldwork, and for lemon-ginger tea. Juliet, your ability to simultaneously grasp details, spot inconsistencies, understand implications, and put it all together, is unparalleled. Whether it was zooming out to interpret experimental results within a broader context, or trying to make sense of a pile of model results, your ability to reframe material and bring me back to the main point has been invaluable. Joseph, your knowledge of Andalusian Spanish and statistical analysis have been crucial for this dissertation. Thank you also for introducing me to a Sevillian to record stimuli, and for your support and encouragement throughout the process.

Thanks also to Laurel MacKenzie and Greg Guy for guiding me through variationist projects, coming to various talks I gave, asking questions, and pushing me to consider other perspectives on the topics at hand. Throughout the years, I have also learned so much from Renée Blake, Ailís Cournane, Chris Collins, Stephanie Harves, Alec Marantz, John Singler, Anna Szabolcsi, and Gary Thoms.

This work would not have been possible without the dozens of people who helped me create materials for the studies, and Sevillians who completed many seemingly odd perception tasks for me. Thanks to Julio López Otero and Deniz Özyıldız for recording stimuli (during the pandemic!), and to Paloma Jeretic, Sandy Abu El Adas, Suhail Matar, and Julien Dirani for helping translate experimental materials to Arabic and French. And thanks to a small army of Sevillian participants, who quite literally made this dissertation possible even though the pandemic canceled my data collection plans. I may never meet you, but thank you for your participation, for sharing my studies with your friends and family, and for your encouraging messages of support. Finally,

thanks to Marina Barrio and Mary Hely Cuenca of the phonetics lab at the Universidad de Sevilla for offering to host my work there, and for sharing my studies with their students when those plans fell through.

I chose to come to NYU in large part because of the community of grad students. You all did not disappoint. I leave with so many good memories of snack and share, dinners, work parties, game nights, happy hours, and bake-offs. One of the hardest parts of the pandemic was being separated from this fantastic community.

Thanks to my cohort, Hagen Blix, Maxime Tulling, and Sora Yin, as well to adopted cohort-mates Suhail Matar and Sandy Abu El Adas. Hagen, you are the reason I decided to learn LaTex. I'm sure everyone reading my dissertation will thank you for the improved legibility of the phonetic symbols. And I always enjoyed our conversations over happy hour, especially about German compounds. Sora, thanks for always bringing us to excellent mochi and Chinese food, and for being my phonology buddy as I got started. Maxime, your resilience and ability to find lightness in anything is stunning. I am convinced this is one of the most important qualities to have around to make it through a PhD. I know it was a rough, awkward start, but thank you for not giving up on me during the first semester. I'll always remember fondly the hours we spent discussing stimuli design, talking about kippies and smoothies, making wugs, going to find chocolate milk and cinnamon, wandering the city at night, doing French-Dutch Duolingo, and you speaking Dutch very loudly to make me understand (I did not). Thank you for being up for anything, and for making sure I was ok. Having you here made this journey so much more enjoyable, and I look forward to many more years of laughter and adventures.

Sandy and Suhail, I am so glad you allowed yourself to be adopted into our cohort! You have brought so much laughter, good food, kindness, and spice to my time here, and I can't imagine what these years would have looked like without you.

The rest of the cohorts are pretty great too. Kate Mooney, you are the hilarious phonological-banter-buddy I never knew I needed. I have so enjoyed our late-night phonology chats, cheese curds, joking about glib titles, trocheepants, and have rarely laughed so hard in my life. Chiara

Repetti-Ludlow, thanks for your support and friendship over my final semesters here. Thank you for emergency ramen, Italian, for being my Thai-at-11am buddy, and for understanding. Alicia Chatten, you always have the most surprising and delicious snacks. Thank you for lots of carrot cake, for helping me move, for game nights and gatherings, and for your friendship over the years. Thanks also to Guy Tabachnick, Alicia Mason, and Jai Pena, for interesting phonology-related discussions, and to Naomi Lee for being a fantastic office-mate and sharing many wonderful (and sometimes odd) conversations. I have also greatly enjoyed getting to know Omar Agha, Anna Alsop, Kim Baxter, Julien Dirani, Özlem Ergelen, Nigel Flower, Selikem Gotah, Ioana Grosu, Soo-Hwan Lee, Alicia Parrish, Sarah Phillips, Alex Warstadt, and Zhuoye Zhao. This year's group of first-years—José, Simone, Kaustubh, Cara, Alden, and Ross—has brought much-needed energy and excitement, and I look forward to seeing what you do in the coming years.

Thanks also to the graduate students who came before me and showed me (continue to show me!) the ropes, especially: Suzy Ahn, Isaac Bleaman, Dan Duncan, Zack Jaggers, Becky Laturnus, Yining Nie, Natalie Povilonis de Vilchez, Ildi Szabó, and Adina Williams. Thanks also to Sheng-Fu Wang, for hours working in the PEP Lab together, for being my weekend department buddy, and for your photos of people enjoying life, which always made my day (even when I was not particularly enjoying life). Special thanks to Mary Robinson. Mary, you took me under your wing when I first arrived, and taught me the ins and outs of linguistics and this department. I will always value memories of our weekly trips to Mille-Feuille, how much you've taught me about myself, and your support across an ocean.

Finally, thanks to Teresa Leung, Teresa Colaizzo, Star Jimenez, Seena Berg, and Hannah Katz. Without you, this department would literally not run. I am always grateful for your thoroughness, kindness, and patience.

All the people who have been with me since before my time at NYU and who are still here, I owe so much of this to you.

Although 2012 feels like a lifetime ago, that was when I met Marcos Rohena-Madrazo. If he hadn't let a random undergraduate into his class, and proceeded to invest an unfathomable

amount of time and energy into me, I would not be a linguist. Thank you for your faith that, as an undergrad with almost zero linguistics training, I could do it. Because of you, I did. Thank you for your piercingly-honest advice, sharing your own experience, asking the hard questions, and unconditional support, kindness, and thoughtfulness. You have been the mentor and friend everyone deserves, and few are lucky enough to have.

To my friends from Middlebury—Jimin Kang, Jen Melgar, Sarah Pollnow, Annie Pruitt, and Esther Steves—thank you for coming to visit me in NYC, for being curious (but not too curious) about what exactly I'm doing, for letting me crash at your apartments, and for making me feel at home wherever we are. Thanks also for online game nights during the pandemic that made me laugh so hard I cried, and for being the best travel buddies. I also owe much of my sanity during the last few years to Rachel Weissler and Carol Ready, who I first met while visiting graduate schools as a prospective student. I have so enjoyed—and needed—our long conversations about work, life, and the ups and downs. You both have taught me so much, and I am grateful that our shared experiences brought us together.

Thanks also to those who helped keep me sane through the PhD by being the best company in the pool and climbing walls. JP C. JP E., and Jessica, I will miss our random little swim group and 15-second chats between sets. I am grateful to have found pool buddies here even if I didn't seem thrilled at first. Special thanks to JP C. for always asking how I was doing, and caring about the answer. Thanks to my Vermont swim buddies as well—Emma Aspell and the entire UVRays team—for the water company and fun through summers and winters. Climbing was an unexpected addition to my life, and one I'm glad I let myself be talked into. Juliet and Ida, you are two of the best climbing buddies one could ask for. I will always levitate across a country or ocean to climb with you (without chalk, of course).

Before I even got to college, several people provided more understanding, support, and inspiration than I can describe. Martha, thank you for opening my mind to my own capabilities, encouraging me to do more than I knew I could, and for always being willing to listen. Nicola, I have never been as flawlessly on the same page with another person as I am with you. I have

always found comfort and inspiration in your quiet strength. Mary, my love of language is largely thanks to you. Thank you for fostering my interest, confidence, and sense of self. I don't know where I would be today if we hadn't crossed paths, but it certainly wouldn't be here.

They say a PhD is not only about the academic achievement, but about who you become in the process. They are right. LR, LG, KC, MJ, SJ, and RM, thank you for teaching me how to navigate life as a human being, and for putting up with me as I figure it out.

Finally, thanks to my family for always offering love and support, and for taking care of me when I come home. Elle, thank you for showing me that life doesn't have to be quite so serious. Mom and Dad, from you I learned many things, including stubbornness (I have been told this can be reframed more positively as 'determination'), which has been maybe the most important skill during my time here at NYU. More than you know, I appreciate that you have encouraged me to follow my interests, no matter how impractical they seem or where in the world they take me.

Abstract

This dissertation investigates the interaction between phonetic change, representational change, and contrast formation. Sound changes do not usually affect the number of contrasts in a language (Gurevich 2003); rarely, however, they result in a new contrast. This dissertation investigates the representational status of sounds in an ongoing change in Sevillian Spanish, metathesis in /s/-voiceless stop sequences: /sp st sk/ → [hp ht hk] → [ph th kh]. The resulting stop-h sequence could plausibly come to be represented as an aspirated stop, creating a new contrast, as argued by O'Neill (2009) and Gylfadottir (2015).

I present a series of production and perception experiments testing the representational status of stop-h forms and the interaction between metathesis and other parts of the phonological system. Results suggest that, although surface forms have changed, underlying representations have not. Based on a production study, I argue that stop-h sequences are derived from underlying /sC/ sequences and arise by gestural overlap. Results from a fill-in-the-blank perception task corroborate the /sC/ representation: Sevillian listeners map [Ch] sequences to underlying /sC/ clusters. Listeners of other dialects do not, indicating that the phonological grammar plays a crucial role in perception. Furthermore, results from a stress judgment task show that Sevillian listeners judge stress as if metathesis has not occurred. The interaction between metathesis is opaque—stress operates on a non-surface level of representation—and Sevillian listeners apply this opaque interaction to novel words.

Based on these results, I present a formal analysis in which the Sevillian change occurs in steps of coda /s/ reduction and, finally, metathesis. The steps in this analysis capture synchronic within- and between-dialect variation, and provide a mechanism to understand why listeners of other Spanish dialects do not perceive [h] in stop-h sequences. I propose that their grammars provide no direct path from /sC/ to [Ch], in either production or perception. Stress applies before metathesis, reflecting Sevillians' response patterns in the stress judgment task, where they judged stress patterns as if metathesis had not occurred.

Finally, results from a perceptual discrimination task suggest that [hC] → [Ch] metathesis is not perceptually optimizing, casting doubt on analyses that frame it as such (e.g. Flemming 1996; Cho 2012; Yoon 2012). For Sevillian, I investigate other explanations, including articulatory and structural accounts.

The results of these studies contribute to our understanding of the relationship between surface forms and underlying forms, specifically of how listeners map surface forms to underlying forms when there is a mismatch. Furthermore, the results raise interesting questions about what conditions are necessary and sufficient for sound changes to occur and for new categories to develop. The series of studies also provides a template for using perception experiments to diagnose underlying phonological representations, which are difficult to pin down, and paves the way for future work on the formation and stability of sound categories, the role of variation in phonologization, and learning.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

This dissertation investigates the interaction between sound change, phonological representations, and contrast formation. Sound changes tend to maintain existing contrasts in a phonological inventory, or, less frequently, neutralize them (Gurevich 2003). While some of these changes force reorganization of the phonological system (e.g. vowel chain shifts), very few changes create new segmental contrasts. I investigate one such case in Sevillian Spanish. This dissertation presents a series of production and perception experiments that test the underlying representation of segments in an ongoing sound change from /sC/ → [hC] → [Ch] in Sevillian /s/-voiceless stop sequences. I refer to this change as metathesis. The outcome of this change, [Ch], has the potential to result in a new series of contrastively aspirated stops /C^h/, as argued by O'Neill (2009) and Gylfadottir (2015).

The perception tasks test the representation of [Ch] sequences, the mapping between surface and underlying forms, the place of metathesis in the phonological grammar, and the role of perceptual factors in facilitating or causing metathesis. Results suggest that the sound change has not resulted in the creation of new segments or contrast in Sevillian Spanish, and that other phonological processes—like stress—treat [Ch] sequences as if they were still /sC/ clusters. I propose that several factors may conspire to prevent the creation of a new category: the presence of ex-

tensive phonetic variation and phonological alternations across word and morpheme boundaries. Finally, an ABX discrimination task suggests that perceptual factors are not likely to be the motivation behind the change. The results of all of the perception experiments are best explained by the role that phonological grammars of speakers' native languages play in shaping perception.

Based on the experimental results, I also develop serial analyses that capture listeners' behavior on the perception tasks. The experimental results show that Sevillian listeners map [Ch] to underlying /sC/ in perception, and that metathesis is invisible to stress. Sevillian listeners evaluate stress location as if metathesis had not occurred, suggesting that metathesis is derivationally late. My analysis thus treats Sevillian [Ch] sequences as deriving from /sC/ clusters that have undergone reduction and metathesis in several derivational steps, and forces stress to occur before metathesis. I argue that listeners' (both Sevillian and of other dialects) behavior on the perception tasks is directly linked to what their grammars produce: listeners whose grammars do not produce the forms presented as experimental stimuli either do not perceive them or do not accept them.

1.0 Sevillian Spanish metathesis

The change in Sevillian Spanish from *h-stop* to *stop-h* in <sp st sk> sequences is ongoing, and is illustrated in (1) (Torreira 2006; O'Neill 2010; Parrell 2012; Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016). Coda /s/ debuccalizes to [h], and [h] then metathesizes with the adjacent stop.

(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'

Both debuccalization and metathesis (1) are possible synchronic variants in Sevillian Spanish (alongside other variants; see Section 3.4.1.2). While many dialects of Spanish have h-stop sequences resulting from coda reduction in /sC/ clusters, several dialects of Spanish spoken in

southern Spain—including in Seville—are unique in having stop-h sequences. These sequences arise morpheme-internally and across word and morpheme boundaries: in some [Ch] sequences, [C] and [h] derive from different morphemes (Ruch 2008; Horn 2013).

Both articulatory and perceptual motivations for metathesis have been proposed. On the articulatory side, Parrell (2012) argues for Sevillian Spanish that metathesis is caused by gestural timing preferences that exert pressure for [C] and [h] to realign. He finds some experimental evidence supporting this hypothesis, although not for all speakers. On the perceptual side, several explanations have been posited for metathesis, attributing it to perceptual difficulties associated with some sequences. For example, Ruch and Harrington (2014) suggest, based on a perception study, that Spanish-speaking listeners may mishear the order of [h], which facilitates reinterpretation of [h] in the non-historical location. This argument is similar to one that has been made in other cases, where ambiguity is crucial for metathesis (e.g. Blevins and Garrett 1998; Hume 2004). Cross-linguistically, both metathesis in the direction of [hC] → [Ch] and the rarity of preaspirated stops (/hC/) have been attributed to perceptual factors. Metathesis has been argued to be driven by the need to optimize the perception of [h] by placing it after the stop rather than before it (e.g. Flemming 1996; Cho 2012; Yoon 2012). Preaspiration has been argued to be rare because [h] is not very perceptible in coda position (e.g. Bladon 1986; Silverman 2003). These claims about perceptual factors are mostly untested.

I test the perceptual optimization hypothesis in an ABX discrimination task in Chapter 5, and find no evidence that metathesis improves the perceptibility of [h], or that listeners are prone to mishearing the location of [h]. In Chapter 6, I discuss alternative accounts.

1.1 Contrast development and sound change

Sevillian [Ch] sequences are representationally ambiguous between being realizations of underlying aspirated stops (2a) and underlying /sC/ sequences (2b).

(2) Representational options for Sevillian [Ch] sequences

a. Segment analysis: [Ch] is a surface realization of an underlying aspirated stop /C^h/.

b. Cluster analysis: [Ch] is a surface realization of an underlying /sC/ cluster.

I argue for the cluster analysis throughout this dissertation. The segment analysis, however, raises interesting possibilities when considering types of attested phonological changes. Under the segment analysis, [Ch] would have phonologized into an aspirated stop /C^h/, creating a new contrast among the stop series in the language. Since the diachronic source of [Ch] sequences is an underlying /sC/ cluster, the new series of aspirated stops would be the result of metathesis of [hC] → [Ch], coalescence of those segments into a single segment, and reinterpretation of [h] as part of the stop. Aspirated stops would contrast with existing voiceless unaspirated and voiced stops in the inventory.

If Sevillian [Ch] were reanalyzed as a series of underlying aspirated stops, this would be unusual because it creates a new contrast, whereas most sound changes do not. Some changes, like mergers, neutralize contrasts. For example, many varieties of American English have merged the COT-CAUGHT vowels, so that the two are not contrastive (Labov 1994). Labov (1994: 313) points to Herzog's principle, which reflects the preponderance of contrast loss: 'mergers expand at the expense of distinctions.' Other changes avoid neutralizing contrasts: one member undergoes phonetic changes, but the contrast is maintained in another way (Gurevich 2003). Lenition/spirantization changes provide one example (Gurevich 2003). Many dialects of Spanish have intervocalic voicing and further lenition of voiceless stops /ptk/, which creates the possibility of contrast neutralization with voiced stops /bdɡ/, which spirantize in most positions. However, even when voiceless stops lenite, they are phonetically distinct from lenited voiced stops, even if not in the traditional parameters like voicing (Lewis 2001; Broś et al. 2021).

Some changes do result in the creation of new contrasts, but the conditions do not apply to the Sevillian case. Demergers (splits) can create 'new' contrasts when a previously merged sound category separates back into the two original categories, such as the reversal of the CARD-CORD (/ɑhr/ - /ɔhr/) merger in St. Louis (Labov 2010). In earlier work, demerger is considered to be

impossible (Garde's principle; Labov 1994: 311), because a word's membership in a phonological class is arbitrary: once words are all produced with the same Sound A, speakers (and learners) would not know which words to produce with Sound A, and which to restore Sound B to. However, some studies indicate that demergers can occur under the right circumstances. One proposal is that demerger can occur only if the merger was not complete (Labov 1994). Another is that demerger can occur through dialect contact, when merged speakers come into contact with speakers who maintain the distinction (Maguire et al. 2013; Nycz 2013; Johnson and Nycz 2015; Regan 2017; Gylfadottir 2018). Finally, orthography can support demerger (e.g. Regan 2017). None of these conditions applies to Sevillian: unaspirated voiceless stops are not the result of a merger of aspirated and unaspirated voiceless stops, Sevillian is not in contact with a language that contrasts voiceless aspirated and unaspirated stops, and the contrast is not supported by orthography.

The most likely path to the creation of a new contrastively aspirated stop category in Sevillian is through phonologization. Phonologization can be defined as the exaggeration of a phonetic feature beyond what is phonetically natural (Hyman 2013). Then, the feature is reinterpreted as belonging to a different source, creating a new contrast (Ohala 1981b). One of the clearest examples comes from the creation of contrastively nasal vowels. In many languages, including Romance languages, contrastively nasal vowels are hypothesized to have developed along the path in (3) (Sampson 1999; Beddor 2009).

(3) Hypothesized path of phonologization for nasal vowels in Romance
$[\mathrm{Vn}] \rightarrow [\tilde{\mathrm{V}}\mathrm{n}] \rightarrow [\tilde{\mathrm{V}}]$
$/\mathrm{Vn}/ \quad \rightarrow \quad /\tilde{\mathrm{V}}/$

The hypothesis is that the nasal consonant induced natural allophonic nasalization on the preceding vowel, the extent of this coarticulation increased, and eventually the nasal consonant deleted. Listeners would have reinterpreted the nasality as belonging to the vowel, creating a set of contrastively nasal vowels. This path to contrastively nasal vowels has been proposed for many languages.

The phonetic effects that are phonologized are often argued to originate from a combination of coarticulation and perception. One segment has coarticulatory effects on others, which creates ambiguity and opportunity for perceptual reanalysis. Listeners can mis-hear the location of a phonetic cue, because the features may be realized over a long domain (Ohala 1981b; Blevins and Garrett 1998; Ohala 1993). In recent experiments, Beddor (2009) finds evidence that a combination of coarticulation and perceptual mis-attribution is a plausible path to the phonologization of nasal vowels. In production, she finds that the duration of the nasal consonant and nasalization on the vowel are inversely correlated, so that as more coarticulatory nasalization is realized on the vowel, the duration of the nasal consonant decreases. In perception, she finds that listeners perceive nasality coming from a nasal and nasality on a vowel as equivalent sources of information. She argues that these factors indicate that listeners can attribute nasalization entirely to the vowel, resulting in contrastively nasal vowels.

Sevillian [Ch] sequences could plausibly phonologize into aspirated stops through a similar process of coarticulation and listener misperception (4). In Andalusian Spanish, as well as several other varieties of Spanish, Torreira (2006) found that the stop release is longer in /s/-voiceless stop productions than in intervocalic voiceless stop productions, which appears to be a natural coarticulatory effect.[1] In Andalusian varieties, this natural coarticulatory effect has been exaggerated (the motivation behind the exaggeration is irrelevant for now). At this point, hypothetically, exaggerated coarticulation could result in fully metathesized [Ch] sequences, evidence of the origin of metathesized [h] could become infrequent or disappear entirely, and listeners could attribute [h] as belonging to the stop itself, rather than to an underlying preceding /s/.

(4) Possible path of phonologization for Sevillian [Ch] sequences
[hC^{h}] → [hCh] → [Ch]
/sC/ → /C^{h}/

Although this path is not too different from that of contrastive nasalization, aspirated stops typologically do not usually arise in these ways. In a small survey (Section 6.3.2), I found that

[1] For /st/ sequences specifically, Ruch and Peters (2016: 28) make a similar argument. It is fairly clear that /st/ clusters lead the change. They suggest that a slightly naturally longer release in /st/ clusters may make them 'particularly prone to imitation and further lengthening.'

very few languages develop aspirated stops through reinterpretation of adjacent stops and [h] as a single segment, or through reinterpretation of coarticulatorily-derived aspirated as part of the stop. One potential case of coalescence comes from Indo-European, for which Rasmussen (1987) argues that one origin of aspirated stops was an independent /h/ phoneme that became adjacent to a voiceless stop. Ikalanga (Bantu) is a case where aspirated stops developed from the phonologization of a coarticulatory effect. For this language (and other Bantu languages), Mathangwane (1996: 208) argues that one source of aspirated stops was a coarticulatory effect that arose when stops were followed by high close vowels /i̧, u̧/. High, close vowels cause longer aspiration for aerodynamic and articulatory reasons (Ohala 1981a: 112), and Mathangwane hypothesizes that speakers reinterpreted this coarticulatory effect (long aspiration) as belonging to the stop.

Beyond the exaggeration of the coarticulatory effect, phonologization also requires listeners to misattribute that effect to a non-historical source. My series of perception experiments finds no evidence that this misattribution is not occurring in Sevillian.

1.2 Segmenthood

Sevillian metathesis presents an ideal opportunity to look at the interaction between sound change, contrast formation, and segmenthood. Metathesis creates a mismatch between surface forms and underlying representations, and the resulting stop-h sequences are of ambiguous representational status. Although they are historically and orthographically /sC/ sequences, they are similar to aspirated stops on the surface. As discussed in the previous section, the long release following the stop could be reinterpreted by listeners as belonging to the stop itself, resulting in phonologization and a new series of aspirated stops. This would create a three-way contrast between voiceless aspirated, voicelesss unaspirated, and voiced stops.

Distinguishing between complex segments and clusters has been the subject of much debate, and has been discussed for prenasalized stops, affricates, doubly articulated stops, pre- and postaspiration, among other sequences (Trubetzkoy 1939; Haugen 1958; Thráinsson 1978; Her-

bert 1986; Sagey 1986; Lombardi 1990; Riehl 2008). Many diagnostics have been used to try to make the distinction, including phonetics, how the sequences interact with syllable structure, phonotactics, and separability of the parts (e.g. Trubetzkoy 1939; Riehl 2008). Distinguishing between segments and clusters has received so much attention in part because it raises learning questions: when learners hear a sequence of sounds in their language, how do they determine whether it should be stored as a single contrastive unit, or as two independent units? Gouskova and Stanton (2021) take up this question using computational learning simulations, and argue that distributional properties of the components of the sequence are crucial for learning sequences as segments vs. clusters.

While there are many proposals that attempt to define segmenthood, they are difficult to test experimentally. Many experimental approaches compare the phonetics of the hypothesized complex segment to the phonetics of clusters in the language. While some studies find consistent differences between segments and clusters (e.g. Riehl 2008), others do not (e.g. Javanese, Adisasmito-Smith 2004; Russian vs. English, Shaw et al. 2021; see also discussion in Stanton 2017: 57-8). The phonetic effects on surrounding segments are also an unreliable diagnostic. Prenasalized stops have been found to affect preceding vowels in different languages cross-linguistically. In some Bantu languages, vowels preceding prenasalized stops are longer than those preceding regular stops (Herbert 1986: 134-139). However, in some Austronesian and Australian languages like Fijian, vowel duration is similar preceding prenasalized stops and other consonants (Ladefoged and Maddieson 1996: 123).[2] Other experimental work attempts to determine segment status by asking participants to perform tasks like reversing words (by phoneme or by syllable), and repeating words produced by model talkers (Berg 2001; Béland and Kolinsky 2005). These tasks, however, can be cognitively difficult for speakers and are influenced by orthography, and results are not always clear.

Regarding aspiration specifically, stop-h and h-stop sequences have been argued to be segments in some languages and clusters in others. For example, in English, tautosyllabic stop-h

[2]See Gouskova and Stanton (2021: 183-6) for a thorough overview of how acoustic differences are unreliable diagnostics for segmenthood.

sequences are usually considered to be aspirated stops, with [h] belonging to the stop (e.g. [p^hɑt], 'pot'). In other languages, however, tautosyllabic stop-h sequences have been analyzed as clusters. In Acehnese (Austronesian), for example, initial stop-h sequences are analyzed as clusters because the stop and [h] components can be split by an infix (e.g. *meu-k-eun-eu-hie*, from *khie* 'it tastes a little bit like cooking oil'; Asyik 1987: 16). In most languages, h-stop sequences have been argued to be clusters, including Spanish and English: [h] is an independent segment with relatively free distribution, and it is not tied to the following stop. In many fewer languages, [h] can be argued to be part of the stop itself (e.g. Scottish Gaelic, Icelandic; see Clayton 2010 for typological survey). I discuss these cases further in Chapter 3, Section 3.1.1.

Sevillian draws a line between phonological process and segment: a phonological operation operates on two independent segments (metathesis), and gives rise to a sequence that could be interpreted as a single segment. If stop-h sequences were to be reanalyzed as aspirated stops, stop-h sequences would be faithful realizations of the underlying form, rather than the result of a phonological operation. This case study allows us to investigate changes in phonetics, phonological representations, and the mapping between the two, at an intermediate stage of the change. Can the result of a phonological operation result in the formation of a new contrast? At what point does a bundle of phonetic characteristics come to be represented as a segment? And how can we test representational status?

1.3 This dissertation: Questions and results overview

The broad questions of the dissertation are the following:

- Is there phonetic evidence that Sevillian metathesis from [hC] to [Ch] occurs via gestural overlap?
- What is the role of the phonological grammar in mapping surface forms to underlying forms, and how can we observe this mapping experimentally?

- What is the role of intermediate derivational phonological representations, and what level of representation is visible to which processes?
- For a change like laryngeal metathesis, what is the role of universal vs. language-specific perception in facilitating, or causing, metathesis?

I probe these questions in a series of experiments. In a small production task (Chapter 2), I look for phonetic evidence that [Ch] sequences arise through a phonological process of gestural overlap of the /s/ and /C/ gestures. I investigate whether this change applies in all /sC/ sequences, or if it is limited to /s/-voiceless stop sequences. Then, I probe the representation of surface stop-h sequences in a fill-in-the-blank perception task (Chapter 3), by asking Spanish speakers of different varieties to map [Ch] to underlying forms in a morphologically-derived environment where the preceding segment could be /s/. In a stress judgment task (Chapter 4), I test how surface [Ch] forms interact with other phonological processes at a more abstract level in the Sevillian grammar by probing the interaction between metathesis and stress. Finally, in a cross-linguistic ABX perception task (Chapter 5), I test potential perceptual motivations behind the metathesis change.

In short, the results are as follows. There is phonetic evidence of gestural overlap in /s/-voiceless stop and /s/-voiced stop clusters. In /s/-voiceless stop clusters, overlap results in full metathesis; in /s/-voiced stop clusters, the result appears to be overlap without full metathesis. For both of these types of sequences, /s/ and /C/ are independent segments that are superimposed on top of each other. /s/-sonorant sequences show no evidence of overlap. In perception, the fill-in-the-blank task suggests that Sevillians map stop-h sequences onto underlying /sC/ sequences, and the stress judgment task suggests that other phonological processes operate as if metathesis has not occurred. Stress operates at an early stage of the derivation. Finally, the ABX discrimination task provides no evidence that metathesis from [hC] to [Ch] is perceptually optimizing, either at a universal level or for listeners of specific language backgrounds. Regardless of listeners' native languages, [h] is more difficult (or equally as difficult) to perceive *following* a stop than *before* it. Beyond this universal result, listeners' ability to perceive [h] in different locations can be explained by the way the experimental stimuli map onto listeners' native language phonological

categories. One surprising result is that Sevillians do not differ from other Spanish speakers in this task. Their experience with surface [Ch] sequences does not aid them in perceiving these sequences. In comparison with results from the other tasks, this suggests that Sevillians only perceive [h] in [Ch] sequences when [h] can reasonably be derived from /s/.

The role of the phonological grammar is evident in the results of all three perception tasks. In the fill-in-the-blank task, the phonological grammar helps Sevillians 'undo' metathesis to arrive at an underlying /sC/ representation. In the stress judgment task, the phonological grammar determines how metathesis and stress interact: stress operates on a level of representation where metathesis has not occurred. And in the ABX task, most results can be explained by listeners' native language phonological categories, and by how forms presented in the experiment map onto them.

Based on these results, I present formal analysis of metathesis and stress. The analysis is cast in Harmonic Serialism (Prince and Smolensky 1993), which allows me to account for listeners' behavior in the perception experiments. In Chapter 3, the serial analysis treats the path from /sC/ $\rightarrow$ [Ch] as consisting of several steps of coda reduction and metathesis. This derivational path corresponds to Sevillian listeners' behavior in the perception task, where they mapped surface [Ch] to underlying /sC/: their grammars provide a direct path between the forms in both production and perception. In contrast, Mexican and Argentinian listeners did *not* map [Ch] to /sC/ in the perception task. I argue that these listeners' phonological grammars differ, so they have no direct path between /sC/ and [Ch]. When presented with surface [Ch], they are unable to 'undo' the steps required to arrive at underlying /sC/. In Chapter 4, Sevillian listeners evaluate stress as if metathesis had not applied. The serial analysis captures this behavior by forcing stress to occur before metathesis, so that stress operates on an intermediate representation where /s/ has not yet metathesized out of coda position. Together, these analyses captures the experimental results in that Sevillians 'undo' the steps of the derivation to arrive at the intended /sC/ representation, and in that metathesis occurs derivationally late.

1.4 Terminology

Before continuing, I clarify the terminology used in this dissertation, some of which differs substantially from previous work on Sevillian Spanish. I have chosen different terminology to be clear about my assumptions about representations, which are central to the dissertation.

First, I refer to the Sevillian change as *metathesis*, following Ruch (2013). My experimental results suggest that speakers represent [C] and [h] as separate segments; thus, the change from [hC] to [Ch] is a change in location of two elements. Almost all previous studies on Sevillian refers to this process as a change from preaspiration to postaspiration. Furthermore, previous studies refer to the duration of the stop release in [Ch] sequences as voice onset time (VOT). I avoid all of these terms because they imply a phonological representation in which [h] belongs to the stop, which is the opposite of my central claim. I use the terms *preaspiration* and *postaspiration* to refer only to cases where [h] can be considered to be representationally part of the stop (e.g. Icelandic, Haugen 1958),[3] and *VOT* to refer only to aspiration on contrastively aspirated stops (e.g. English). Where VOT is not part of a contrastively aspirated stop, I simply call it a *stop release*.

I also make a distinction between segments, clusters, and sequences. By *segment*, I mean a single, contrastive phonological unit; visually, I represent these as /C^{h}/, [C^{h}]. By *cluster*, I mean two adjacent segments; I represent these as e.g. /sC/, [hc], [Ch]. By *sequence*, I mean two adjacent sounds without a specified phonological representation (following Riehl 2008). I use *sequence* when referring to [Ch], without wanting (or needing) to imply that it is decidedly a single segment or a cluster.

Finally, within Andalusian Spanish, there are two main varieties: Western Andalusian Spanish (WAS) and Eastern Andalusian Spanish (EAS). Seville is a variety of WAS. The two dialect regions differ substantially. Most relevant for the current study, WAS varieties are undergoing the metathesis change, while EAS varieties do not have robust metathesis, although they may be moving in this direction (Ruch and Harrington 2014; see Herrero de Haro 2017 for overview

[3]Other scholars have argued that [h] should be treated independently from the stop in Icelandic (e.g. Thráinsson 1978). See Heimisdóttir (2015) for an overview of previous accounts.

of EAS properties). Much of the existing literature specifies only a broad dialect region (WAS or EAS), even though there is extensive variation even within these regions. In order not to misrepresent claims about certain varieties, or add a level of specificity not intended, I use the terms used by the authors when discussing their findings. I use the broad term *Andalusian Spanish* only when necessary to refer to properties that hold of the entire dialect region, including both Eastern and Western Andalusian, and when the original sources do not specify region. I use *Sevillian Spanish* to refer to the Spanish spoken in Seville, which is where my participants come from.

1.5 Roadmap

This dissertation is organized as follows.

Chapter 2 provides a more thorough description of Sevillian metathesis, including previous experimental and theoretical work that documents the change and proposes different motivations for it. In particular, I give an overview of the articulatory description of the change, most clearly laid out by Parrell (2012) and Torreira (2006), which predicts specific phonetic effects in these clusters. Then, I present results from a small production study on /sC/ clusters of different types, providing the phonetic basis for treating metathesis as gestural overlap.

Chapter 3 tests the underlying representation of Sevillian [Ch] sequences in a morphologically derived environment. In this environment, there is an explicit possibility that [h] in [Ch] derives from an underlying /s/ on the preceding verb. The experiment manipulates the duration of [h] and asks listeners to decide which verb form precedes [Ch]. Sevillian listeners map [h] to an underlying preceding /s/, suggesting that [Ch] is a sequence, not a single segment. Listeners of other Spanish dialects (Mexican and Argentinian) do not perform this mapping. In the formal analysis, I propose that this difference in perception is because the dialects have different phonological grammars that map surface forms to underlying representations differently. The Sevillian grammar is able to 'undo' metathesis through various steps, to arrive at a /sC/ representation, while other dialects' grammars do not provide a path for this mapping.

Building on the finding in Chapter 3 that Sevillians decompose [Ch] sequences into underlying /sC/ clusters, Chapter 4 tests how metathesis interacts with a higher-level phonological process, stress. Results from a stress judgment task show that Sevillian listeners treat surface-open syllables derived from /sC/ clusters ([CV.CHV]) the same as surface-closed syllables ([CVC.CV]). In the formal analysis, metathesis is invisible to stress because it operates early in the derivation, on a representation where metathesis has not yet occurred.

It is sometimes suggested that metathesis happens for perceptual reasons. [h] is argued to be more perceptible following a stop than before it, and may be more confusable around some consonants than others. Chapter 5 tests this explanation by probing listeners' accuracy in determining (a) the presence vs. absence of [h] in different locations, and (b) the linear order of [h]. Results provide no evidence of a perceptual motivation for metathesis, but rather show that listeners' perception of [h] is based mostly on how they map [h] to their language's phonological categories. Additionally, results show that Sevillians pattern with other Spanish speakers: their experience with [Ch] sequences is context-dependent, and is not applied to environments where [h] has no plausible source. Perception is filtered through phonological categories, and [Ch] is not a phonological category for any Spanish speakers.

Chapter 6 ties together the experimental results in a broader discussion that aims to raise questions for future research. I compare Sevillian metathesis-via-overlap to other metathesis patterns that have been argued to occur by overlap, and call for more studies on the phonetics of metathesized sequences. I also lay out possible articulatory and structural accounts for Sevillian metathesis, given that perceptual optimization (tested in my ABX task) does not seem to be the cause. Then, I discuss Sevillian metathesis in light of the literature on phonologization, comparing it to the develoment of contrastive aspiration in other languages. While a more thorough comparison is warranted, my small survey suggests that the Sevillian path will likely not lead to contrastive aspiration. Finally, I compare stress-metathesis interactions to stress-epenthesis, which have received much more attention. I suggest that there may be an asymmetry between metathesis and vowel insertion, where metathesis is cross-linguistically limited to being derivationally late,

while vowel insertion can be either early or late. This comparison is speculative, and invites deeper investigation of the architecture of the phonological grammar.

Chapter 7 concludes.

CHAPTER 2

Phonetics of Sevillian /sC/ sequences

2.0 Introduction

This chapter presents data from a small production study on Sevillian Spanish clusters of /s/-voiceless stop, /s/-voiced stop, and /s/-sonorant. While many previous studies look at the acoustics of /s/-voiceless stop sequences, fewer look at clusters of coda /s/ with other consonants. Investigating the acoustics helps establish the mechanism behind how the change is implemented, the reach and progression of the change, and what acoustic cues might be available to learners as they determine the segmental status of [Ch] sequences. If learners are exposed to [Ch] forms and must decide if these forms are segments or metathesized versions of clusters, what phonetic information do they have?

In this chapter, I review the reduction processes leading to metathesis, and findings from previous studies on the phonetics of Sevillian /sC/ sequences (Section 2.1). Then, I present the methods (Section 2.2) and results (Section 2.3) of the production study. Looking at these results together (Section 2.4), I suggest that acoustics provide evidence that metathesis in /s/-voiceless stop clusters occurs via overlap, and that a similar process may apply in /s/-voiced stop clusters

too, although the phonetic outcome is different. There is no evidence for a similar process in /s/-sonorant clusters. I also suggest that stop-h sequences appear to be integrating into the broader phonological system, as they are now partially subject to intervocalic voicing, a process which typically applies to intervocalic voiceless stops. [Ch] sequences also now arise in a novel context, /pt kt/ sequences (e.g. /akto/ → [at(ː)ho]), suggesting that the change is expanding.

2.1 Coda /s/ reduction and Sevillian metathesis

Coda /s/ reduction is one of the most common and well-studied processes across varieties of Spanish, and Sevillian metathesis can be conceptualized as part of this continuum. Many varieties debuccalize coda /s/ to [h], or delete it entirely (Hammond 2001). Debuccalization represents one stage of lenition, since the oral gesture of /s/ is lost. Deletion represents the final stage, since all gestures have been lost (Mason 1994). Variation in coda /s/ production is extensive within and across communities, and is conditioned by linguistic, extra-linguistic, and social factors, including surrounding phonological environment, position in the word, morphological status, speech rate, speaker demographics (e.g. sex, age, socioeconomic status, education), and speech style. See Erker (2012) for a thorough overview of how linguistic and social factors condition coda /s/ weakening in Spanish.

In addition to /s/, other coda obstruents also undergo reduction in many varieties of Spanish (Campos-Astorkiza 2012). This synchronic reduction has been argued to reflect the diachronic tendency in modern Spanish to move towards open syllables (Malmberg 1965; Catalán 1971). Moya Corral (2007) and Moya Corral and Tejada Giráldez (2020: 205) argue that Sevillian metathesis can be treated as part of the reduction continuum because it changes the syllable structure, moving phonetic material related to /s/ out of the coda ([tʃis.pa] (HL) → [tʃi.pha] (LL) 'spark') (see Section 4.1.1 for justification of this syllabification). Gemination, which diachronically preceded metathesis and often co-occurs with metathesis, can also be seen as a way to improve on

coda restrictions: it maintains some information about /s/, while creating an open syllable (Morris 2000).

2.1.1 Properties of Sevillian metathesis

Metathesis in varieties of Western Andalusian Spanish, including Sevillian, has been widely reported in /s/-voiceless stop clusters (e.g. Torreira 2006; Ruch 2008; O'Neill 2010; Parrell 2012; Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016).

Metathesis can occur in any /s/-voiceless stop sequence, regardless of other phonological properties of the environment. Metathesis occurs in all stress environments: when the syllable containing /s/ is stressed (5a), unstressed and followed by a stressed syllable (5b), or surrounded by unstressed syllables (5c) (Horn 2013; Torreira 2012).

(5) Sevillian metathesis occurs in all stress environments

a.	/ˈpasta/	[ˈpatha]	'pasta'	Ví̍s.CV
b.	/amisˈtad/	[amiˈthað]	'friendship'	Vs.CV́
c.	/distinˈθjon/	[dithinˈθjon]	'distinction'	Vs.CV.CV́

Metathesis also applies in all morphological contexts (Ruch 2008; Horn 2013). In addition to morpheme-internally (6a), metathesis occurs across word-internal morpheme boundaries (6b) and across word boundaries (6c).

(6) Metathesis occurs in all morphological environments

a.	/ˈpasta/	[ˈpatha]	'pasta'	Morpheme-internal
b.	/des-taˈpar/	[dethaˈpar]	'to uncover'	Morpheme boundary
c.	/ˈmas ˈtorta/	[ˈma ˈthorta]	'more cake'	Word boundary

Sevillian metathesis in /s/-voiceless stop sequences is an ongoing change in progress. Ruch and Harrington (2014) is the first study to document the progression of the change in Western Andalusian by looking at social patterning of realizations of /st/ sequences. In WAS, /s/-voiceless stop sequences have longer post-posed [h] than intervocalic voiceless stops, and the social patterns show a change in progress. Younger speakers show advanced metathesis, with short pre-posed

[h] and long post-posed [h]. In contrast, older speakers produce /s/-voiceless stop sequences with longer pre-posed [h] and shorter post-posed [h] than younger speakers, suggesting that metathesis is not as advanced for them. In subsequent work, Ruch and Peters (2016) expand the scope to /sp/, /st/, and /sk/ sequences, and find that pre-posed [h] is shortening and post-posed [h] is lengthening for young WAS speakers in all of the sequences. Older WAS speakers, in contrast, have longer stop releases for /st/ than /t/, but not for /sk/-k/ or /sp/-/p/. The change may have started with /st/ sequences, a point I return to in Section 2.4.2.

While the generational patterns in Western Andalusian Spanish clearly point to a change in progress, the influence of other language-external factors is less straightforward. Horn (2013) finds that no social factors predict the presence or duration of post-posed [h], although it is numerically more frequent and longer for younger women with university or trade school levels of education. Ruch (2008) does find an effect of speech style: [Ch] forms are frequent in conversational and read speech, but less common in word lists, as would be expected for a variant that is not entirely standard.

A further progression of the change has also been reported in Seville and other cities in Western Andalusia. Metathesis in /st/ sequences has progressed from a simple reordering ([Ch] $\rightarrow$ [th]) to *postaffrication*, which involves fortition of [h] to [s] ([th] $\rightarrow$ [ts]) (Moya Corral 2007; Ruch 2008; Ruch 2010; Vida Castro 2016; Del Saz 2019; Moya Corral and Tejada Giráldez 2020). The change towards affricated variants shows clearer social stratification (level of formality, level of education) than metathesis in cities where it is reported (Ruch 2010; Moya Corral 2007; Vida Castro 2016; Moya Corral and Tejada Giráldez 2020), possibly because it is newer.

Because /s/-voiceless stop sequences are undergoing change, there are many phonetic variants present in the communities, which I discuss in the next section.

2.1.2 Previous phonetic studies

2.1.2.1 /s/-Voiceless stop sequences

Sevillian /s/-voiceless stop sequences have received relatively robust attention, and have been found to differ phonetically from intervocalic voiceless stops. The main cues of /s/-voiceless stop sequences reported in the literature are coda [h] (debuccalization: [hC]), a long stop release (metathesis: [Ch]), and lengthening of the stop following /s/ (gemination: [Cː]).

Some variants of Sevillian /sC/ are shown in (7) (adapted from Ruch 2008: 33-4; Ruch and Harrington 2014). Parentheses indicate that the cue is sometimes present, and sometimes not. Variants include retention of coda /s/ as [s], debuccalization of /s/ to [h], gemination, deletion, and metathesis. These cues are variably realized and can co-occur. Specifically, note that [h] can be split across the stop, as in (c).

(7) Some variants of Sevillian <sC> clusters (/pasta/ 'pasta')

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.)

The main acoustic property of the realization of /s/-voiceless stop sequences in Sevillian and other Western Andalusian varieties is long stop releases. Stop releases in Western Andalusian varieties range from 40-60ms, in comparison to intervocalic voiceless stops in these same varieties, which have short releases of 15-30ms, depending on the place of articulation of the consonant (Torreira 2006; Torreira 2007; O'Neill 2010; Ruch and Harrington 2014; Ruch and Peters 2016; Galvano and Henriksen 2021).

The presence and duration of the stop release in Sevillian /s/-voiceless stop sequences is also conditioned by the place of articulation of the stop, the morphological context, and possibly stress.

Cross-linguistically, the VOT of voiceless stops follows a cline based on place of articulation, illustrated in Table 2.1a. Velar stops have the longest VOT, followed by alveolar and bilabial

stops, a pattern usually proposed to be due to how aerodynamics relates to place of articulation (Cho and Ladefoged 1999). In Peninsular Spanish (specifically, Madrid), Rosner et al. (2000) found the same cline, although although labial and dental stops have similar length VOTs (/k/ = 26.5ms, /t/ = 14.0ms, /p/ = 13.1ms) (Table 2.1b).

Table 2.1: VOT/Stop release clines cross-linguistically and in varieties of Spanish

		Intervocalic	/sC/
a.	Cross-linguistic cline	$k^h > t^h > p^h$	—
b.	Madrid Spanish (Rosner et al. 2000)		—
c.	Younger Sevillian Spanish (Ruch and Peters 2016)		$k^h \sim t^h > p^h$
d.	Older Sevillian Spanish (Ruch and Peters 2016)		$k^h > t^h > p^h$

According to data in Ruch and Peters (2016: 14), the durations of intervocalic voiceless stop releases appear to follow this cline (Table 2.1c, d). Data from their same study suggest that durations of stop releases in [Ch] sequences deriving from /sC/ follow the expected cline for older speakers. However, for younger speakers, the group most advanced in the change, they do not: the stop release of /st/ and /sk/ sequences is the same (around 50ms). Galvano and Henriksen (2021) and Horn (2013) also found that /st/ and /sk/ had approximately equal length releases, and Horn (2013) found that the *presence* of metathesis is more frequent with /st/ and /sk/ clusters as compared to /sp/ clusters. However, it is not actually clear whether long stop releases deriving from metathesis would be expected to follow the universal cline for voiceless stops, since the stop release comes from leftover gestures from another segment. I discuss this further in Chapter 3, Section 3.2.1.

Metathesis occurs in all morphological contexts and in all stress environments, as mentioned in Section 2.1.1. The duration of stop releases has been found to be longer morpheme-internally than across word boundaries (/es'poso/ 'spouse' > /los 'peros/ 'the dogs') (Horn 2013). The results for stress are mixed: Torreira (2012) and Ruch (2008) do not find an effect of stress on duration of the stop release, while Horn (2013) finds that stop releases are longer when the preceding vowel is stressed (['gatha] '3SG spends' > [ga'tho] '3SG spent'). Regardless of the differences

in phonetic detail, what matters is that metathesis occurs across word boundaries: my perception studies rely on this fact. Furthermore, because the effect of stress is unclear, my studies match stimuli items for stress location.

Another cue to /sC/ sequences is gemination of the consonant following /s/, as in (7b), (7d), and (7f). Gemination has been widely reported in both Eastern Andalusian Spanish (e.g. Gerfen 2001; Gerfen 2002; Campos-Astorkiza 2003; Bishop 2007) and Western Andalusian Spanish (e.g. Torreira 2006; Ruch and Harrington 2014).

One point that will be crucial for my analysis of metathesis in Chapter 3 is that gemination appears to have arisen before metathesis and, as metathesis progresses, there appears to be degemination. Ruch and Harrington (2014: 20-1, 24) find that the young Western Andalusian speakers leading the change have a strong negative correlation between the duration of the following consonant and the duration of post-posed [h]. As post-posed [h] lengthens, the duration of the preceding consonant shortens. Older Western Andalusian speakers, who are not as advanced in the metathesis change, do have long closures, but do not show this correlation. Dialectal differences also suggest that gemination occurs before, and somewhat independently of, metathesis. As already mentioned, Eastern Andalusian Spanish has gemination of the consonant following /s/ (Gerfen 2002; Campos-Astorkiza 2003; O'Neill 2010), and incipient signs of metathesis among younger speakers (Ruch and Peters 2016). O'Neill (2010) compares the two dialects directly and finds that Eastern Andalusian speakers have longer stop closures and shorter stop releases, while Western Andalusian speakers have shorter closures and longer releases. Because Eastern Andalusian has gemination but limited metathesis, Ruch and Harrington (2014) deduce that Eastern Andalusian represents an early stage of the change, and that gemination also preceded metathesis in Western Andalusian varieties. Additionally, early studies of Western Andalusian Spanish describe long closure duration for the consonant following /s/, before metathesis was reported (Alvar 1955: 291), and other varieties of Spanish have gemination but not metathesis (e.g. e.g. Puerto Rico, Galarza et al. 2014; Cuba, Terrell 1979). Ruch and Harrington (2014: 24) conclude that debuccalization and gemina-

tion preceded metathesis, and that once metathesis is established, degemination begins to occur. Together, these findings suggest that gemination and metathesis are somewhat independent.

Less investigated acoustic cues to Western Andalusian /s/-voiceless stop sequences include breathiness on the preceding vowel (Torreira 2007; O'Neill 2010; Torreira 2012), resistance to intervocalic voicing (O'Neill 2010; Galvano and Henriksen 2021), and spectral properties of the noise ([Ch̲] or [Cs̲]) following the stop release (Moya Corral 2007; Gylfadottir 2015; Ruch and Peters 2016; Vida Castro 2016; Del Saz 2019).[1] While some of these findings are based on acoustic analyses, others appear to be based on impressionistic observations. The resistance to intervocalic voicing is particularly relevant for the current study.

In sum, previous work has established that Sevillian /s/-voiceless stop clusters differ phonetically from intervocalic voiceless stops. /s/-voiceless stop clusters are realized variably, with long stop releases and gemination of the following consonant as the most robust properties. The duration of the stop release has been found to vary by place of articulation of the voiceless stop, stress, and the presence of a word boundary. Furthermore, gemination has been argued to be a precursor to metathesis, and degemination is an ongoing change once metathesis has been established.

While the phonetic properties of /s/-voiceless stop sequences are relatively well-investigated, those of other sequences have received much less attention. I discuss reported properties of those sequences in the following sections.

2.1.2.2 /s/-Voiced stop sequences

/s/-voiced stop sequences in Andalusian Spanish have been reported to differ phonetically from intervocalic /bdɡ/. In almost all varieties of Spanish, intervocalic voiced stops undergo spirantization in most phonological contexts (/bdɡ/ → [βðɣ]). The voiced stops /bdɡ/ are only realized as stops [bdɡ] when they follow homorganic nasals/liquids or a pause (Table 2.2).

[1] 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. Herrero de Haro 2016; Henriksen 2017; see Herrero de Haro 2017 for overview).

Table 2.2: Spirantization of voiced stops in Spanish

Realization	Context			
Spirantized	Intervocalic	/bodeɡa/	[boðeɣa]	'winery'
	Following non-homorganic nasal/liquid	/alba/	[alβa]	'dawn'
		/arde/	[arðe]	'3SG burns'
		/alɡo/	[alɣo]	'something'
Stop	Following homorganic nasal/liquid	/ambos/	[ambos]	'both'
		/aldea/	[aldea]	'town'
		/ˈangulo/	[angulo]	'angle'
	Post-pausal	/bino/	[bino]	'wine'

In the contexts where it occurs, spirantization is obligatory, but the degree of constriction is conditioned by factors such as regional dialect (Carrasco et al. 2012; Butera 2018), stress (Cole et al. 1999; Ortega-Llebaria 2004), and surrounding vowels (Cole et al. 1999; Ortega-Llebaria 2004; Simonet et al. 2012). [βðɣ] are usually considered to be approximants, as opposed to fricatives.

The pronunciation of /s/-voiced stop clusters in varieties of Western Andalusian Spanish is reported to be highly variable (Alvar 1955; Romero 1995). Generally, they are produced as a 'single segment with the same point of articulation as the spirantized stop' (8) (Romero 1995: 132).

(8) /s/-voiced stop sequences in Andalusian Spanish (adapted from Martinez-Gil 2012: 122)

a.	/resbalar/	[refːalar] ∼ [reɸːalar]	'to slip, slide'
b.	/desde/	[deθːe]	'since'
c.	/musɡo/	[muxːo]	'moss'

As compared to spirantized intervocalic voiced stops, the coalesced productions resulting from /sb sd sɡ/ sequences are consistently longer (geminated) (Alvar 1955; Romero 1995; Hualde 1989a; Martinez-Gil 2012). They have also been reported to be realized more as fricatives than approximants—that is, with higher constriction and more noise (Romero 1994; cited in Romero

1995: 10-11)[2]—and partially or fully voiceless (Alvar 1955, Hualde 1989a; Martinez-Gil 2012). See also Morris (2000) (and references therein) for an overview of how /s/-voiced stop sequences behave in different Peninsular Spanish varieties.

2.1.2.3 /s/-Sonorant sequences

Like /s/ clusters with voiceless and voiced stops, /s/-sonorant clusters are reported to have gemination of the following consonant (Gerfen 2001; Martinez-Gil 2012; Morris 2000 and references therein). Martinez-Gil (2012) also reports that the first half of the geminate may be voiceless (9).

(9) /s/-sonorant sequences in Andalusian Spanish (adapted from Martinez-Gil 2012: 121)

a. /isla/	[ilːa]	∼	[il̥la]	'island'
b. /mismo/	[mimːo]	∼	[mim̥mo]	'same'
c. /asno/	[anːo]	∼	[an̥no]	'donkey'

2.1.2.4 Voiceless stop-Voiceless stop sequences

Other sequences involving sequences of coda obstruents and /t/ may undergo metathesis as well. Gerfen (2001) gives several examples of coda obstruents debuccalizing and causing gemination of the following consonant in Eastern Andalusian Spanish (10).

(10) Debuccalization and gemination in /pt kt/ clusters in Eastern Andalusian (from Gerfen 2001: 194)

a. /apto/	[ahtːo]	'apt'
b. /piθka/	[pihkːa]	'pinch, small amount'
c. /akθjon/	[ahθːjon]	'action'
d. /obtuso/	[ohtːuso]	'obtuse'

Although metathesis has not been experimentally verified for these sequences, it would not be unexpected in Sevillian. Like coda /s/, these coda obstruents debuccalize and cause gemination of the following consonant. Metathesis could apply, just as it does with [h] resulting from coda

[2]Romero (1995) argues, however, that differences in constriction and noise between intervocalic spirantized voiced stops and /sb sɡ sɡ/ is largely because the clusters are simply longer.

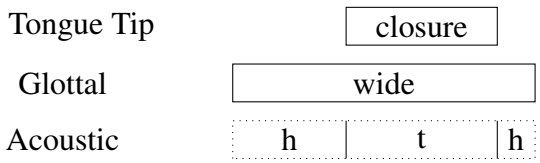
/s/. Consonant sequences with coda obstruents are infrequent in the Spanish lexicon, and my study includes the two most frequent sequences, /pt kt/. I look at these sequences because they could represent an expansion of the metathesis change to a new context, with coda stops.

2.1.3 Articulatory descriptions and acoustic trade-offs

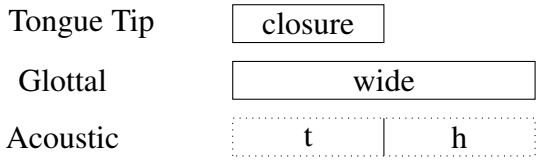
Most descriptions of Sevillian metathesis cast it in articulatory terms. Metathesis can be described as a change in the gestural coordination between the closure gesture for /ptk/ and the spread glottis gesture for [h], as most clearly proposed by Parrell (2012) (see also Torreira 2006). In [hC] sequences, the glottal gesture for [h] starts well before the beginning of the closure gesture and extends just past the closure release (11a). In [Ch] sequences, in contrast, the glottal gesture and stop closure gesture begin at the same time, and the glottal gesture extends well past the stop release (11b). Only the relative timing of gestures changes; their durations stay the same. Finally, intermediate [hCh] forms look like (11c): the [h] is split around the stop acoustically, because the closure gesture has not fully realigned.

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

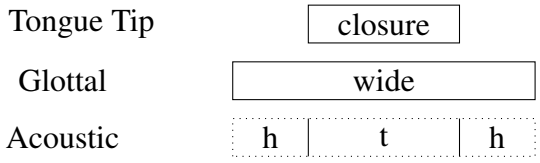
a. Sample gestural score for [ht] sequence



b. Sample gestural score for [th] sequence



c. Sample gestural score for [hth] sequence



A gestural account of Sevillian metathesis makes predictions about the phonetics of these sequences. Phonetically, /s/-voiceless stop sequences should have long stop releases (as already found in previous studies) and may also resist passive voicing. Many varieties of Spanish have extensive voicing of intervocalic /ptk/ (e.g. Torreira and Ernestus 2011). Although /sC/ sequences in Sevillian appear to be intervocalic when they surface as metathesized [VChV], the glottal gesture that overlaps the stop closure, as in (11b), may prevent passive voicing.

For /s/-voiced stop and /s/-sonorant sequences, this same kind of gestural realignment and overlap could predict (a) voicelessness or (b) noise during the constriction phase. Voiced stops spirantize in this context, so there is no release on which to 'dock' [h], unlike voiceless stops. In line with gestural predictions, devoicing has been reported for both types of clusters, and increased constriction and noise have been reported for /s/-voiced stop clusters (see Sections 2.1.2.2 and 2.1.2.3).

The articulatory description also makes predictions about what cues should be durationally correlated with each other, which I discuss in the next section.

2.1.3.1 Cue trade-offs

The cues of coda [h], long stop release, and gemination are variable, can co-occur, and are sometimes found to show trading relationships (e.g. Ruch and Harrington 2014; Parrell 2012; Torreira 2012). The articulatory account makes predictions about how the durations of cues should correlate with each other. For example (assuming that the length of gestures is stable), there should be a correlation between the duration of pre-posed [h] and post-posed [h], since the stop closure gesture simply slides earlier. Ruch and Harrington (2014) find no evidence for this correlation. Parrell (2012) does find evidence for this correlation, as does Cronenberg et al. (2020), using a method that determines gestural length in more abstract terms (instead of non-overlapping, linear order on a spectrogram).

Correlations between other cues are not predicted or explained by the articulatory model. One sticking point is closure duration: the articulatory account of metathesis does not explain

gemination, or how gemination interacts with stop release duration, because the gestures simply realign in relation to each other without changing duration. As discussed in Section 2.1.2.1, it is important to note that gemination and metathesis are independent processes. Gemination can be analyzed as compensatory lengthening for /s/ reduction (Campos-Astorkiza 2003) that occurs before metathesis (Ruch and Harrington 2014). Gemination—without metathesis—has been reported in other varieties of Spanish as well (e.g. Puerto Rico, Galarza et al. 2014; Cuba, Terrell 1979). In WAS, studies consistently find a negative correlation between closure duration and stop release: shorter closures occur with longer releases (O'Neill 2010; Parrell 2012; Torreira 2012; Ruch and Harrington 2014). Simple gestural realignment does not account for this acoustic trade-off.

One possible explanation for the negative correlation between stop closure and release duration is that, as release duration becomes a more salient cue to /s/, gemination is no longer needed, and the consonant shortens (Ruch and Harrington 2014). Where gemination of the following consonant used to be the main cue to the /C/-/sC/ contrast (as is still the case in EAS), long releases may be taking over this role in WAS. This would mirror other phonologization changes, like the development of contrastive nasal vowels in many languages: as nasalization on the vowel increased, the nasal consonant dropped away (Beddor 2009).

Gemination is also related to stop release length and pre/postaspiration, and may create the conditions for metathesis in Sevillian (I return to this point in Chapter 3). Synchronically, in languages with preaspirated stops, these are often realizations of a fortis or geminate stop and can alternate with geminates (Silverman 2003). Even in languages where [h] optionally precedes a stop but is not phonological preaspiration, there is still a connection with geminates. For example, in Sienese Italian, [h] was reported preceding geminates /pː tː kː/ (Stevens 2010); a further change appears to result in [h] surfacing *after* these geminates too (Stevens and Hajek 2010). Diachronically, Clayton (2010) argues that, in most cases, preaspirated stops actually developed from voiceless geminates.

In short, some of the trade-offs predicted by the articulatory account are found only sporadically, and other trade-offs are not explainable if the only process is gestural realignment. In

this chapter, I focus on a single cue-trading relationship—that between closure duration and stop release duration—as informative about progression of the change.

2.1.3.2 The articulatory account as description vs. motivation

There is a crucial distinction between the *mechanism* of the change vs. the *motivation* behind the metathesis change. The process described in this section—realignment of gestures—is the mechanism. The *motivation* is a separate question entirely.

Some scholars have argued that the motivation is also articulatory (Torreira 2006; Parrell 2012). In contrast, based on inconsistencies between the articulatory account and previous experimental results, others have argued that the motivation is, at least in part, perceptual (Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016). I return to the issue of articulatory vs. perceptual motivations in 5.1, and the ABX discrimination task in Chapter 5 tests for certain perceptual motivations. The results provide no evidence that metathesis is perceptually motivated or facilitated, and Section 6.2 in Chapter 6 includes a more thorough discussion of the articulatory motivation proposed by Parrell (2012), as well as other potential motivations.

2.2 Production study

2.2.1 Questions and hypotheses

The account of metathesis as gestural overlap makes predictions about the acoustics of /sC/ sequences. For voiceless stops, the realignment that results in metathesis has the obvious characteristic of a long stop release. The articulatory account also predicts maintaining voicelessness during the closure. Finally, the existence of a trading relationship—or lack thereof—between closure duration and stop release duration may shed light on the progression of the change.

For /s/ clusters with voiced stops and sonorants, it is not clear from existing phonetic studies whether there is a process of gestural alignment similar to that which occurs in /s/-voiceless stop sequences. Whether the metathesis is truly limited to /s/-voiceless stop sequences or whether

it applies similarly to /s/-voiced stop and /s/-sonorant sequences is important because the answer could shed light on the cause of the change. For /pt kt/ sequences, whether they undergo metathesis like /s/-voiceless stop clusters is important because it would suggest that metathesis is spreading to new contexts.

To that end, I investigate the following questions:

- Is there phonetic evidence of gestural overlap in the realization of metathesized /s/-voiceless stop sequences? Are these sequences integrating into the system by undergoing other processes that target voiceless stops? Are there trade-offs between closure duration and stop release duration?
- Is there phonetic evidence for gestural overlap in /s/-voiced stop sequences and /s/-sonorant clusters?
- Do /pt kt/ sequences show gemination and metathesis as well?

2.2.2 Materials and task

Participants read a series of paragraphs containing target words (real and nonce).

There were 120 target words. Half contained /sC/ clusters of /s/ and voiceless stops (/ptk/), voiced stops (/bdɡ/), fricatives (/θf/), the affricate (/tʃ/), and sonorants (/mnl/). There were five words per medial consonant (one for each vowel quality /a e i o u/). The other half of the target words contained intervocalic /C/ and were minimally different from each /sC/ word, matching in both surrounding vowels and stress (e.g. /ˈkaspa/ - /ˈkapa/, 'dandruff' vs. 'cape'). Clusters with affricates occurred across word boundaries (/las tʃaketas/ - /la tʃaketa/, 'the jackets' vs. 'the jacket'), because Spanish words do not have /stʃ/ sequences.

There were an additional 10 words with /pt/ and /kt/ clusters, each matched with intervocalic /t/ words (e.g. /kapta/ - /pata/). These words were matched as closely as possible, but two were matched with /p/ and /k/ words instead. When there were no existing words in the lexicon that matched a certain pattern (e.g. /sb sd sɡ/ words are rare), nonce words were invented and

inserted into the paragraph stories as people's last names. Sample target words are illustrated in Table 2.3.

Table 2.3: Sample target words for production task

	/sC/		/C/	
/ptk/	ˈkaspa	'dandruff'	ˈkapa	'cape'
/bdɡ/	resˈbala	'3SG slips'	reˈbaxah	'discount'
/mnl/	ˈmisma	'same'	ˈmima	'3SG spoils'
/pt kt/	ˈkapta	'3SG captures'	ˈpata	'paw'

This resulted in a total of 140 target words, 120 with coda /s/ and their intervocalic pairs and 20 with coda /pk/ and their intervocalic pairs. I analyze only /s/-voiceless stop, /s/-voiced stop, and /s/-sonorant clusters; numbers of tokens analyzed (after exclusions) are in Table 2.4.

Table 2.4: Number of tokens analyzed for Intervocalic /C/ and /sC/ words

	Consonant	Intervocalic	/sC/ (or /CC/)
Coda /s/	p	32	21
	t	30	20
	k	27	22
	b	27	22
	d	34	16
	g	32	19
	m	33	13
	n	31	12
	l	32	15
Coda /p, k/	p	7	18 (/pt/)
	t	45	—
	k	7	18 (/kt/)

For all categories except for /pt pk/, I analyze only tokens without a clearly separate [s], [h] or [ɦ] preceding the consonant. Tokens with preceding [s h ɦ] (with some period of frication, sibilance, or noise preceding the consonant) were excluded (/sp st sk/ = 35; [sb sd sɡ] = 31; /sm sn sl/ = 43). I do, however, briefly discuss /s/-voiceless stop sequences that were realized with some

portion of frication preceding the stop (Section 2.3.1.1). While I do not analyze the phonetics of these sequences, their presence in the data is important for the description of metathesis as gestural overlap. Some /pt kt/ sequences had [s, h, ɦ] preceding the stop closure (as debuccalization of the coda /p k/). There were only 5 of these, and they were included in the data analyzed because token numbers are already so low. Additionally, some /pt kt/ tokens were realized with two distinct closures and stop releases. I use only those tokens where there was a single closure and stop release.

Recordings were done in a sound-proof booth at New York University with a Zoom H4N Pro recorder and an Audio-Technica AT831b lapel microphone. The participants saw paragraphs on a laptop screen, one at a time. They first read the paragraph silently to themselves and answered a simple comprehension question to ensure they had read the paragraph. Then, they read the paragraph out loud. They were told to read at a normal speed, using a speech style approximating one they would use with their friends or in daily life. This was done to counteract the effects of formal lab speech. Participants were well aware of their distinct speech styles, and understood that casual speech was the goal.

2.2.3 Measurements and tokens

Segmentation was done manually in Praat. Boundaries were placed around the following elements, as relevant for each consonant type:

- The vowel preceding /sC/ or /C/ ([pehkʰa])
- Any [s], [h] or [ɦ] preceding the consonant ([pehkʰa]) (these productions were excluded from analysis, except for /pt kt/ sequences)
- The consonant closure/constriction ([pehkʰa])
- The release (voiceless stops only) ([pehkʰa])

For all consonant sequences, preceding vowels were segmented according to the onset and offset of F2. [s] or [h] preceding the consonant were marked as beginning at the zero-crossing at

the offset of F2 of the vowel, which coincided with increased noise in the waveform and attenuated formant structure at the higher frequencies. In practice, this means that not all of the frication is captured in the [s, h, ɦ] segment preceding the target consonant, and that portions of the frication noise are voiced. [3]

For intervocalic voiceless stops and /s/-voiceless stop sequences, the stop closure was marked as beginning at the offset of the preceding vowel's F2, or the offset of frication corresponding to a preceding [s, h, ɦ]; the end boundary was placed at the burst onset (Figures 2.1 and 2.2). The stop release was measured from the onset of the burst to the zero-crossing of the first periodic wave of the following vowel. The measurements were the same for /pt kt/ sequences, since I analyze only those productions that had a single closure and stop release.

For intervocalic voiced stops and /s/-voiced stop sequences, recall that /bdɡ/ undergo spirantization in these contexts. Boundaries were placed around the constriction by locating the region of decreased formant intensity, and, in cases of very light constriction, of changes in formant location (Figures 2.3 and 2.4). When neither of these cues were available, the token was not used for duration measurements. These tokens were still used for intensity measurements, which only use the minimum and maximum intensity measurements within the interval.

For intervocalic sonorants, boundaries were placed at the zero-crossings of the first characteristic waves of each segment, as with the other consonant types. For /s/-sonorant clusters, there was often a gap between [s] or [h] and the following sonorant. This space was segmented as part of the sonorant, because it seemed to result when the oral closure of the sonorant formed before the onset of voicing.

[3]Note that this differs from other methods; for example, Ruch and Peters (2016) determined the duration of aspiration by using a pitch tracker's inability to calculate pitch during voiceless portions.

Figure 2.1: Spectrogram of intervocalic /k/

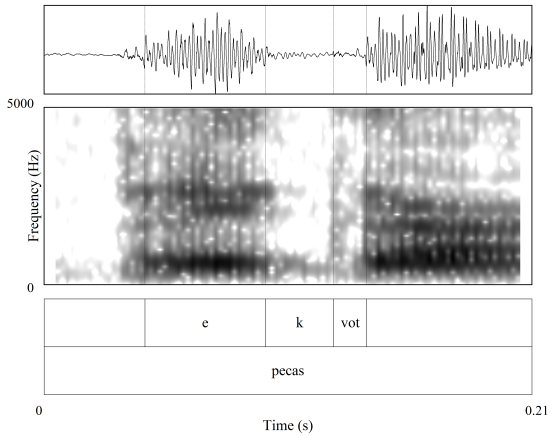


Figure 2.2: Spectrogram of /sk/ sequence

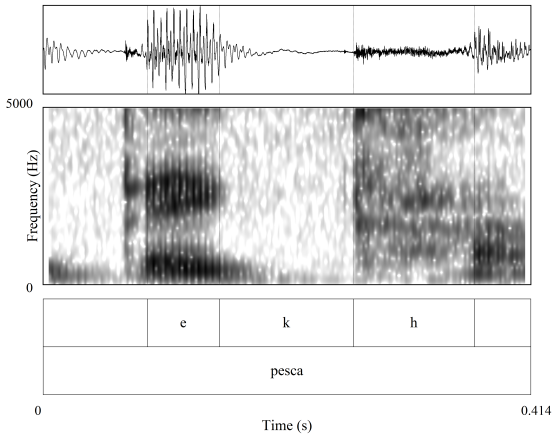


Figure 2.3: Spectrogram of intervocalic /b/

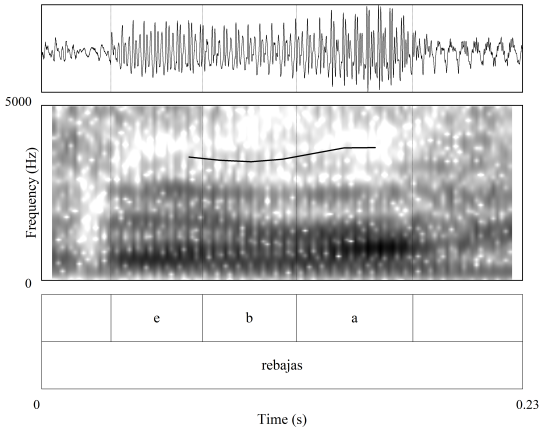
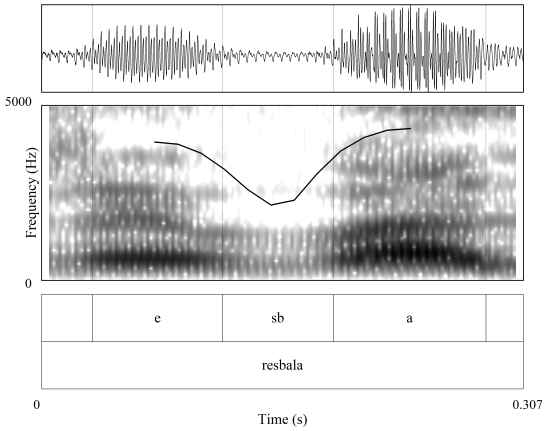


Figure 2.4: Spectrogram of /sb/ cluster



Most measurements were extracted with a Praat script. Cepstral peak prominence was extracted with *praatsauce*, a series of Praat scripts developed by Kirby (2018) to extract spectral measures using existing Praat functions. Measurements taken were as follows. Some measurements were taken for all types of consonants, but others were only relevant for certain consonants.

- **All consonants**: For all sequences, I measured the duration of the preceding vowel, [s, h, ɦ] (if present), and the constriction. Recall that for all but /pt kt/ sequences, I exclude tokens with preceding [s, h, ɦ] from the phonetic analysis.
- **Voiceless stops**: For voiceless stops, I also measured the duration of the stop release, and the amount of voicing during closure. Percent voicelessness during the closure was obtained using Praat's Voice Report and later converted to percent voicing.
- **Voiced stops**: For voiced stops, I also measured the amount of voicing during the constriction (using the same method as for voiceless stops), the constriction degree of the stop, and cepstral peak prominence (CPP) to investigate noise. Constriction degree was calculated by extracting the maximum intensity of the preceding vowel and the minimum intensity during the constriction, and dividing the maximum by the minimum. Higher ratios correspond to higher constriction degrees (more of a dip in intensity for the constriction). This method follows Soler and Romero (1999), Martínez-Celdrán and Regueira (2008), Hualde et al. (2011), and Rogers (2016), among others.

  I use cepstral peak prominence (CPP) to measure the degree of frication between spirantized interovocalic voiced stops, and the coalesced productions of /sb sd sɡ/ seqeunces. CPP is a measure of aperiodicity in the acoustic signal (see Heman-Ackah et al. 2003 for a clear introduction). CPP is most commonly used in speech pathology research and clinical work to measure voice quality (e.g. Hillenbrand and Houde 1996; Heman-Ackah et al. 2003). In phonetics, it has been used as a measure of voice quality in breathy vowels, tones, and stress (Berkson 2019; Keating et al. 2010; Esposito and Khan 2012; Garellek and White 2015; Garellek and Esposito 2021). High CPP indicates a highly periodic signal with low noise. Low CPP indicates that there is more noise in the signal. I use CPP (as opposed to a noise measure like the harmonics-to-noise ratio) for several reasons. First, CPP captures spectral information from the entire frequency spectrum, which is important because fricatives have information above the frequency ranges captured in HNR measures. Second, CPP does

not rely on the ability to track fundamental frequency (Heman-Ackah et al. 2003), which is crucial because some of the /sb sd sg/ realizations are partially devoiced.

Cepstral measures have been used in previous studies to measure properties of consonants. Spinu et al. (2018) use cepstral measures to categorize Russian fricatives, and Colantoni (2006) uses a cepstral measure to analyze periodicity in Spanish palatals and rhotics undergoing changes in frication. Repiso-Puigdelliura et al. (2021) use CPP specifically to analyze the degree of frication in Spanish voiced palatal production, and Santagada and Gurlekian (1989) use a cepstral measure to analyze frication in Spanish spirantized voiced stops, arguing that they are approximants, not fricatives.

CPP was taken every 5ms over the entire constriction interval, using the default settings of *praatsauce*. Plots and models show the average CPP over the interval.

- **Sonorants**: For sonorants, I also measured CPP to investigate signs of overlap between /s/ and the following sonorant. CPP is affected by nasalescence (Madill et al. 2019), but this should not matter because nasalescence should contribute equally to CPP in the relevant comparisons (e.g. /m/-/sm/).
- **Voiceless stop-Voiceless stop sequences:** For /pt kt/ sequences, I also marked multiple stop releases and closures when present, but these tokens were excluded from the current study.

2.2.4 Participants

Seven Sevillians (4 male/3 female) in New York City participated in the study. Data was collected as a pilot study in late 2019, with the intention to collect further data in Seville. Due to the COVID-19 pandemic, these plans were canceled. Most participants had completed or were currently completing postgraduate degrees (6/7), and one was in a technical course after graduating high school. They ranged in age from 18-40. Two had been in the US for 1-2 years, four had been in the US 5 or more years, and one was living in Seville but was recorded on a short-term visit to New York. All had advanced knowledge of English, and several also reported lower proficiency in French and Brazilian Portuguese.

2.2.5 Hypotheses

Voiceless stops: Previous research leads us to expect that /sp st sk/ sequences will have longer stop release durations and longer closure durations than intervocalic voiceless stops. Based on the articulatory model shown in Section 2.1.3, I also hypothesize that /sC/ sequences will resist intervocalic voicing more than intervocalic stops, since the [h] gesture may overlap the stop closure. I also expect some tokens to be realized with [h] split across an intervening stop, based on the gestural description of metathesis and Ruch's (2008) report of variants present in Seville.

Voiced stops: Previous studies also lead us to expect that /s/-voiced stop sequences should be longer than intervocalic voiced stops, and possibly partly voiceless or noisy, due to overlap of the voiced stop and an overlapping gesture from /s/. In terms of CPP, this means that /bdg/ should have higher CPP (less noise) than /sb sd sg/ (more noise). Additionally, although results on constriction degree are conflicting for Andalusian Spanish, high constriction degree in coalesced /sb sd sg/ sequences has been found in other varieties as well (Chilean Spanish, Rogers 2016).

Sonorants: Based on previous work, /s/-sonorant clusters should show longer closure durations than intervocalic sonorants. Previous work also reports partial voicelessness and/or noise during the closure. For noise, this prediction means higher CPP (less noise) for /mnl/ than for /sm sn sl/ (more noise).

/pt kt/: These clusters could show evidence of gemination and metathesis, since the coda obstruents have been reported to debuccalize to [h], like coda /s/.

2.3 Results

Visual interpretations of the plots are confirmed by statistical models. For each measurement on each consonant type, I ran a linear mixed-effects model (*lme4*; Bates et al. 2015) and tested comparisons between other levels of predictors using *emmeans*, Tukey adjusted for multiple comparisons (Lenth 2020). The dependent variable was the measurement (e.g. closure duration, release duration, percent voiced, constriction degree, CPP). The independent variables were Consonant

(e.g. /ptk, bdg, mnl/), Preceding Vowel (/a, e, i, o, u/), and Type (Intervocalic /C/, /sC/). Models also had random intercepts by speaker.[4] I report model results in the text, and relevant *emmeans* comparisons in tables.

2.3.1 Voiceless stops

/s/-voiceless stop clusters differ from intervocalic voiceless stops in stop release duration, closure duration, and amount of voicing during the closure.

Figure 2.5 shows that /s/-voiceless stop sequences have longer releases than their intervocalic counterparts. The model confirms a main effect of Type (Intervocalic vs. /sC/), which indicates that /st/ sequences have longer releases than intervocalic /t/ (β = .03, p < .001). There is also an interaction between Type and Consonant, whereby the effect of Type is not as strong for /p/ as for /t/ (β = -.01, p < .01). The effect of Type for /k/ is not significantly different from that of /t/. Post-hoc tests with *emmeans* (Table 2.5) highlight a different result: /st/ and /sk/ do not have significantly different stop release durations, and both have longer stop releases than /sp/. Note also that the stop release of intervocalic /t/ is not significantly different from that of either /p/ or /k/, which is somewhat unexpected. It is not unusual for /p/ and /t/ to have similar-length stop releases, but they are both usually substantially shorter than /k/, cross-linguistically and in Spanish (Cho and Ladefoged 1999; Rosner et al. 2000). Because my data do not fit this expected pattern, the result of similar-length durations in /st/ and /sk/ sequences should be interpreted with caution.

Table 2.5: *emmeans* comparisons for stop release duration for voiceless stops

contrast	Type	estimate	SE	df	t.ratio	p.value
t - p	Intervoc	0.00	0.00	154.64	1.70	0.21
t - k	Intervoc	-0.00	0.00	155.23	-1.27	0.41
p - k	Intervoc	-0.01	0.00	155.13	-2.95	0.01*
t - p	sC	0.02	0.00	155.07	5.07	0.00***
t - k	sC	-0.01	0.00	155.03	-2.03	0.11
p - k	sC	-0.03	0.00	154.89	-7.27	0.00***

[4] I do not discuss the effects of surrounding vowels, because there are too few tokens per category for meaningful results.

Closure duration of the consonant following /s/ is also longer in /s/-voiceless stop words than in intervocalic words for /p/ and /k/ (Figure 2.6). In the model there is no main effect of Type, confirming the visual interpretation that Type does not affect the closure duration of /t/ (the baseline level of Consonant in the model). The significant interaction between Type and Consonant confirms that, for both /p/ and /k/, /sC/ words have longer stop releases than intervocalic /c/ words (/sC/*/p/: $\beta = .02$, $p < .001$; /sC/*/k/: $\beta = .02$, $p < .001$).

For voicing, the closure portions of intervocalic stops undergo more intervocalic voicing than the closures in /sC/ sequences (Figure 2.7). The model has a main effect of Type, such that /sC/ closures are more voiceless than intervocalic /C/ ($\beta = 19.98$, $p < .001$). There is no significant interaction with Consonant, but, numerically, /st/ clusters have slightly more intervocalic voicing (on average, 18%) than /sp/ and /sk/ sequences (15% and 14%, respectively).

Figure 2.5: Voiceless Stops: Plot of stop release duration

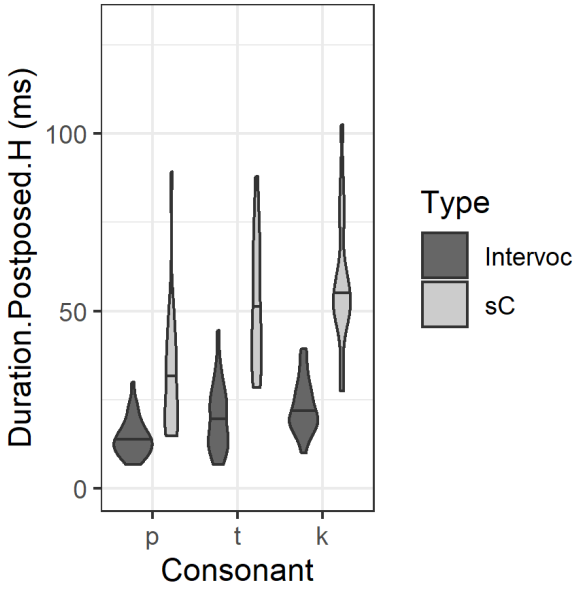


Figure 2.6: Voiceless Stops: Plot of closure duration

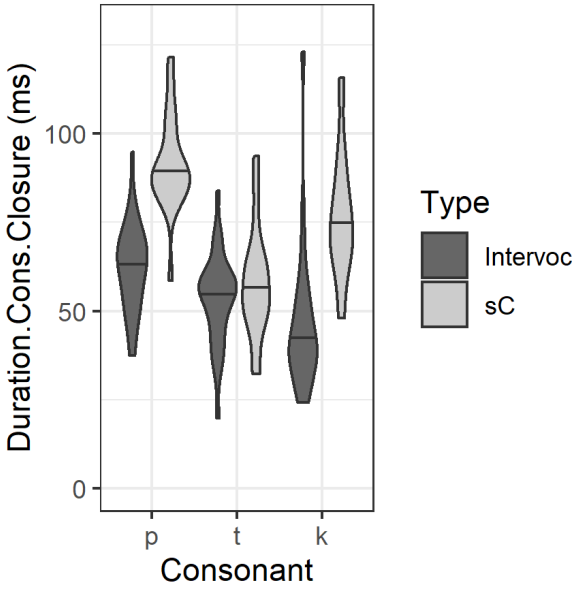
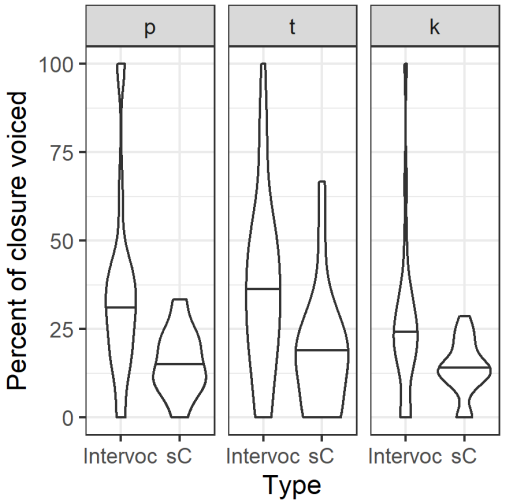
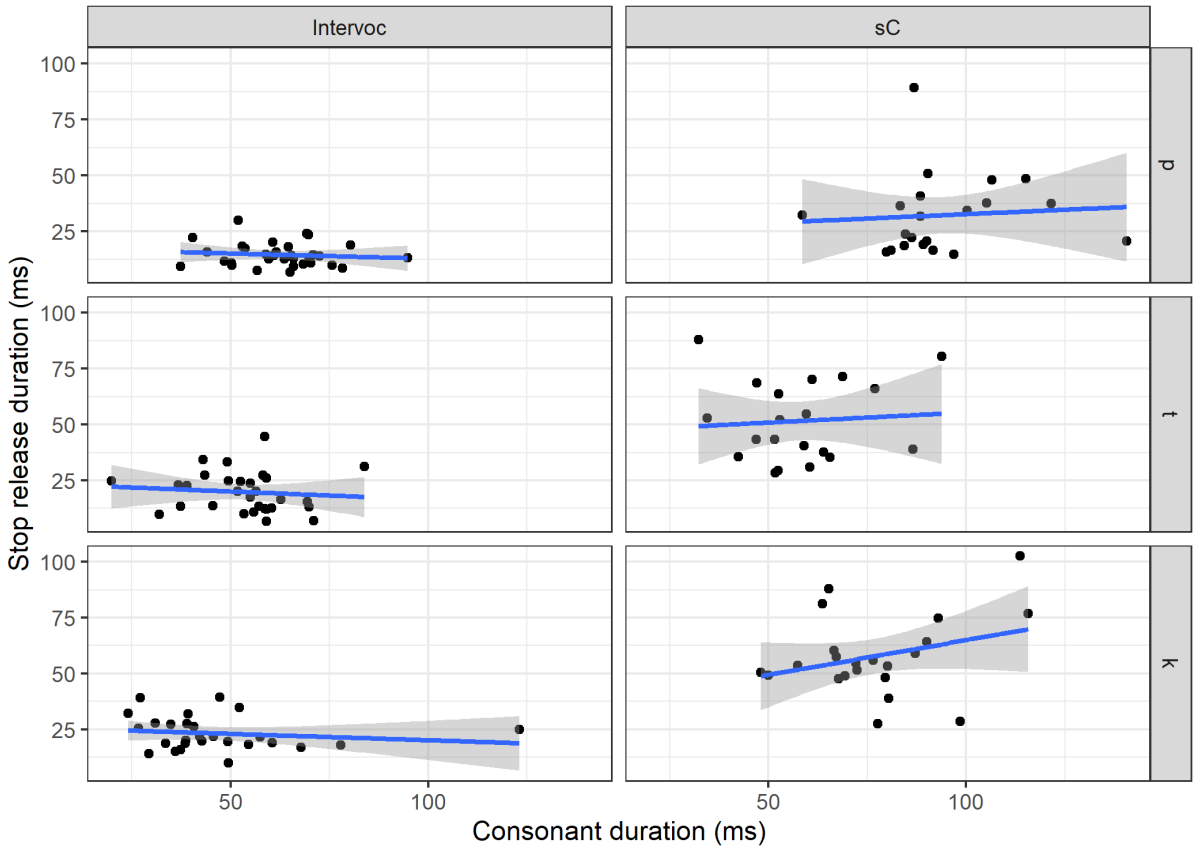


Figure 2.7: Voiceless Stops: Plot of percent voiced



In terms of trading relationships, Figure 2.8 shows that there is no strong positive or negative correlation between closure duration and stop release duration, for either intervocalic voiceless stops or /s/-voiceless stop sequences.

Figure 2.8: Plot of correlation between closure duration and stop release duration



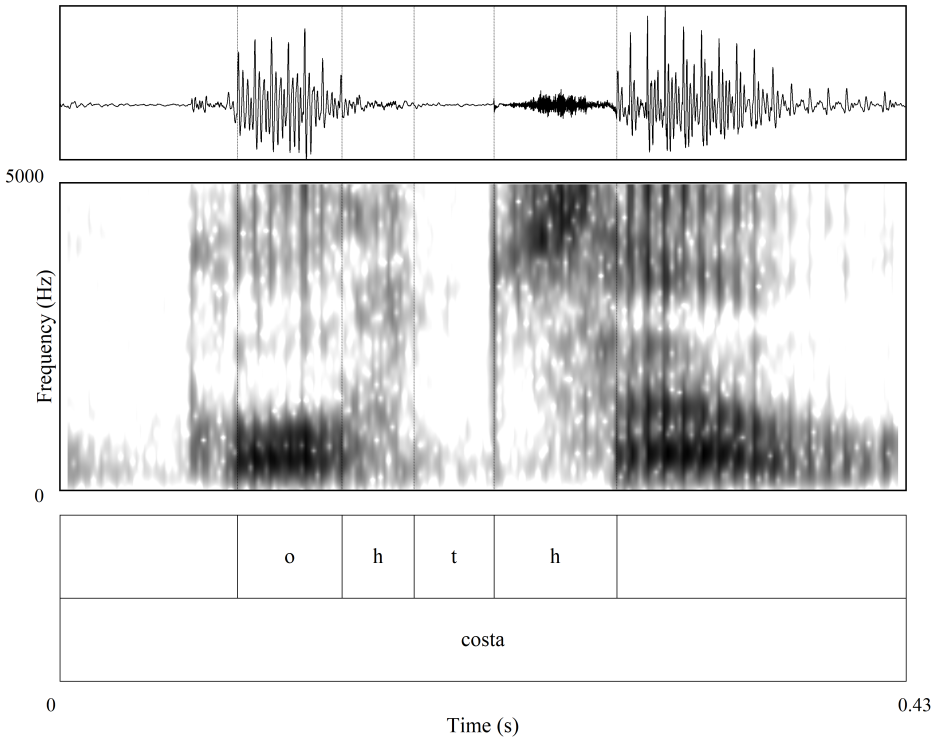
In sum, /s/-voiceless stop clusters (with no [h] or [s] preceding the stop closure) differ from intervocalic stops. /s/-voiceless stop clusters have longer stop release durations, longer closure duration of the following consonant (except for /st/), and undergo less voicing during the closure. The lack of correlation between stop closure and release duration suggests that these cues are independent of each other.

2.3.1.1 Voiceless stops: [hCh] sequences

Recall that I excluded from the phonetic analyses /s/-voiceless stop sequences that were realized with some glottalization or frication preceding the stop. I do not aim to document the phonetic properties of these sequences here, but I wish to document their existence because I believe that they are crucial for the analysis of Sevillian metathesis as gestural overlap.

In Section 2.1.3, I showed that forms like these are expected to arise under an articulatory description of the change, if the stop closure gesture moves partially, but not entirely, across the [h] gesture. The data show that these intermediate forms do exist; a sample production is illustrated in Figure 2.9.

Figure 2.9: Spectrogram of intermediate realization: [h] split across an intervening voiceless stop



In my study 35 /sp st sk/ realizations contained some glottalization or frication preceding /ptk/. My study is restricted to a laboratory reading task, and is relatively small, but in a larger study, Ruch (2008: 78) found that 11% of all /st/ realizations fell into the category [hth]/[sth]. These intermediate realizations are consistently present, and support my argument that metathesis occurs by gestural overlap. When metathesis is complete, it results in [Ch]. When it is incomplete, it results in [hCh].

2.3.2 Voiced stops

/s/-voiced stop sequences differ from spirantized intervocalic stops in similar ways. /s/-voiced stop sequences are longer than intervocalic voiced stops (Figure 2.10), and this is confirmed by a main effect of Type (Intervocalic vs. /sC/) in the model (β = .04, p < .001). The interaction between Type and Consonant is not significant.

/s/-voiced stop sequences also have higher intensity ratios than intervocalic voiced stops, which suggests a higher constriction degree (Figure 2.11). The model has a main effect of Type, suggesting that baseline /b/ has a higher intensity ratio in /sC/ contexts than in /C/ contexts (β = .17, p < .001). There is also a main effect of Consonant, such that, intervocalically, /d/ has a lower intensity ratio than baseline /b/ (β = -.06, p < .01). The interaction between Type and Consonant suggests that the /sC/ context increases the intensity ratio more for /d/ than for /b/ (β = .07, p < .05). In other words, /d/-/sd/ differ most in intensity ratio, while /b/-/sb/ and /ɡ/-/sɡ/ do not differ as much. This is likely because /d/ has lower intensity ratios (less constriction) to begin with, so the increase in ratio (constriction) in /sd/ clusters is larger than for the others, whose intervocalic counterparts have higher ratios to to begin with.

Furthermore, while intervocalic voiced stops are almost entirely voiced, /s/-voiced stop sequences are less voiced (Figure 2.12). The model confirms this result (β = 16.82, p < .001), and there is no interaction between Type and Consonant. /s/-voiced stop clusters also have lower CPP than intervocalic /bdɡ/, indicating that realizations of /sb sd sɡ/ sequences are noisier than intervocalic [βðɣ] (Figure 2.13). The model confirms lower CPP for /s/-voiced stop sequences than for intervocalic voiced stops (β = -2.44, p < .01), and there is no interaction between Type and Consonant.

Figure 2.10: Voiced Stops: Plot of closure duration

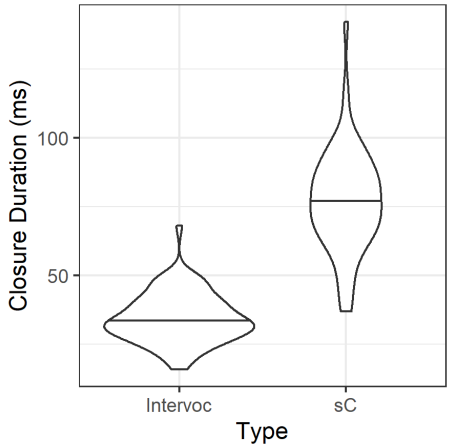


Figure 2.11: Voiced Stops: Plot of constriction degree

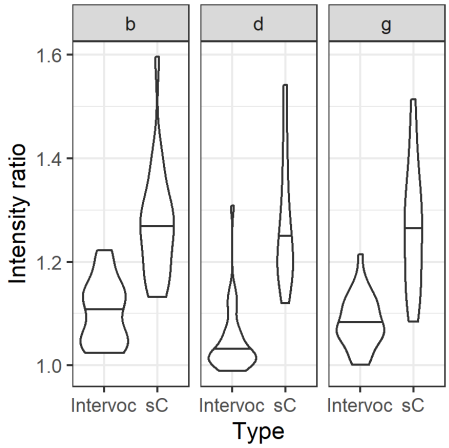


Figure 2.12: Voiced Stops: Plot of percent voiced

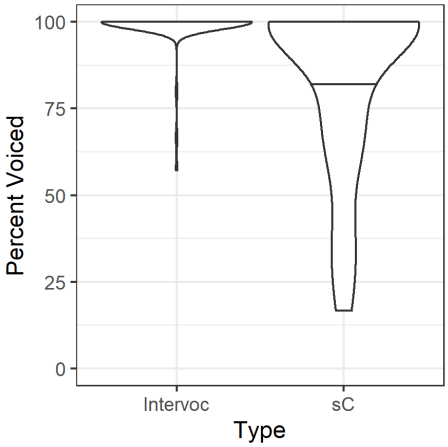
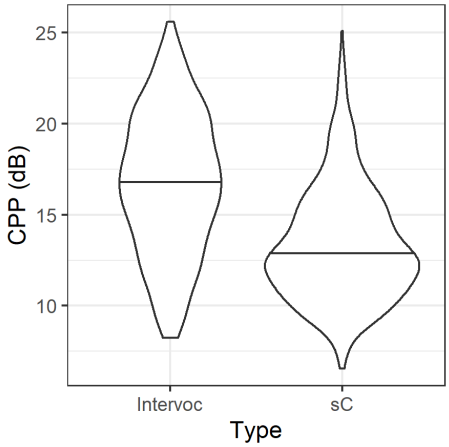


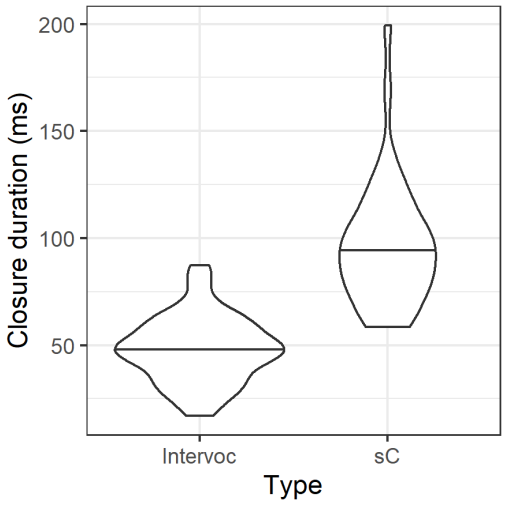
Figure 2.13: Voiced Stops: Plot of CPP



2.3.3 Sonorants

Similarly to clusters of /s/ with voiceless and voiced stops, /s/-sonorant clusters also show lengthening of the sonorant, as compared to intervocalic sonorants (Figure 2.14). The model confirms this effect ($\beta = .05$, $p < .001$).

Figure 2.14: Sonorants: Plot of closure duration



I also investigated voicelessness and noise in /s/-sonorant clusters. Of the 40 /s/-sonorant tokens, 36 tokens were found to be 100% voiced. 4 tokens were found to have partial voicelessness. Upon manual inspection, however, the clusters Praat determined to be partially voiceless were not actually voiceless. The voicelessness appears to have been measurement error.

There was no effect of CPP on sonorants or /s/-sonorant sequences, either visually or in the model. Intervocalic /mnl/ and /sm sn sl/ all have high periodicity and low levels of noise. These sequences undergo gemination but not metathesis/overlap, further suggesting that the two are separate phenomena.

2.3.4 Voiceless stop-Voiceless stop clusters

Figures 2.16 and 2.15 show the closure duration and stop release duration for intervocalic /p t k/ and /pt, kt/ clusters. Recall that some of these /pt kt/ words were matched with intervocalic /p k/ words instead of /t/ words. Data from all of these words is shown for completeness.

Figure 2.15: Plot of stop release duration in /pt, kt/ sequences and intervocalic /p t k/

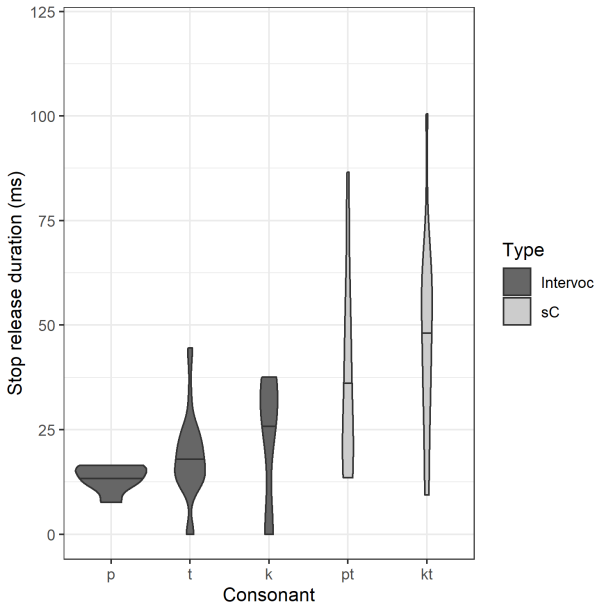
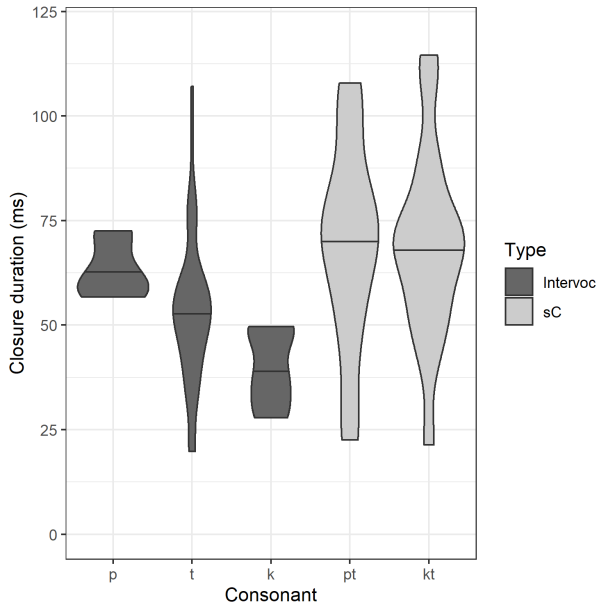


Figure 2.16: Plot of closure duration in /pt, kt/ sequences and intervocalic /p t k/



The models contained a single main effect of Consonant (/p t k pt kt/) and a random intercept by Speaker. For stop release duration, the model has a significant main effect of Consonant. *emmeans* pairwise comparisons indicate that /ptk/ all have shorter releases than /pt/ and /kt/ (Table 2.6; shaded gray). For closure duration, the effect of Consonant is significant. Pairwise comparisons with *emmeans* indicate that /pt/ and /kt/ have longer closure durations than /t/ and /k/, but not /p/ (Table 2.7, shaded gray).

Table 2.6: *emmeans* comparisons for /pt kt/ for stop release duration

contrast	estimate	SE	df	t.ratio	p.value
p - t	-0.00	0.01	107.26	-0.76	0.94
p - k	-0.01	0.01	107.15	-1.46	0.59
p - pt	-0.02	0.01	107.72	-3.77	0.00**
p - kt	-0.03	0.01	108.47	-5.24	0.00***
t - k	-0.01	0.00	107.35	-1.23	0.73
t - pt	-0.02	0.00	109.34	-4.97	0.00***
t - kt	-0.03	0.00	110.48	-7.23	0.00***
k - pt	-0.01	0.01	108.03	-2.82	0.04*
k - kt	-0.02	0.01	109.14	-4.65	0.00***
pt - kt	-0.01	0.00	108.32	-2.04	0.26

Table 2.7: *emmeans* comparisons for /pt, kt/ for closure duration

contrast	estimate	SE	df	t.ratio	p.value
p - t	0.01	0.01	109.29	1.18	0.76
p - k	0.02	0.01	109.03	2.37	0.13
p - pt	-0.02	0.01	110.93	-1.78	0.39
p - kt	-0.01	0.01	112.97	-0.99	0.86
t - k	0.01	0.01	109.49	2.08	0.24
t - pt	-0.03	0.01	114.85	-4.61	0.00***
t - kt	-0.02	0.01	117.05	-3.30	0.01*
k - pt	-0.04	0.01	111.91	-5.28	0.00***
k - kt	-0.03	0.01	114.68	-4.26	0.00***
pt - kt	0.01	0.01	112.12	1.05	0.83

2.4 Discussion

My results largely confirm the results of previous studies, and build on them. In addition to gemination that happens for all types of /sC/ sequences, the phonetics provide evidence that metathesis occurs via gestural overlap for /s/-voiceless stops, and that /s/-voiced stop clusters also undergo gestural overlap, if not full metathesis. These results are as predicted by the articulatory model

in Section 2.1.3. Furthermore, the acoustic patterns provide evidence that the change continues to progress for voiceless stops, further integrating into the phonological system by undergoing intervocalic voicing. The results also document metathesis in /pt kt/ sequences, an extension to a novel context. Finally, the gestural description of metathesis explains many of the observed phonetic properties, but cannot be the only factor at play because of gemination and degemination of the following stop, along with other factors further discussed in Chapter 5, Section 5.1.1. I elaborate on these points in the following sections.

2.4.1 Metathesis as gestural overlap

Given the acoustic analyses presented in this chapter, gestural overlap appears to occur in /s/-voiceless stop and /s/-voiced stop sequences, resulting in what looks like full metathesis in the former, but not in the latter. /s/-sonorant sequences show no evidence of overlap.

Voiceless stops: In /s/-voiceless stop sequences, stop closures in /sC/ sequences appear to resist passive voicing more than intervocalic voiceless stops. In many dialects of Spanish, including Peninsular varieties, intervocalic /ptk/ undergo passive voicing (Lewis 2001; Torreira and Ernestus 2011). Intervocalic voicing has also been reported specifically for Andalusian Spanish (Salvador 1968; Galvano and Henriksen 2021). Cross-linguistically, intervocalic voicing as a reduction pattern is common with unaspirated voiceless stops (Gurevich 2003: 37), but is rarer with aspirated stops.[5] Aspirated stops presumably resist passive voicing because they have a [spread glottis] feature that prevents it (Beckman et al. 2013: 270). That Sevillian [Ch] sequences undergo less intervocalic voicing than intervocalic voiceless stops suggests that the [h] gesture may overlap the closure, preventing some of this voicing.

[5]For one example, Gurevich (2003: 173) reports that Limbu voices aspirated voiceless stops, but they do not lose their aspiration. However, note that Davidson (2018) finds that even aspirated stops (onsets to stressed syllables) in American English have at least some passive voicing. Shih et al. (1999) find that aspirated stops in Mandarin undergo less voicing than their voiceless unaspirated counterparts. However, they also find that features and voicing patterns do not always correlate: for example, Mandarin voiceless aspirated stops pattern with voiceless stops of German (aspiration or unaspirated), Spanish (unaspirated) and Italian (unaspirated.)

Additionally, there were productions with [h] split across the stop, realizations that suggest that [h] completely overlaps the stop and extends on either side. While these productions are not the most frequent, they illustrate an intermediate stage in the process and suggest that full metathesis is just the final product of gestural overlap and realignment. I illustrate this process in (12), in Section 2.5.

Voiced stops: For /s/-voiced stop sequences, the results also suggest gestural overlap. This overlap may be the same in terms of gestural realignment as for /s/-voiceless stop sequences, but may look different on the surface because voiced stops spirantize. In these sequences, /s/ and the spirantized voiced stops appear to be superimposed. One indication of overlap is the increased voicelessness in the coalesced sequences resulting from /sb sd sɡ/ as compared to [βðɣ]: an [h] gesture overlapped with a voiced stop constriction could have this effect. Another sign of overlap is the increased noise in /s/-voiced stop sequences, as compared to intervocalic [βðɣ]. /sb sd sɡ/ are produced with frication in the acoustic signal, which can be attributed to features of underlying /s/. Recall that none of these tokens had a clearly separate [h] or [s] segment: the voiced stop and the fricative are entirely overlapped. Even if the oral gesture of /s/ has been deleted, a remaining glottal gesture could still affect the spirantized voiced stop.

Constriction degree may also reflect gestural overlap, although less clearly. Increased constriction degree makes articulatory sense for /sd/ clusters. In these clusters, the segments share an articulator, if we assume that /s/ is produced as [s]. The tongue tip for intervocalic spirantized /d/ may create very little constriction, but [s] requires the tongue tip to create high constriction in order to produce frication. Overlapping these two gestures could result in higher constriction than for spirantized /d/ alone. A similar explanation could hold for /sɡ/: because /ɡ/ is produced further back in the vocal tract, a tongue tip constriction could lower intensity in the acoustic signal, which is my measurement of constriction. However, these accounts require that /s/ be produced as [s]: if it loses its oral gesture and reduces to [h], the argument no longer holds. Without a tongue tip gesture, [h] cannot—in purely articulatory terms—increase constriction. There is no reason to think

/s/ is retained as [s] in these environments, since it generally undergoes such drastic reduction in coda position in this variety. Furthermore, the argument that increased constriction is due to overlap does not hold for /sb/ sequences, since an additional tongue tip restriction would be behind the bilabial constriction, and would not result in show the acoustic correlate of lower intensity.

Apart from gestural overlap, there are alternative explanations of increased constriction. Regardless of the phonetic details reported, existing literature reports that /sb sd sg/ clusters show constriction *at the same place of articulation* as the voiced stop (Romero 1995: 132), resulting in a series of fricatives like [ɸθx] (Hualde 1989a) or more fricative-like versions of the IPA approximants [βðɣ] that are used to represent spirantized allophones (Romero 1995: 10). Romero (1995: 130-1) argues that the differences in constriction degree between /sb sd sg/ clusters and intervocalic spirantized /bdg/ derive from differences in duration, which are clear in both his study and mine. Fricative-like segments (the outcome of /sb sd sg/) require high constriction to result in turbulent airflow; building up the pressure to produce this turbulent airflow requires the the constriction to be held for long enough. In contrast, a shorter constriction for approximants [βðɣ] leads to less constriction and less noise.

The difference in airflow between fricatives and approximants is reflected in cepstral measures. Santagada and Gurlekian (1989) compare Spanish fricatives /s z/, /l/, and the spirantized allophones [βðɣ], and find that [βðɣ] have cepstral measures more similar to /l/ than to fricatives /s, z/. That is, [βðɣ] have similar levels of noise as the voiced sonorant /l/. Santagada and Gurlekian (1989) thus distinguish [βðɣ] from fricatives. In my data, then, the fact that CPP is significantly higher in realizations of /sb sd sg/ than in intervocalic /bdg/ suggests that the former are more sibilant-like (have more turbulent airflow) than the latter.

Another possibility for the increased constriction in /sb sd sg/ clusters is that gestural overlap started in the coronal context (/sd/), as with /s/-voiceless stop sequences, and high constriction degree became a cue that spread to other sequences where it was not articulatorily motivated.

A final possible explanation is that the increased constriction in realizations of /sb sd sg/ clusters is due to a process of compensatory occlusion. Increased constriction in these clusters

has been found in other Spanish varieties as well, and the authors have put forth compensatory occlusion as a possible explanation (e.g. Puerto Rican Spanish, Galarza et al. 2014; Chilean, Rogers 2016).

I believe an explanation combining duration (along the lines of Romero 1995) and the need to avoid contrast neutralization is plausible. /sb sd sɡ/ sequences have two timing slots, so gemination of the voiced stop in the wake of /s/ reduction would lead to increased duration in comparison to intervocalic /bdɡ/. But what prevents shortening, and what keeps constriction high? Shortening may not occur in order to keep /sb sd sɡ/ acoustically distinct from [βðɣ]. Duration has been found to distinguish voiced vs. (lenited) voiceless stops in Spanish, in both production (Broś et al. 2021) and perception (Melero-García 2022). Duration may be an important cue for distinguishing /sb sd sɡ/ vs. /bdɡ/, too. And if /sb sd sɡ/ stay long, constriction degree can be reached and maintained. Studies of Spanish show that the degree of spirantization of intervocalic /bdɡ/ depends on duration: shorter consonant productions have less constriction. Most authors argue that this is because reduction occurs both to the duration and magnitude of the gesture (e.g. Soler and Romero 1999). Soler and Romero (1999) propose that a longer gesture provides more time for the articulators to make a constriction, so this constriction can be more complete.

Sonorants: My results do not show evidence of gestural overlap in /s/-sonorant clusters. Although they show gemination of the following consonant, like /s/-stop sequences (voiced and voiceless), they do not have devoicing or noise. These results contrast with anecdotal reports of partial devoicing in varieties of Spanish spoken in the region (see Section 2.1.2.3).

In my data, [h] appears to simply delete instead of overlapping with the nasal. There are several possible explanations for why overlap may be avoided. One is aerodynamic, and is often captured in featural terms. Assuming that [h] is the acoustic manifestation of a wide glottal gesture ([sg]) and the following sonorant is voiced ([voice]), then Sevillian [h] sliding over a sonorant would result in simultaneous [sg] and [voice] articulations and features. These are incompatible because [voice] requires the vocal folds to vibrate, and [sg] requires them *not* to vibrate. Solé

(2007) argues that aerodynamic requirements make both (a) the combination of nasalization and frication and (b) the combination of nasality and voicelessness difficult. Pulleyblank (1989) posits a co-occurrence restriction against nasality and voicelessness, and Steriade (1994: 212) uses a restriction against the combination of voicing and spread glottis. Either way, overlapping [h] with a nasal violates the restriction.

Another possibility is that the voicelessness or noise of [h] overlapped on a sonorant would be difficult to perceive, and [h] would essentially be lost. Most of the literature on [h] and sonorants focuses on nasals. [h] and nasals share phonetic properties and have been argued to be perceptually confusable (Ohala 1980; Ohala 1993). If this is true, then it makes sense that Sevillian deletes [h]: overlapping it with a following nasal would be essentially the same as deleting it. At first glance, the results from my ABX discrimination task in Chapter 5 do not appear to support this hypothesis, because [h] is quite perceptible next to sonorants. However, in those stimuli, [h] was clearly distinct from the adjacent sonorant and there was no overlap. Those results suggest that [h] is perceptible before a sonorant, but say nothing about how perceptible it would be if it were overlapped with the sonorant.

Finally, voiceless nasals are cross-linguistically rare (Ladefoged and Maddieson 1996: 106). Chirkova et al. (2019) report that out of 307 languages with nasals in the UPSID database, only 12 have voiceless nasals. Whether the reason is featural incompatibility or perceptual difficulty, the typology supports deletion of [h] in Sevillian rather than overlapping a following nasal.

2.4.2 Progression of the change in /s/-voiceless stop sequences

In my study, /s/-voiceless stop sequences were realized with longer stop release and longer closure duration than intervocalic voiceless stops, in line with previous findings (Torreira 2006; Torreira 2007; O'Neill 2010; Ruch and Harrington 2014; Ruch and Peters 2016; Galvano and Henriksen 2021). My results provide further evidence that /st/ clusters are more advanced than /sp sk/ in the change, as proposed by Ruch and Peters (2016) and Galvano and Henriksen (2021). This evidence

comes from the effect of place of articulation on stop release duration, intervocalic voicing, and degemination.

First, /st/ and /sk/ did not differ in stop release duration. Because /t/ and /k/ also did not differ in stop release duration, I do not read too much into the similarity between /st/ and /sk/, but note that Ruch and Harrington (2014) found the same. They took this 'reversal' of the universal VOT hierarchy as evidence that /st/ clusters were most advanced in the change. /k/ usually has a longer release than /t/, so if the release is similar in [th] and [kh] (← /st, sk/), the long release of [th] must be due to metathesis. This explanation predicts that, as the change becomes more established in /sp/ and /sk/ sequences, the universal order /k/ > /t/ will be reestablished. My data do show a trend of longer stop releases for /sk/ than /st/, although it does not reach significance. At minimum, this suggests that /sk/ is catching up to /st/, and may eventually surpass it.

A second piece of evidence of progression of the change is that [ph th kh] sequences undergo some intervocalic voicing, integrating into the system by undergoing the same process as intervocalic [p t k]. While stops at different places of articulation did not differ significantly from each other, /st/ clusters underwent slightly more voicing than /sp/ and /sk/ clusters, suggesting that they are the most integrated. A comparison with Galvano and Henriksen (2021) shows further progression. Galvano and Henriksen (2021) find that the probability of complete voicing is highest for intervocalic /ptk/, but between 55%-65% for /s/-voiceless stop sequences. My data are presented differently—in terms of percent voiced—but none of the /s/-voiceless stop closures are 100% voiced. This difference could reflect differences in speaker demographics, which, in turn, could reflect advancement of the change. My speakers ranged in age from 18-40; those in Galvano and Henriksen (2021) were slightly younger (18-30). More likely, however, the difference reflects regional variation, since my speakers were from Seville and theirs were from Jerez de la Frontera, a city further south. In either case, the more extensive intervocalic voicing in their study likely represents further progression of the change.

Finally, /st/ sequences show degemination, while /sp sk/ sequences do not. I find gemination only in /sp sk/ sequences; in /st/ sequences, the closure duration is similar to intervocalic /t/.

This result contrasts with Ruch and Peters (2016: 19), who found that all sequences had longer stop closures in /sC/ contexts than intervocalic /C/ contexts. If gemination arose prior to metathesis (Ruch and Harrington 2014; see Section 2.1.2.1), then degemination in /st/ sequences may reflect the advanced stage of /st/ clusters in the change: gemination is no longer a primary cue to the /t/-/st/ distinction. /sp sk/ clusters lag behind /st/, not having lost gemination yet. Recall that neither gemination nor degemination is explained by the articulatory model. In the analysis presented in Chapter 3, I suggest that this degemination occurs because /s/ is no longer in coda position, and thus can no longer bear a mora.

Metathesis has spread to /pt kt/ sequences as well. Although debuccalization of coda obstruents had been previously reported, metathesis has not. That metathesis affects these clusters as well suggests that the change has advanced to new contexts. Furthermore, it suggests that metathesis is surface-level: it applies to any *surface* coda [h], regardless of what consonant that [h] derived from.

2.4.3 Limitations and outstanding questions

This small-scale phonetic study has several limitations. First, there were few participants, and all but one were currently living in New York. All had grown up and lived most of their lives in Seville, and are thus expected to be representative of Sevillians living in Seville. But because metathesis is a change in progress, it will be important to do more extensive phonetic analysis with speakers currently residing there.

A second limitation is that the data was collected in a laboratory setting. Participants were asked to speak informally, as if to a friend; they recognized the stylistic differences between two registers of speech when asked to do this. However, read laboratory speech is different from spontaneous speech, both in terms of level of formality (e.g. Labov 2001) and the amount of variability produced (Gilbert 2022). My participants did produce metathesis in the majority of tokens for voiceless stops, in line with Ruch's (2008) findings that [th] sequences are almost as frequent in read speech as in conversational speech. My data from /s/-voiced stop and /s/-sonorant clusters

may be less representative. Metathesis with /s/-voiceless stop sequences is well-established, and appears not to be stigmatized, but it is possible that gestural overlap in other clusters is more socially salient, and more subject to formality constraints. Spontaneous speech might show more overlap in the form of voicelessness or noise during the closures of these clusters.

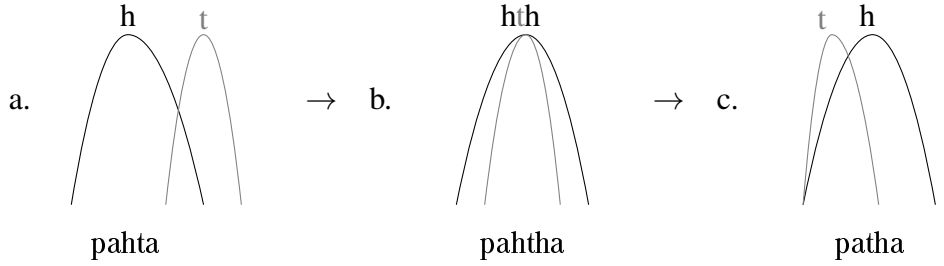
Finally, it is difficult to measure the duration of spirantized stops. Although the differences in my study were large enough to likely not be reducible to segmenting inconsistency (and were in line with results of articulatory proposals, Romero 1995), future phonetic studies should determine boundaries in the acoustic signal with an algorithm, like that developed in Ennever et al. (2017).

Future studies should focus on clusters in addition to /sp st sk/ in order to more fully understand the motivations and extent of the change. Does it apply everywhere, but is only phonetically realized as metathesis in /s/-voiceless stop sequences? Or is metathesis limited to just these clusters? Articulatory work would also shed light on the gestural coordination hypothesized to result in metathesis.

2.5 Implications

In sum, the phonetic data points to the change from h-stop to stop-h in /s/-voiceless stop sequences as occurring by metathesis, as illustrated in (12).

(12) Gestural realignment of [h] and [t] results in metathesis (following Parrell 2012)



I propose a sequence of steps like the following, following Parrell (2012) and Torreira (2006) . The [t] gesture originally follows the gesture for [h] (12a) and then slides to the middle so they overlap and [h] surfaces on both sides of [t] (12b). My data contained productions like this, with [h] split across an intervening stop. Finally, [t] slides to align with the beginning of the [h] gesture, so that

[h] is only acoustically present following the stop (12c). The percept is of segmental metathesis, but it has occurred through multiple steps of gestural overlap and realignment.

In /s/-voiceless stop sequences, metathesis is well-established and may be giving rise to new cues, including degemination, and integrating into the existing phonological system by undergoing intervocalic voicing. These cues appear to be most robust in /st/ sequences, which were also the first context in which metathesis applied. The change is also expanding to new contexts (/pt kt/).

The phonetic evidence presented in this chapter points to overlap, but not full metathesis, for /s/-voiced stop clusters. The process could be similar to that in /s/-voiceless stop clusters, but the realignment does not result in [h] extending past the constriction for voiced stops. It is possible that, since /bdɡ/ spirantize in this context, the acoustic result of 'metathesis' is different than for /s/-voiceless stop sequences, even though the process of realignment is the same. However, the process could also be fundamentally different with voiced stops—possibly because there is no stop release to 'attract' the [h] gesture (Kingston 1990). Finally, there is no evidence of overlap in /s/-sonorant clusters.

The terminological debate between calling /bdɡ/ fricatives vs. approximants is long-standing (e.g. Quilis 1981; Martínez-Celdrán 1984; Santagada and Gurlekian 1989). My results support a distinction between the two categories, realizations of which arise in different contexts: fricative-like realizations arise in /sb sd sɡ/ sequences in some dialects, while intervocalic /bdɡ/ are approximant-like. Transcribing productions of /bdɡ/ and /sb sd sɡ/ differently, the former as approximants ([βðɣ]) and the latter as more constricted, noisy, and fricative-like (β̞ ð̞ ɣ̞]), more accurately reflects the phonetic properties. This transcription follows that in Romero (1995).

The lack of correlation between closure duration and stop release duration in my data suggests that gemination and metathesis are, to some extent, independent. Both appear to occur in /s/-voiceless stop sequences and result in full metathesis; /s/-voiced stop sequences also appear to show both, but the phonetic result indicates overlap, not full metathesis. /s/-sonorant sequences only show gemination. With /t/, one novel finding is that /t/ and /st/ do not differ in closure duration, but do differ in stop release duration. This degemination in /st/ sequences may be a sign

that metathesis is well-established in these sequences. But it also suggests that stop release and closure duration operate independently, and that the ongoing change is not one solely of gestural realignment.

Recall from Section 2.1.3.2 that the inability of the articulatory model to account for all aspects of metathesis and gemination has inspired speculation that the change is perceptually motivated (e.g. Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016), a hypothesis I test in Chapter 5. I find no evidence that perceptual factors are *motivating* the change. However, the articulatory *description* of metathesis by overlap presented in this chapter—supported by phonetic evidence—is independent of the motivation, and can describe how metathesis occurs, regardless of *why* it occurs.

CHAPTER 3

Testing representations in Sevillian metathesis: Segments vs. clusters

3.0 Introduction

Sevillian stop-h sequences are ambiguous between being underlying aspirated stops /C^h/ or clusters /sC/. In the previous chapter, I presented phonetic evidence that suggests that surface stop-h sequences in Sevillian arise from metathesis of underlying /s/-voiceless stop clusters. In this chapter, I present further evidence for the cluster representation from a perception experiment: when they hear [Ch] sequences, Sevillian listeners map these sequences to underlying /sC/ clusters. I also present a formal analysis corresponding to the perception results, based on the hypothesis that Sevillians are able to map [Ch] to /sC/ in perception because their grammar provides a direct derivational path between the two. In contrast, listeners of other Spanish varieties do not map [Ch] to /sC/ in the perception experiment because their grammars provide no derivational path between the two forms, in either production or perception.

Determining the status of sequences of sounds is not always straightforward. In some languages, an [h]-like sound is analyzed as part of the adjacent stop (e.g. English, [p^{h}ɑt]); in others [h]

is analyzed as a segment in its own right (e.g. Acehnese, where the components of [Ch] can be split by an infix; Asyik 1987). Sevillian [Ch] sequences have been treated, on one hand, as underlying /sC/ clusters having undergone debuccalization and metathesis (Parrell 2012; Torreira 2012; Ruch 2013; Chapter 2 of this dissertation). This analysis assumes that the change in pronunciation has not resulted in a change of representation. On the other hand, there is another possibility: Sevillian [Ch] sequences may be phonologizing into aspirated stops (/$p^h t^h k^h$/) (O'Neill 2009; Gylfadottir 2015). Although few argue for this position, I take it seriously. Changes that result in different representations (and new segmental contrasts) are uncommon, as is the ability to study them while they are changing. If the surface form is phonetically similar to an aspirated stop, why don't learners posit this representation? Neither side of the debate leverages 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.

Based on data from a forced-choice perception study, I argue that Sevillian [Ch] sequences are realizations of underlying /sC/ clusters, not of /C^h/ segments. In the experiment, Sevillian listeners attribute [h] in [Ch] sequences to a morphological marker /-s/ on the preceding word. That is, [C] and [h] are separable and attributed to different morphemes. Because they are separable, I treat the mapping between underlying /sC/ and surface [Ch] as metathesis (following Ruch 2013) that occurs via gestural realignment (Parrell 2012; Torreira 2012; see also Chapter 2). Sevillian listeners 'undo' metathesis when they hear [Ch], arriving at underlying /sC/. Participants who speak other dialects of Spanish—Mexican and Argentinian Spanish—do not perform this same mapping, and do not glean linguistically-relevant information from [h] in [Ch] sequences.

I argue that listeners of the three dialects behave differently in perception because their phonological grammars differ. Sevillian listeners' grammars have a direct derivational path between /sC/ $\leftrightarrow$ [Ch], while other listeners' grammars do not. When Sevillian listeners hear [Ch] sequences, they follow the derivational path, 'undoing' metathesis to arrive at underlying /sC/. For other listeners, this path does not exist, and they do not 'undo' metathesis. To capture these differences, I develop a formal analysis for each dialect, cast in Harmonic Serialism. The derivational

steps in the analysis represent steps on a coda /s/ lenition continuum that plays out both diachronically and synchronically. Mexican Spanish is the most conservative (coda /s/ retention), Sevillian is the least conservative (metathesis), and Argentinian falls in the middle (debuccalization). Mexican and Argentinian Spanish grammars map /sC/ to [sC] and [hC], respectively, in production. Their grammars do not produce [Ch] forms; because there is no path between /sC/ and [Ch] in production, they cannot map [Ch] to /sC/ in perception. In contrast, the Sevillian grammar has a direct derivational path from /sC/ to [Ch] in production, so these listeners were able to 'undo' the steps of the mapping in the perception experiment.

Finally, I discuss what 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 phonotactic and phonetic (e.g. Trubetzkoy 1939; Riehl 2008), I briefly consider how the presence of sociolinguistic variation and phonological alternations may prevent single-segment representations. Both may provide evidence that [h] in [Ch] sequences is a realization of underlying /s/, and that it can belong to a distinct word or morpheme. These properties may make it nearly impossible for listeners to interpret [h] as a contrastive feature of the stop itself. The difficulty of constructing a sensible analysis with aspirated stops, as well as the oddness of the resulting phonological system, further suggest that this analysis—formally developed or posited by children—is implausible.

3.0.1 Roadmap

I begin by briefly discussing [Ch] sequences as clusters and segments, how Sevillian stop-h sequences fit into this classification, and some phonetic and perceptual properties of these sequences in Sevillian (Section 3.1). Then, I present the perception experiment showing that Sevillian listeners—but not listeners of other dialects—attribute [h] in [Ch] sequences to a preceding word (Section 3.2), and a serial analysis treating the differences in perception as the result of different grammars acting on the same underlying representations (Section 3.3). Finally, I suggest that the presence of variation and alternations in /sC/ realizations, as well as the difficulty of analyzing a system with

aspirated stops, makes positing underlying aspirated stops implausible (Section 3.4). Section 3.5 concludes.

3.1 Background

3.1.1 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* ([hC]), as in Icelandic (Haugen 1958; Thráinsson 1978; Arnason 1986; Suh 2001). I focus on sequences where [h] *follows* a stop, which are found in Sevillian Spanish. Finnish provides a straightforward example of stop-h sequences whose properties point towards a cluster analysis (Suomi et al. 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.

Tautosyllabic stop-h sequences can also be analyzed as either clusters or segments. Stop-h sequences can be analyzed as clusters in Acehnese (Austronesian), Old Khmer (Austroasiatic), and Taba (Austronesian). Acehnese [ph th kh] sequences can be split by an infix *-eun-* (13a) (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 how the parts can split by an infix.

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

a.	meu-k-eun-eu-hie	'it tastes a little bit like bad cooking oil'	Acehnese
	(from *khie*)		(Asyik 1987: 16)
b.	[kamhəŋ]	'anger'	Old Khmer
	(from [khəŋ] 'to be angry' + -m-)		(Schiller 1994: 314-5)

c.	[ˈmhonas]	'sick'	Taba
	[ˈnhik]	'bat'	(Bowden and Hajek 1999)
	[ˈkhan]	'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 (13b) (Schiller 1994; Minegishi 2006). Furthermore, the language allows many onset clusters, so analyzing [Ch] as a cluster fits with broader phonotactic patterns (Minegishi 2006). In Taba (13c), stop-h sequences can be analyzed as clusters because [h] forms onset sequences with many consonants, not just voiceless stops (Bowden and Hajek 1999). Furthermore, word-initial [Ch] sequences often arise from single-consonant prefixes, so the components come from different morphemes. Both phonotactic patterns and the fact that the components come from separate morphemes suggest a cluster analysis. Languages where tautosyllabic stop-h sequences are analyzed as clusters are rare, possibly because they would be confusable with aspirated stops.

In contrast, most English stop-h sequences ([ph th kh]) are usually analyzed as single segments: they are allophonically aspirated realizations of underlying voiceless stops. These sequences are tautosyllabic in monomorphemic contexts, and count as simple onsets.[1] For example, in a word like *intelligence* [ɪnˈthɛlədʒəns], [th] is tautosyllabic (*[ɪnt.ˈhɛlə.dʒəns]). English allows complex onsets, but not with [h] as the second consonant. Thus, [th] is best analyzed as a single segment, not a cluster. Analyzing [ph th kh] as underlying clusters would require an arbitrary exception in which /h/ could form clusters only with voiceless stops.

I now turn to Sevillian stop-h sequences, which are ambiguous between a cluster representation and an aspirated stop representation.

3.1.2 Spanish segments and clusters

Sevillian [Ch] sequences are ambiguous between a representation as /sC/ clusters, and an aspirated stop representation (/C^h/), closer to their current surface form. There are two related—but

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

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? I address each in turn.

First, I argue throughout this dissertation that stop-h sequences are realizations of underlying /sC/ clusters. Although the surface form has changed, the underlying form has not. Because /sC/ sequences are the origin of stop-h sequences, I briefly discuss the diachronic and synchronic relationship between them.

Stop-h sequences are diachronically and orthographically /sp st sk/ (<sp st sk>) clusters, where /s/ is a coda of the preceding syllable and /ptk/ is an onset. Stop-h sequences arise from coda /s/ reduction and metathesis. Coda /s/ reduction in Spanish is extensive (Hammond 2001), and variants—including debuccalization, gemination, and deletion—can be seen as steps on a reduction continuum (e.g. Mason 1994). This continuum is both diachronic and synchronic. 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). Today, coda /s/ variation appears to be stable in most varieties (Labov 1994: 584). Dialects differ in what variant is most common, but all dialects have socially and linguistically-conditioned variation between more and less reduced forms (see Erker 2012 for a thorough overview). I treat metathesis as part of the same reduction continuum because the underlying motivation is similar. Like deletion, it removes /s/ from coda position and leaves an open syllable (Moya Corral 2007; Moya Corral and Tejada Giráldez 2020). In contrast to most varieties where coda /s/ realizations are in stable variation, Ruch and Peters (2016) show that Sevillian coda /s/ variation is currently undergoing change.

The results of my experiment support this analysis of Sevillian [Ch] as deriving from an underlying /sC/ cluster. Assuming this representation, the analysis must map /sC/ to [Ch]. This is the analysis that I pursue in Section 3.3. A serial account, whose intermediate steps correspond

to forms with differing degrees of coda /s/ reduction, explains both diachronic and synchronic processes (further discussed in Section 3.3.1).

Second, is aspiration underlyingly contrastive? Because I treat [Ch] as a surface representation of an underlying /sC/ cluster, this question is not relevant. However, I believe it is important to consider the alternative that [Ch] is a contrastively aspirated stop, and return to this debate later (Section 3.4.2). 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, none of which have aspirated stops or three-way laryngeal contrasts. This 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.

3.1.3 Previous experimental studies on stop-h sequences

My perception experiment and analysis build on previous acoustic and perception work for Sevillian Spanish. The acoustic findings provide the basis for stimuli design; the perception findings provide the basis for the choice of participants and predictions.

Chapter 2 discusses previous studies of the phonetic properties of stop-h sequences at length (Section 2.1.2.1), as well as my own findings on these properties. What is crucial for my experimental design is that these sequences have long stop releases. Crucial to my analysis is that these sequences have gemination, that gemination appears to have arisen before metathesis in Western Andalusian Spanish, and that there is ongoing degemination in the sequences where the metathesis change is most advanced (see Section 2.1.2.1).

My perception experiment relies on one further observation: that metathesis occurs across word and morpheme boundaries (Ruch 2008; Horn 2013). The aspect of the design is crucial because the only way to show a segment is underlyingly present is through phonological alternations across boundaries. The rate and extent of metathesis differ by morphological context (Ruch 2008; Horn 2013), but what matters is that it occurs frequently in all contexts.

In perception, previous studies on Sevillian stop-h sequences establish that: (1) Sevillians perceive a distinction between stop-h ([VChV]) and intervocalic stops ([VCV]); (2) Sevillians are better than other listeners at distinguishing stop-h ([gakho]) from intervocalic stops ([gako]) and from other coda consonants ([gafko], [galko], 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.

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*. Listeners heard words from two continua and chose the corresponding orthographic representation. The continua were built off two base words, [pah$_1$th$_2$a] and [pata], and the duration of [h$_1$] and [h$_2$] were manipulated (for *pata*, [h$_2$] was inserted on [t]). Sevillian listeners—particularly younger ones—distinguished between *pata-pasta* based solely on the duration of [h$_2$] ([patha]). 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, which cannot be assumed to match underlying representations. [Ch] could be represented as a cluster (/sC/ → [Ch]) or an aspirated stop (/C^h/ → [Ch]), and these options cannot be distinguished with monomorphemic words. Listeners could map [patha] to /pasta/ or /pat^{h}a/, and both to orthographic <pasta>. My experiment avoids this limitation by using contexts where alternations at word boundaries give evidence for underlying /s/.

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] (← /pasta/ 'pasta') and [pata] (← /pata/ 'paw') and asked them to choose the corresponding orthographic representation. For words like [pahtha] ([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 respond to it, or identified it as another consonant.

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 /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 set-duration long [h$_2$]. In the second continuum, the same set of [h$_1$] steps were combined with a set-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 [$\underline{\text{h}_1}$C$\underline{\text{h}_2}$], then familiarity with /s/ debuccalization might help with perception in both [hC] *and* [Ch] sequences. Speakers of non-debuccalizing varieties—like Bedinghaus' (2015)—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 above perceptual findings motivate three listener groups for my perception experiment: Mexico ([sC], non-debuccalizing), Argentina ([hC], debuccalizing), and Seville ([Ch], debuccalizing and metathesizing). To preview, I did not find evidence that listeners hear [h] as a unit regardless of its location, in either this experiment or the ABX discrimination task in Chapter 5.

3.2 Fill-in-the-blank perception experiment

A binary forced-choice task was designed to test the following question: do listeners map [h] in [Ch] sequences to an underlying preceding /s/? Participants heard short sentences consisting of a verb and nonce word with the subject removed (e.g. *** *tiene p(h)ali* '*** has *pali*'). The sentences were ambiguous between the presence or absence of morphemic /-s/ on the verb, since the subject was removed, 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 [ph] 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.

3.2.1 Methods

3.2.1.1 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/ (14a vs. 14b).

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

a. Juan /tjene-∅ pali/.
Juan has-3SG *pali*.
Juan has *pali*.

b. Tú /tjene-s pali/.
You have-2SG *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 /ptk/ and /aiu/: *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 (14). 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.

In the experiment, listeners hear short phrases like (15): the subject is removed, the verb has no [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.

(15) Sample test sentence

Hear: *** [ˈtjene ˈp<u>h</u>ali] **Answer choices:** Juan Tú

The test stimuli sentences were created as illustrated in Table 3.1, starting from two sentences: one has underlying /s/ and one does not (Table 3.1a). The 3SG (no /s/) sentence is the base sentence (Table 3.1b). From this sentence, the subject was removed and replaced with a 30ms pure tone, inserted immediately preceding the verb (Table 3.1c). 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 | $\underline{\text{ˈphali}}_2$) (Table 3.1d). This ensures that all other acoustic cues—like closure duration and duration 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 form from the 3SG (no /s/) recording, and the nonce word from the 2SG (/s/) recording.

Table 3.1: Steps in stimuli creation. [p] = Step-0; [p^h] = Step-1; [ph] = Step-2/Naturally long.

a.	Two original sentences	1. Juan tiene $pali_1$ 2. Tú tienes $pali_2$	[ˈhuan ˈtjene ˈpali] [ˈtu ˈtjene ˈphali]
b.	(1) is base sentence	Juan tiene $pali_1$	[huan ˈtjene ˈpali]
c.	Subject in (1) replaced with tone	* 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^hali] [* ˈtjene ˈpali]

Next, an H-Step continuum was created by shortening [h] in nonce word-initial [ph th kh]. Recall that this [h] was naturally long because it was recorded a 2SG context (Table 3.1e). Continua were calculated independently for each nonce word sentence pair, as illustrated in Table 3.2. 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.[2] Step-1 is 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 3.2: H-Step continua for three nonce words

	VOT/H-duration in unmanipulated sentences		Manipulated H-duration		
	Intervoc. VOT	/sC/: [Ch]			
UR	/xuan tjene.../	/tu tjenes.../			
	[xuan tjene p_ali]	[tu tjene phali]	Step-0	Step-1	Step-2 (orig.)
/pali/	11.0ms	66.7ms	9.41ms	38.9ms	66.7ms
/tali/	17.2ms	82.7ms	16.1ms	49.4ms	82.7ms
/kali/	22.7ms	96.6ms	22.7ms	60.1ms	96.6ms

[2]Although metathesized [h] and VOT are not the same, the experiment shows that listeners accept shortened [h] as normal intervocalic stop releases.

The duration of post-posed [h] in my stimuli reflects the common cross-linguistic pattern for languages with aspirated stops, where 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—show this cline as well is notable (see Chapter 2, Sections 2.1.2.1 and 2.3.1 for further discussion about place of articulation).

Test sentences for one nonce word are illustrated in Table 3.3 (see Section 3.6 for 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 3.3: Set of test items for *pali*. ([p] = Step-0; [p^h] = Step-1; [ph] = Step-2)

Phrase recorded in	Word	Answer 1	Answer 2	Listeners hear
Juan tiene pali (3SG)	pali[Step-0]	Juan	tú	[* ˈtjene ˈpali]
Juan tiene pali (3SG)	pali[Step-1]	Juan	tú	[* ˈtjene ˈp^hali]
Juan tiene pali (3SG)	pali[Step-2]	Juan	tú	[* ˈtjene ˈphali]

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 in 3SG and 2SG sentences, with 2SG and 3SG answer choices. 50 filler sentences consisted of 5 nonce words (*pali, pumi, kuma, kina, tinu*) with different combinations of subject choices that made the answer unambiguous. For example, subjects heard '*** [tenemo phali]' ('1PL have *pali*'), and the answer choices were *tú* (2SG) and *nosotros* (1PL). The verb forms corresponding to the answer choices, <tienes> and <tenemos>, are different enough that the 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 on the internet 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.

3.2.1.2 Participants

Listeners of Sevillian (33), Mexican (30), and Argentinian (24) Spanish participated in the study; 8 were later excluded (see Section 3.2.1.4). 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. In Seville, 15 also had 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.

3.2.1.3 Predictions

The three 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 3.4). Gray cells in each column are the same across dialects.

Table 3.4: Properties of Argentinian, Mexican and Sevillian Spanish

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

All dialects have the same underlying representation for 3SG verbs (Table 3.4). 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: [xuan ˈtjene] vs. 2SG: [boh teˈneh]).[3] When Argentinian listeners hear [ˈtjene ˈphali], confusion may 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 3.5. The table also shows what the prediction would be

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

if [Ch] sequences were represented as aspirated stops. '≫' means strong preference, '>' means preference, and '∼' means no preference.

Table 3.5: Predicted response patterns (≫ indicates much greater; > indicates greater)

	Predicted under cluster analysis		
	Step-0	Step-1	Step-2
Mexico	3SG (no /s/) ≫ 2SG (/s/)	3SG (no /s/) ≫ 2SG (/s/)	3SG (no /s/) ≫ 2SG (/s/)
Argentina	3SG (no /s/) > 2SG (/s/)	3SG (no /s/) > 2SG (/s/)	3SG (no /s/) > 2SG (/s/)
Seville	3SG (no /s/) ≫ 2SG (/s/)	3SG (no /s/) ≈ 2SG (/s/)	3SG (no /s/) ≪ 2SG (/s/)
	Predicted under aspirated stop analysis		
Seville	3SG (no /s/) ≫ 2SG (/s/)	3SG (no /s/) ≫ 2SG (/s/)	3SG (no /s/) ≫ 2SG (/s/)

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.

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 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 work—is difficult for listeners of non-debuccalizing varieties (see Section 3.1.3).

Prediction 3: Argentinian listeners' choice of subject may depend somewhat on [h] duration, but I expect them to have a slight preference 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 3.1.3). 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.[4]

3.2.1.4 Statistical analysis

The data were analyzed in a logistic mixed-effects regression in R (RCoreTeam 2020), fit with the BOBYQA optimizer. The *lmerTest* package was used to calculate *p* values (Kuznetsova et al. 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 shown in Section 3.7). 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).

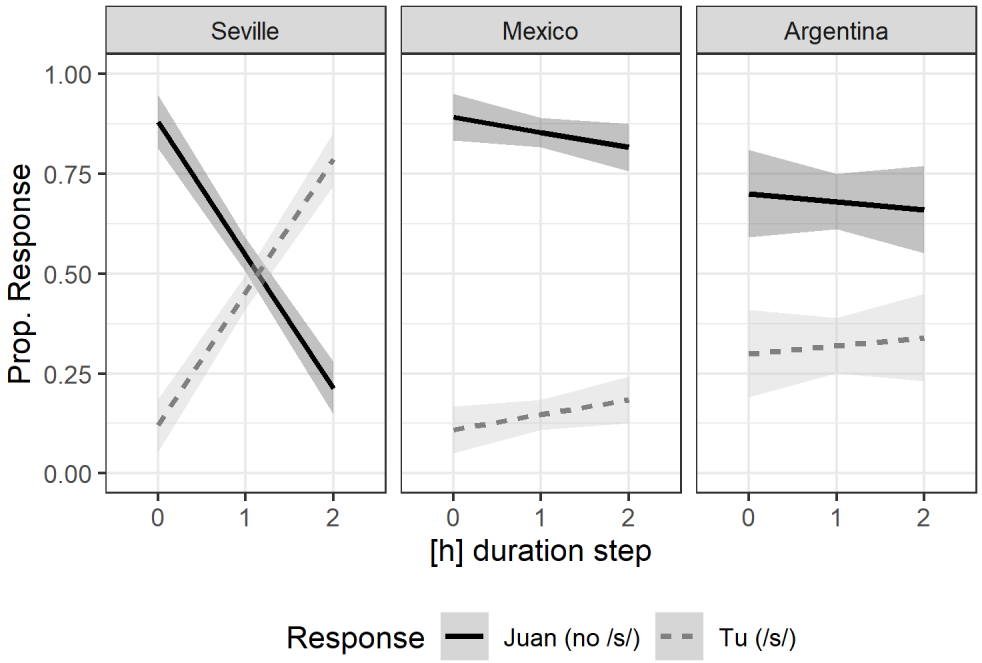
Eight participants were excluded. Four Sevillian participants were excluded for accuracy below 75% on the control items (50%-66% accuracy). Sevillians should perform well on these natural recordings of their native dialect. Two Mexican and two Argentinian listeners were excluded for accuracy lower than 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 and Argentinian listeners lack familiarity with these cues. Fillers had unambiguous answers.

[4]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.

3.2.2 Results

Sevillian listeners' responses depend on H-Duration Step (Figure 3.1, 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.[5] Sevillian listeners use the duration of [h] in [Ch] as information about the preceding word—specifically, about the morphological distinction between 2SG (/s/) and 3SG (no /s/) verb forms.

Figure 3.1: Response rate for listener groups at each [h] duration step



[5]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 sentence and the nonce word comes from a 2SG sentence. As reported for some varieties of EAS (Henriksen 2017), the 2SG verbs in my stimuli did show some vowel laxing in the wake of /s/ weakening (/tjenes/ → [tjenɛ]), as compared to 3SG verbs. This is somewhat unexpected, since laxing has not been reported in WAS, and the speaker who produced my stimuli spoke one of these varieties (Sevillian). 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.

Mexican listeners' responses (Figure 3.1, middle panel) differ drastically from Sevillian listeners'. They respond 3SG (no /s/) at rates higher than 75% at all h-duration steps, while 2SG (/s/) responses range from 10%-18%, even at Step-2 ([h] is 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 3.1, right panel) respond with 2SG (/s/) subjects between 30%-34% at all h-duration steps. They 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. Argentinian listeners do fall between Mexican and Sevillian listeners, but differently than expected. Instead of being a less extreme version of Seville, their responses are qualitatively like Mexican listeners' but are closer to chance.

The statistical modeling confirms these visual interpretations. The main model including all three regions (shown in Section 3.7, Table 3.15) 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 3.6), the effect of H-Duration Step is significant: Sevillians give more 2SG (/s/) responses at Step-1 than at Step-0 (β = 2.44, $p < .0001$), and more at Step-2 than at Step-1 (*emmeans* post-hoc test: β = -1.84, $p < .001$).[6] For Mexican listeners (Table 3.7), 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

[6]Negative estimates from *emmeans* indicate that the likelihood of a 2SG response is lower at one H-Duration step than at another. The directionality is reversed from the effects in the models, where negative estimates indicate higher likelihood of a 3SG response. This is because *emmeans* compares the likelihood of a 2SG response at two H-Duration steps, whereas the models predict the likelihood of a 3SG vs. 2SG at a given 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].

Table 3.6: Model 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 3.7: Model 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.[7]

3.2.3 Distribution of responses

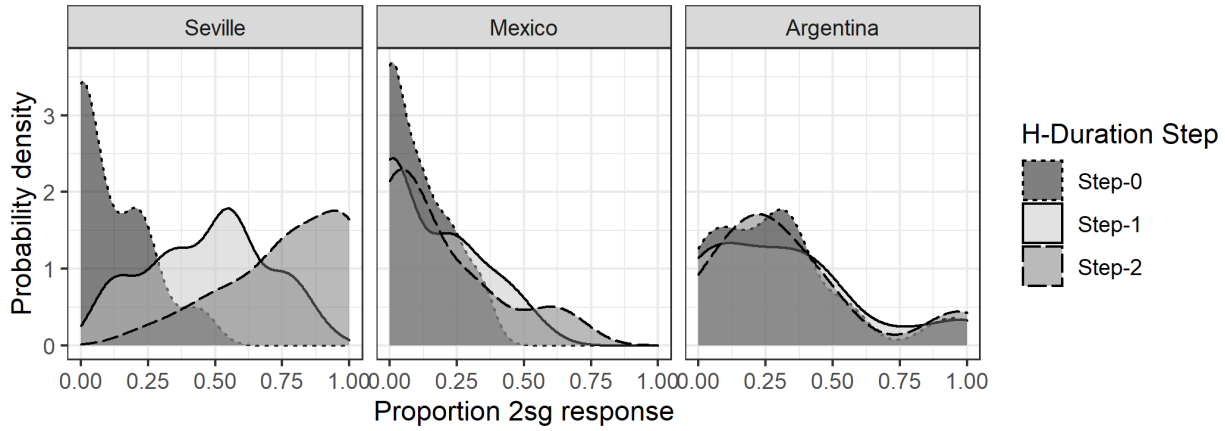
Neither Mexican nor Argentinian listeners perform like Sevillian listeners. But although their patterns are qualitatively similar, the distribution of responses differs. The large confidence intervals in the Argentinian data (Figure 3.1, right panel) could hide distinct individual patterns, or could reflect uncertainty at the level of the individual.

The density plots in Figure 3.2 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, partici-

[7]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 Region*H-Step*POA), the only statistically-significant effect is for /st/ in Seville at Step-2. That is, there is an 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 3.2). 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 (see Section 2.4.2 in Chapter 2). 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.

pants 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.

Figure 3.2: Density plots of 2SG (/s/) responses by H-Duration Step



3.2.4 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.[8] Mexican listeners do not consistently parse [h] in [Ch] sequences as underlying preceding /s/. Regardless of the duration of [h], they respond mostly 3SG (no /s/). 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

[8]Another interpretation would be that listeners parse it as a morphological distinction, but not underlying preceding /s/. I argue against this account in Section 3.4.3.

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 these listeners 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. They could also have contact with other debuccalizing or deleting varieties of Spanish, and may have learned to associate unfamiliar noisy productions with /s/.

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 those in Ruch and Harrington (2014). Most importantly, Argentinian listeners' native dialect uses different 2SG forms than those presented in my stimuli (see Section 3.2.1.3 and Table 3.4). The experiment stimuli had subtle phonetic differences between forms ([ˈtjene ˈpʰ̱ ali]-[ˈtjene ˈp̱hali]), and all of these forms differ drastically from what Argentinian listeners are most used to, in terms of the stress and first vowel of the verb ([teˈneh ˈpali]). 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 for them to connect it to the forms heard in the experiment. Both perception and lexical representations are necessary.

Finally, could the difference in [h] duration between the two studies may have affected Argentinian listeners' accuracy, since [h] tends to be shorter across word boundaries (/s#t/) than within words (/st/)? While plausible, this explanation does not explain their difficulty, because [h] in my stimuli ranges from 16ms-82ms for /s#t/ sequences, which is actually *longer* than [h] in Ruch and Harrington's (2014) monomorphemic /st/ word (29ms).

3.3 Analysis

The analysis is designed to capture the connection between perception and production. In the perception experiment, listeners of three Spanish dialects mapped the surface form they heard to underlying forms differently. I focus on behavior when [h] was longest and differences between dialects were clearest. I treat production as the inverse process of perception, in that speakers map from underlying forms to surface forms. My analysis treats the perception differences as a direct result of production differences. The groups have different phonological grammars that affect mappings in both directions, summarized in Table 3.8.

Table 3.8: 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]

The listener groups share underlying forms, but map them differently in both production and perception. Sevillian Spanish maps /tjenes pali/ to [tjene phali] in production. In perception, there is a direct path in their grammar from [tjene phali] back to /tjenes pali/ (2SG). Argentinian and Mexican listeners do not map /tjenes pali/ to [tjene phali] in production. In perception, when they hear this unfamiliar form, they map it to underlying /tjene pali/ (3SG). These listeners have no path between the metathesized surface form and the underlying 2SG verb form. Only the Sevillian grammar produces the surface form heard in the experiment, and only these listeners 'undo' the mapping to arrive back at the intended UR.

The mapping between /tjenes pali/ and [tjene phali] in Seville requires multiple steps (Table 3.9), some of which are steps in the Argentina and Mexico analyses as well. I first discuss the content of these steps, before implementing them in Harmonic Serialism, a derivational implementation of Optimality Theory (Section 3.3.1).

Table 3.9: Derivational steps from UR to surface forms in the three varieties

	Step	Seville	Argentina	Mexico
		/tjenes pali/	/tjenes pali/	/tjenes pali/
Step 1	Prosodification	[tje.nes$_{\mu}$.pa.li]	[tje.nes$_{\mu}$.pa.li]	[tje.nes$_{\mu}$.pa.li]
Step 2	Debuccalization	[tje.neh$_{\mu}$.pa.li]	[tje.neh$_{\mu}$.pa.li]	—
Step 3	Gemination & Metathesis	[tje.ne.pːh$_{\mu}$a.li]	—	—
Step 4	Degemination	[tje.ne.pha.li]	—	—

At Step 1, the candidates are assigned prosodic structure (moras and syllables, indicated by the period (.)). Then, coda reduction happens in steps. In Step 2, /s/ debuccalizes to [h]. Although chosen as optimal at this step, coda [h] is still problematic because [h] is a poor mora bearer. In Step 3, Sevillian Spanish geminates the following consonant, so that [h] is no longer the sole mora bearer. I assume that gemination places both segments unordered within the same timing slot. This is a somewhat non-standard assumption that I justify in Sections 3.3.2 and 3.3.6. At this same step, metathesis occurs. Instead of surfacing as unordered and overlapped, the overlap is resolved in favor of the metathesized order, and [pːh$_{\mu}$] is resyllabified as an onset (see Section 3.3.6 in this chapter, and Section 4.1.1 in Chapter 4 for further justification of syllabification). While treating metathesis as overlap may seem unusual, it has been proposed for Sevillian Spanish before (see Section 3.3.6) and is evident in the phonetics (Chapter 2). Metathesis-via-overlap has also been proposed for other languages (see Section 6.1 in Chapter 6).

Finally, at Step 4, the mora is lost (because [h] is no longer a coda), leading to degemination. The result is that coda /s/ and its mora have been removed, but the laryngeal feature of /s/ is retained as metathesized [h].

Metathesis and degemination can be seen as a way to satisfy coda restrictions. As already mentioned in Chapter 2 (Section 2.1), metathesis allows for an open syllable, without losing /s/ entirely (Moya Corral and Tejada Giráldez 2020). Mexican and Argentinian derivations converge much before metathesis would be possible.

The steps in the analysis represent within- and between-dialect differences.

Spanish dialects vary in the realization of coda /s/. While there is synchronic variation within dialects, each region tends towards specific variants: Mexican Spanish tends towards retention of sibilant [s], Argentinian Spanish tends towards debuccalization, and Sevillian Spanish tends towards metathesis. For the sake of simplicity, I treat the final step for each dialect as the categorical output. My analysis derives these variants for Mexican, Argentinian, and Sevillian Spanish, respectively. In constraint-based frameworks, variation can be derived in various ways (e.g. multiple grammars; see Coetzee and Pater 2011 for an overview of variation in phonological theory).

Before turning to the analysis proper, I describe the theoretical framework (Harmonic Serialism) (Section 3.3.1) and lay out my assumptions about gemination and metathesis (Section 3.3.2), which are relevant for all analyses, but most crucial for Seville. I also lay out my assumption that the gestures I have used to describe the metathesis change so far are tightly linked to the features manipulated in the analysis (Section 3.3.3). Then, I build up the analysis starting with Mexico, moving to Argentina, and finishing with Seville, the dialect whose derivation involves the most steps (Sections 3.3.4, and 3.3.5, 3.3.6, respectively). I finish with a brief summary (Section 3.3.7).

3.3.1 Harmonic Serialism and coda reduction

I cast the analysis in Harmonic Serialism (Prince and Smolensky 1993), a constraint-based framework related to Optimality theory that uses derivational steps. I follow McCarthy (2008a) in using this framework to analyze coda reduction. Derivations proceed by gradually removing features of the coda: [pa<u>s</u>ka] → [pa<u>h</u>ka] → [paka]. In HS, the requirement of *gradualism* means that the available candidates at each step are limited to ones that differ by a single change from the input. In [pa<u>s</u>ka], for example, the [s] cannot be deleted in one fell swoop. Place features must be deleted first; only then can the remaining feature delete, resulting in full segmental deletion.

What counts as a single change in HS is not settled; I follow McCarthy (2007) in defining a single change as a faithfulness violation. The analysis is affected in two ways by this definition.

First, (re)syllabification happens for free and can occur at any step, because it does not incur a faithfulness violation (McCarthy 2008b). Multiple syllabifications of the same input compete at each step, and I focus only on the plausible ones. Second, gemination and metathesis can occur at the same step, because I treat metathesis as violating only markedness constraints. I discuss this further in Section 3.3.2.

To control faithfulness to features, I follow McCarthy (2008a) in using MAX constraints rather than IDENT constraints to regulate features (see also Lombardi 2001; Farris-Trimble 2008). MAX constraints enforce gradual reduction because they are violated by deletion of individual features.

As shown in ((16)-(18)), different outcomes are obtained by slight rerankings of a small constraint set (McCarthy 2008a), so that the derivation converges at different steps and no further changes are harmonically improving. These constraints are defined in (28), (29), (38) and (39).

(16) Retention: MAX, MAX[PLACE], HAVEPLACE » CODACOND

(17) Debuccalization: CODACOND, MAX » MAX[PLACE], HAVEPLACE

(18) Deletion: CODACOND » HAVEPLACE, MAX[PLACE] » MAX

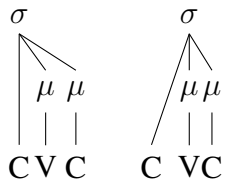
I focus on between-dialect synchronic variation, but the steps in the analysis also represent within-dialect variation and change over time along the lenition continuum. As McCarthy (2008a) points out, we should not be surprised to see debuccalization and deletion in variation because intermediate stages of lenition are part of the derivation. Since sound change over time grows out of synchronic variation (Ohala 1989), forms produced in a serial derivation provide options—and, ultimately, a path—for long-term change. Although there are few HS accounts of diachronic change (see Torres-Tamarit et al. 2012 for one), the connection between synchronic variation and diachronic change is captured in a serial derivation.

3.3.2 Gemination and metathesis

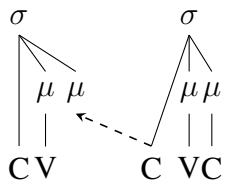
My analysis treats gemination and metathesis as inherently connected (Clayton 2010; Blevins and Garrett 1993). See also Section 6.3.2 for cross-linguistic discussion about the connection between gemination and [h].

Mechanics: I assume that gemination occurs when the root node of one segment is structurally linked to the mora of another ((19)-(20)). This kind of spreading is typical of moraic theories of gemination and compensatory lengthening (e.g. Hayes 1989), and has been previously applied to Spanish (e.g. Hualde 1989b; Martínez-Paricio and Lloret 2017; Martinez-Gil 2012).

(19) Moraic coda consonant



(20) Gemination via compensatory lengthening

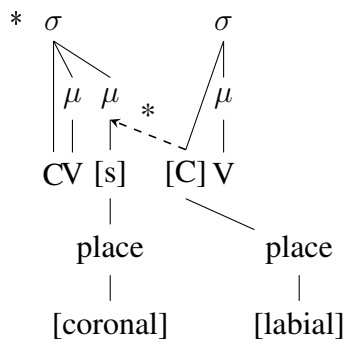


In terms of syllabification, I assume that the geminate consonant is in onset position, and [h] is in the coda of the preceding syllable ([h.Cː]).

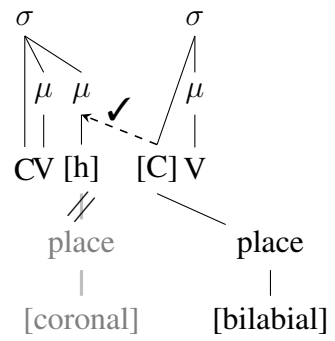
Motivation: Compensatory lengthening usually occurs once a segment has been entirely deleted. In my analysis, gemination is a process of compensatory lengthening triggered by the loss of place features from coda /s/, but not necessarily full deletion of the segment. When moraic coda [s] loses place features, [h] is left to bear the mora. The insertion of a structural link from the onset stop to the mora of /s/ saves [h] from being the sole mora bearer.

Although coda [s] and [h] are both obstruents, and thus poor mora bearers (Zec 1995), gemination can occur after coda /s/ reduces to placeless [h], but not if it is retained as place-ful [s]. I treat this as a result of the desire to avoid having multiple place features attached to the same timing slot. Gemination cannot occur in an [sp] sequence, because both [s] and [p] have their own place features that would be linked to the same moraic slot (21). Gemination can occur in [hp], however, because [h] has no place features (22). This is the basis of the ONEPLACE constraint, which I define in Section 3.3.4.

(21) No gemination with coda [s]
(e.g. *[tjenes$_{\mu}$p:ali])



(22) Gemination with coda [h]
(e.g. [tjeneh$_{\mu}$p:ali])



This connection between gemination and reduction reflects a generalization across Spanish dialects: gemination co-occurs with coda weakening (to place-less [h]) or deletion. Some dialects have gemination following coda /s/ debuccalization/deletion (/pasta/ → [pat:a] 'pasta') (Puerto Rican, Galarza et al. 2014; Cuban, Terrell 1979) and following /r l/ deletion (/kulpa/ → [kup:a] 'blame') (e.g. Cuban, Martinez-Gil 2012). Other dialects have debuccalization without gemination (/pasta/ → [pahta]) (e.g. Argentinian; Torreira 2006), but I have not found descriptions of dialects that have gemination *without* coda weakening (/pasta/ → *[past:a]). Dialects that weaken codas and geminate the following onset have a general dispreference for codas, and gemination appears to compensate for coda weakening or deletion.

Why gemination is important for metathesis: Gemination is crucial for my analysis because puts the two segments in the same timing slot (see tree in (20)), and I assume that they are unordered within it. If they are unordered within it, then surface order is determined by markedness constraints rather than faithfulness constraints to underlying order. The motivation to treat metathesis with markedness rather than faithfulness constraints is that gestural timing is not contrastive. If it is not contrastive, it is not subject to faithfulness constraints (following Hall 2003, a.o).

These assumptions about gemination and metathesis interact with the derivational steps of Harmonic Serialism. Gemination incurs a faithfulness violation (by inserting a structural link), while metathesis does *not* incur a faithfulness violation because it is controlled by markedness

constraints only. Because metathesis does not count as an extra step, it is available whenever geminated candidates are available.

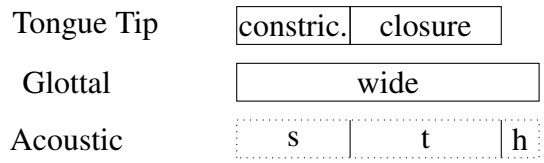
I save further discussion of gemination until the Seville Analysis (Section 3.3.6, where it is most relevant. For the sake of parallel presentation across dialects, I include geminated candidates in the Mexico and Argentina analyses, as applicable.

3.3.3 Gestures and features

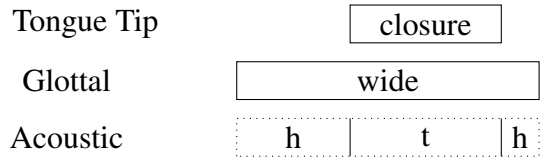
I have thus far discussed Sevillian metathesis as occurring via gestural retiming and overlap, as illustrated in the gestural diagram in Section 2.1.3. While the relationship between gestures and features is likely based on multiple factors (e.g. Zsiga 1993; Ridouane et al. 2011), I assume a tight link between featural specifications and gestures, along the lines of laryngeal realism (e.g. Honeybone 2005; Beckman et al. 2013; Schwarz et al. 2019). Specifically, I assume that [s] and [h] both have a wide glottal gesture, which I represent as [sg]. Additionally, [s] has a tongue tip constriction, represented with a place feature in the featural diagrams. For example, the featural representations for [st] (the structure in (32)) and [ht] (the structure in (33)) correspond to the gestural scores in (23). The wide glottal gesture in these diagrams is intended to represent a normal stop release, not metathesis.

(23) Gestural scores for [st] and [ht]

a. Sample gestural score for [st] sequence (with stop release)



b. Sample gestural score for [ht] sequence (with stop release)



The candidates and constraints in my analysis are framed in terms of features, and, for now, I assume equivalence between features and gestures. The constraints I use to control metathesis operate on the wide glottal gesture, which characterizes [h]. The constraints *HC, *CH are not framed gesturally, but I apply them as if they refer to alignment of this glottal gesture. In future research, they could be framed to operate explicitly on gestures instead (e.g. by preferring certain timing coordinations over others). This kind of account would require determining which parts of phonology operate on features vs. gestures, which is an open question beyond the scope of this dissertation. For my purposes, the phonology must manipulate gestures during metathesis, but the distinction is irrelevant elsewhere.

3.3.4 Mexico

Mexican Spanish listeners perform the mappings in (24) (repeated from Table 3.8a):

(24) Mexican Spanish listeners' mappings

a. Production: /tjenes pali/ (2SG) → [tje.nes.pa.li]

b. Perception: [tje.ne.pha.li] → /tjene pali/ (3SG)

In my analysis, the production grammar faithfully maps /tjenes pali/ → [tjenes pali], the most common realization in Mexican Spanish. There is no derivational path from /tjenes pali/ → [tjene phali]. The goal is to account for this production mapping, which determines listeners' behavior in perception.

The first step is prosodification, where the UR is assigned moraic and syllabic structure. Codas are assigned moras through Weight-by-Position (Hayes 1989), which requires coda consonants to be moraic. I implement this requirement with the constraints in (25) and (26), adapted from Rosenthall and van der Hulst (1999). *APPEND assigns violations for codas without moras, while *μ/OBS assigns violations for moraic coda obstruents.

(25) *APPEND: Assign a violation for a non-moraic syllable appendix.

(26) *μ/OBS: Assign a violation for a coda obstruent that is the sole sponsor of a mora.

Controlling mora assignment with constraints allows languages to assign moras to different sets of consonants, most often defined by sonority. More sonorous consonants are more cross-linguistically likely to bear moras, while less sonorous consonants (like obstruents) are the poorest mora bearers (Zec 1995). I treat all coda consonants in Spanish as mora-bearing because they pattern similarly in weight-sensitive stress patterns in the lexicon and in experimental work (Bárkányi 2002; Shelton 2007; Fuchs 2018; Chapter 4 of this dissertation).[9]

The tableau in (27) illustrates mora assignment and syllabification, which results in a fully faithful, prosodified winner (McCarthy 2007: 61). The ranking *APPEND ≫ *μ/OBS results in assigning a mora to the coda obstruent (b); this is better than having a mora-less coda (a). The syllable boundary is between [s] and [p]. Other possible syllabifications (e.g. [V.spV], [Vsp.V]) would be ruled out by high-ranked markedness constraints that disallow these types of syllable structures, which are generally not found in the language.

(27) Assigning prosodic structure (mora and syllabification)

	/tjenes pali/	*APPEND	*μ/OBS
a.	tje.nes.pa.li	*!	
☞b.	tje.nes$_{\mu}$.pa.li		*

Mora assignment works the same way in all dialects (and does not interact with the processes that *do* differ between dialects), so (27) is the first derivational step in the Argentinian and Sevillian grammars as well.

Two constraints control coda behavior. CODACOND (28) disprefers place features that are not linked to an onset, and MAX[PLACE] (29) requires that place features present in the input be present in the output..

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

[9]Readers may wonder about the difference between lexical /s/ and /s/ that is a morpheme. In words with lexical /s/ (e.g. /inˈgles/ 'English'), the syllable closed with /s/ attracts tress. In words where /s/ is a suffix (e.g. /pata-s/ 'paws'), it does not. Stress is pre-determined in morphological paradigms. I assume that both types of /s/ are assigned a mora, but a syllable can be heavy without attracting stress. Furthermore, in dialects where /s/ reduction triggers gemination, both lexical and morphemic /s/ pattern alike (Galarza et al. 2014). This suggests that they both occupy a moraic timing slot into which the following consonant can lengthen.

(29) MAX[PLACE]: Assign a violation for each input place feature lacking an output correspondent (McCarthy 2008a).

Recall from Section 3.3.2 that geminated candidates are available as part of the candidate set because they differ from the input by a single faithfulness violation: the insertion of a structural link to a mora. Recall also that gemination occurs in order to avoid a coda obstruent being the sole mora bearer, by sharing the mora with the following onset. The fact that coda /s/ is moraic is what allows for—and exerts pressure for—gemination. And as I will illustrate in the analyses of Mexico, Argentina, and Seville, gemination only occurs if doing so does not result in a single timing slot being linked to more than one set of place features. Gemination is optimal only ones [s] debuccalizes to [h].

The constraints in (30) and (31) formalize the motivations and restrictions on gemination. DEP penalizes the insertion of the structural link, and ONEPLACE penalizes mora slots that are linked to more than one set of place features. Recall also *μ/OBS (26), which penalizes moraic coda obstruents that are the sole bearers of a mora. While ONEPLACE and DEP are redundant in the Mexico analysis, both are necessary for Argentina and Seville. The ranking of ONEPLACE, DEP $\gg$ *μ/OBS prevents gemination while coda /s/ has place features.

(30) DEP-LINK: Assign a violation for structural association lines that are present in the output but not in the input (Fill-Link; Itô et al. 1995).

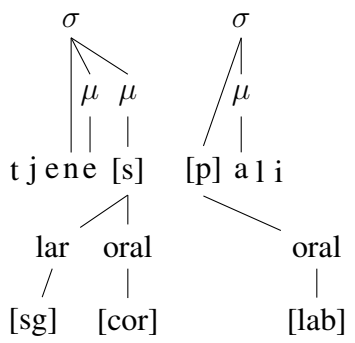
(31) ONEPLACE: Assign a violation for a mora slot linked to more than one set of place features.

The candidates available in the Mexico analysis are illustrated in (32)-(34). The fully faithful form (32) has moraic, place-ful, coda [s]. In the form with debuccalization (33), the place feature in the oral node has been delinked, leaving a placeless coda [h] defined only by a [spread glottis] laryngeal feature.[10] The form with gemination (34) has a moraic, place-ful coda [s], and the

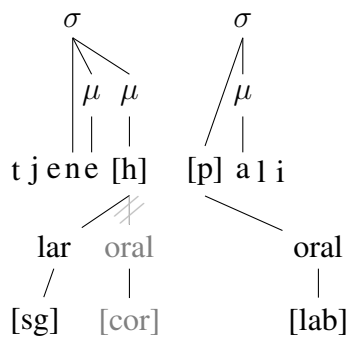
[10]This analysis differs from the typical assumption for Spanish, which is that fricatives are laryngeally unspecified because they do not contrast for voicing and stops do not contrast for [sg] (Beckman and Ringen 2009. However, some previous research on Spanish also assumes that /s/ is specified as [sg] (Widdison 1995; Vaux 1998; Morris 2000; Gerfen 2002: fn.11). For Sevillian Spanish, several findings support the [sg] specification of /s/: /s/ causes devoicing of surrounding segments (/desde/ → [deθ̱:e] 'since') (Hualde 1989a; Martinez-Gil 2012; Hualde and Colina 2014: 160-1), which also means it resists voicing assimilation common in other dialects (e.g. /desde/ → [dezðe], Martinez-

root node of [p] has spread to the mora of [s]. This is a possible candidate, but is disallowed since [s] still has place features, and the mora slot is thus linked to two sets of place features. The final structure, in (35), illustrates a candidate where both debuccalization and gemination have occurred. Gemination is allowed in this case, because debuccalization has removed place features from [h]. This candidate is not available in the Mexico analysis due to the HS requirement of gradualism, but I show it alongside the others because it *is* a candidate in the analyses for Argentina and Seville.

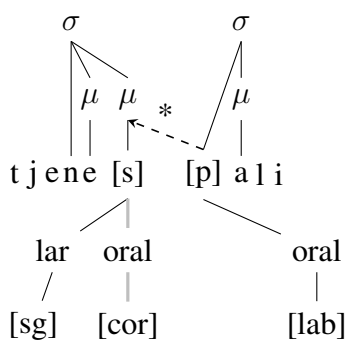
(32) Retention
[tjenes$_{\mu}$ pali]



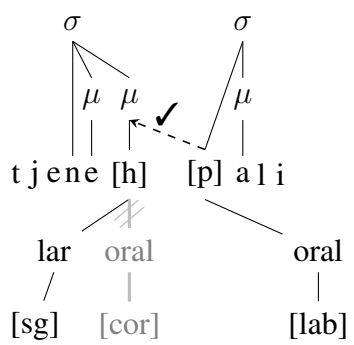
(33) Debuccalization
[tjeneh$_{\mu}$ pali]



(34) Gemination with coda [s]
*[tjenes$_{\mu}$pːali]



(35) Gemination with coda [h]
[tjeneh$_{\mu}$pːali]



In Step 1, the UR is prosodified (36a), as was shown in (27). The tableau in (36) shows the rest of the derivation. In Step 2, the principle comparison is between fully faithful candidate

Gil 2012; Hualde and Colina 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 [sg] for the sake of simplicity. While the specific constraints are affected by this assumption, the structure of the analysis does not hinge on it.

(a), with /s/ retention as [s], and candidate (b), with debuccalization. Candidate (a) wins over (b) because the ranking MAX[PLACE] ≫ CODACOND favors retaining places features.

(36) Mexico derivation

a. Step 1 | **Prosodification**: *APPEND ≫ *μ/OBS
/tjenes pali/ → [tje.nes$_{\mu}$.pa.li]

b. Step 2 | **Convergence (No debuccalization)**: MAX[PLACE] ≫ CODACOND
[tje.nes$_{\mu}$.pa.li] → [tje.nes$_{\mu}$.pa.li]

# 2	**tjenes$_{\mu}$ pali**	ONEPLACE	DEP	*μ/OBS	MAX[PL]	CCOND
☞a.	tje.nes$_{\mu}$.pa.li			*		*
b.	tje.neh$_{\mu}$.pa.li			*	*!	
c.	tje.ne(s$_{\mu}$.pː)a.li	*	*!			*

Output: tje.nes$_{\mu}$.pa.li **Derivation:** /tjenes pali/ → [tje.nes$_{\mu}$.pa.li]

The input thus maps to itself, and the derivation converges. No further changes are harmonically improving under this constraint ranking. A candidate with deletion is not available because full segmental deletion requires deleting multiple features, violating the gradualness requirement of HS.

For the derivation to continue, candidate (b) would need to win at Step 2. It does not, because the constraint that drives alternations—CODACOND—is too low-ranked to compel debuccalization. Deriving [ph] from /sp/ requires debuccalization and gemination as precursors to metathesis. In Mexican Spanish, CODACOND cannot even compel debuccalization.

The geminated candidate ((c) in (36), structure in (34)) merits some discussion, for the sake of parallel presentation across dialects. Gemination is available because it differs from the input by a single faithfulness violation: insertion of a structural link. Recall from Section 3.3.2 that gemination puts segments unordered within the same timing slot; I represent this with parentheses ([(s$_{\mu}$.pː)]). I mark syllabification as if the segments were linearized, with the geminate in the onset and [h] in the coda of the preceding syllable. I leave the mora marked on /s/ (its original affiliation), although /p/ is also attached to this timing slot.

Inserting the structural link in candidate (c) satisfies *μ/OBS by linking the moraic coda obstruent to another segment. But doing so violates high-ranked DEP and ONEPLACE, because the timing slot of coda [s] is linked to both its own place features and those of onset [p]. The ranking that prevents gemination while [s] has place features is ONEPLACE, DEP $\gg$ *μ/OBS. Candidate (c) also violates CODACOND because the place feature of [s] is not linked to an onset.

Recall that a geminated candidate means that a candidate with metathesis is available at the same step, since metathesis only violates markedness constraints. A candidate with gemination and metathesis (e.g. [tjenep:.s$_\mu$ali], and other syllabification options) will be ruled out by the same high-ranked constraints that disprefer candidate (c), since all have gemination and thus violate DEP. They will also be dispreferred by markedness constraints prohibiting gestural overlap, complex segments, and specific sequences across a syllable boundary, as I show in the Seville analysis.

In my perception experiment, Mexican listeners did not complete the intended mapping between surface [tjene phali] and /tjenes pali/ because there is no path between these forms in their grammar. There also may be a perceptual reason. Mexican listeners do not have experience perceiving [Ch] sequences, which do not exist in their variety. Their ability to perceive [Ch] sequences is tested in the ABX discrimination task in Chapter 5. That task does not require undoing metathesis or mapping to a lexical representation of a verb, and in that context, Mexican listeners have low accuracy perceiving [h] in [Ch] sequences. Mexican Spanish does not have aspirated stops or surface [Ch] forms, so there is no viable mapping from [Ch] to underlying segments or phonological categories. Together, results from the fill-in-the-blank and ABX tasks suggest that Mexican listeners do not perceive [h] in [Ch] sequences, even when perceiving [h] does not require connecting metathesis to an unmetathesized lexical representation.

3.3.5 Argentina

In production, Argentinian Spanish goes one step further than Mexican Spanish, mapping underlying coda /s/ to debuccalized [h]. Although one step closer to Sevillian Spanish in production,

Argentinian listeners behaved like Mexican listeners in the perception experiment. The mappings performed by Argentinian listeners are in (37) (repeated from Table 3.8b).

(37) Argentinian Spanish listeners' mappings

a. Production: /tjenes pali/ (2SG) → [tje.neh.pa.li]

b. Perception: [tje.ne.pha.li] → /tjene pali/ (3SG)

My analysis shows that grammar maps /tjenes pali/ → [tjeneh pali] in production; there is no path from /tjenes pali/ → [tjene phali]. Recall that the most common 2SG present tense form in Argentinian Spanish differs from that of Mexican and Sevillian Spanish (Section 3.2.1.3; Section 3.2.4). Despite this difference, if given /tjenes pali/ as a 2SG form instead, the phonology maps it to [tjeneh pali], with debuccalization.

Because Argentinian Spanish proceeds to debuccalization, additional constraints are needed to deal with newly available candidates that have full deletion. MAX (38) is violated by the deletion of a segment (a root node) (McCarthy 2008a). Recall that deletion of a segment can happen only when a single feature remains, due to the gradualism restriction on GEN in Harmonic Serialism (see Section 3.3.1). The conflicting constraint, HAVEPLACE (39), penalizes placeless segments.

(38) MAX: Assign a violation for deletion of a root node (McCarthy 2008a). Violated by deletion of a segment.

(39) HAVEPLACE: Assign a violation for a segment with no place specification (McCarthy 2008a: 279)

Like in Mexican Spanish, prosodification in Step 1 assigns a mora to coda [s] and syllabifies [s] into coda position (27). The tableau in (40) shows Steps 2 and 3. The candidates in Step 2 are the same as for Mexico, and the main comparison is between a fully faithful candidate (a) and one with debuccalization (b). Candidate (b) wins because CODACOND outranks MAX[PL] and HAVEPLACE (the opposite ranking of in Mexican Spanish). Candidate (c) is eliminated for the same reason as in Mexican Spanish: ONEPLACE and DEP outrank *μ/OBS. While Argentinian Spanish does not go all the way to metathesis, CODACOND is ranked high enough to compel

alternations, setting the stage for metathesis, in case the other constraints are ranked appropriately (in this case, they are not).

(40) Argentina derivation

a. Step 1 | **Prosodification**: *APPEND ≫ *μ/OBS
/tjenes pali/ → [tje.nes$_\mu$.pa.li]

b. Step 2 | **Debuccalization**: CODACOND ≫ MAX[PLACE], HAVEPLACE
[tje.nes$_\mu$.pa.li] → [tje.neh$_\mu$.pa.li]

c. Step 3 | **Convergence (No gemination)**: DEP ≫ *μ/OBS, HAVEPLACE
[tje.neh$_\mu$.pa.li] → [tje.neh$_\mu$.pa.li]

# 2	**tje.nes$_\mu$.pa.li**	ONEPLACE	DEP	MAX	CCOND	*μ/OBS	MAX[PL]	HAVEPL
a.	tje.nes$_\mu$.pa.li				*!	*		
☞b.	tje.neh$_\mu$.pa.li					*	*	*
c.	tje.ne(s$_\mu$.pː)a.li	*	*!		*			
# 3	**tje.neh$_\mu$.pa.li**	**ONEPLACE**	**DEP**	**MAX**	**CCOND**	***μ/OBS**	**MAX[PL]**	**HAVEPL**
☞d.	tje.neh$_\mu$.pa.li					*		*
e.	tje.ne(h$_\mu$.pː)a.li		*!					*
f.	tje.ne.pa.li			*!				

Output: tje.neh$_\mu$.pa.li **Derivation:** /tjenes pali/ → [tje.nes$_\mu$.pa.li] → [tje.neh$_\mu$.pa.li]

The output of Step 2 is passed to Step 3, in which the debuccalized candidate (d) wins. Although (d) has a moraic coda [h] that violates *μ/OBS, this is better than geminating the following consonant (e) or deleting [h] (f). The candidate with gemination (e), has [h] and [p] overlapped in the same timing slot. This candidate satisfies *μ/OBS because coda [h] is no longer the sole mora bearer, but is eliminated by higher-ranked DEP. Full deletion (f) violates high-ranked MAX.[11] The derivation converges after Step 3; no further changes are harmonically improving. CODACOND is ranked high enough to cause debuccalization in Argentinian Spanish, thus allowing for the possibility of gemination and metathesis. But high-ranked DEP prevents the derivation from continuing.

[11] No candidates in this step violate MAX[PL], because they are compared to the input to this step where [h] is already placeless, or CODACOND, because coda [h] is placeless.

As in the Mexico analysis, the geminated candidate (e) implies the availability of candidates that linearize [h] and [C] as either [hCː] or [Cːh]. These candidates are no better than (e), because any candidate with gemination fares worse than the winner on DEP. The markedness constraints that could decide between different candidates with gemination are irrelevant, because they are too low-ranked.

Because there is no path between underlying /tjenes pali/ and a metathesized [Ch] form, Argentinian listeners do not complete this mapping in perception either.

3.3.6 Seville

In the perceptiont ask, Sevillian listeners map the metathesized surface form back to underlying 2SG (/s/) forms, which I propose is because their grammars provide a a direct derivational path between these forms. The goal of the Seville analysis is to show that the production and perception mappings are symmetrical, unlike for Mexico and Argentina. The mappings are in (41) (repeated from Table 3.8c).

(41) Sevillian Spanish listeners' mappings

a. Production: /tjenes pali/ (2SG) → [tjene.phali]

b. Perception: [tjene phali] → /tjenes pali/ (2SG)

The derivation for Sevillian Spanish goes beyond debuccalization, to gemination, metathesis, and degemination. In Section 3.3.2, I discussed my assumption that gemination is a necessary component of metathesis because it puts two segments unordered within the same timing slot, allowing the surface order to be controlled by markedness (not faithfulness) constraints. I now justify these assumptions.

My assumption that [h] and [C] are unordered in a timing slot is reminiscent of proposals in which the components of complex segments share a timing slot. Both affricates (Lombardi 1990) and other types of complex segments (e.g. labiovelars, Sagey 1986) have been treated as consisting of underlyingly unordered features in a timing slot that are linearized on the surface. Sagey

(1986: 105) furthermore argues that some features that appear to be metathesized (e.g. labialization, palatalization) are underlyingly unordered, and are realized as metathesized for 'acoustic or articulatory' reasons. Results from my ABX task do not support metathesis for acoustic or perceptual reasons, but articulatory reasons are plausible. Finally, Takahashi (2019) proposes that CC metathesis occurs through fusion and fission. Two consonants first fuse together, and then break apart in the opposite order. While Takahashi's (2019) fusion and splitting operations are not framed articulatorily, they could easily be cast in those terms. Representationally, her 'fused' stage could be recast as gestural overlap, where the segments are unordered within a timing slot.

My assumption that gestural timing is a matter of markedness constraints comes from proposals that gestural timing is not contrastive (Hall 2003), and thus is not protected by faithfulness constraints. Instead, timing is determined by markedness constraints. This argument has also been made explicitly for laryngeal features. Across the world's languages, the timing of laryngeal features is never contrastive (Ladefoged and Maddieson 1996: 73). For example, for languages with preaspirated stops, many proposals argue that the timing of aspiration in relation to the stop (preaspirated, postaspirated) is not underlying (e.g. Anderson 1974: 267; Kehrein and Golston 2004; Bals Baal et al. 2012). Where the aspiration feature surfaces is determined by other factors, like syllable position and sonority (Golston et al. 2013). I do not treat Sevillian [Ch] sequences as complex segments, but I do treat them as passing through a stage where they are linked to the same timing slot and are unordered within it. This sets the stage for linearization, variations of which violate only markedness constraints.

The markedness constraints interacting to motivate metathesis are as follows. *OVERLAP prevents two gestures from being realized as completely overlapped, which would lead to poor recoverability (modeled on Hall's 2003 *GESTUREINGESTURE).

(42) *OVERLAP: Assign a violation for two segments that are overlapped in the realization. (Assumption: segments that are unordered in a timing slot surface as overlapped if pronounced as-is.)

(43) *HC: Assign a violation for a sequence of [h] followed by a voiceless stop [C].

(44) *CH: Assign a violation for a sequence of a voiceless stop [C] and [h]. (Not violated by single-segment, aspirated stops, or by overlapping segments)

The markedness constraint *HC deserves further justification. Sevillian metathesis has been argued to occur by re-timing of the glottal and closure gestures (Torreira 2006; Parrell 2012; Torreira 2012; Cronenberg et al. 2020). The gestural mechanism could also provide a motivation, by providing pressure for the in-phase coordination that results in the order [Ch] (see Section 6.2.1). But the motivation does not need to be gestural in order for a gestural *description* of the change to hold (see Section 2.1.3.2). Previous studies suggest that several aspects of the change may be perceptually motivated, but results from my ABX discrimination task in Chapter 5 (and those reported in Clayton 2010) show no support for this hypothesis. Instead, my *HC constraint could be justified in several ways, in addition to the gestural timing preference for in-phase coordination already mentioned. These typological and articulatory explanations are similar to those proposed for other markedness constraints like *NC (Pater 1999) and *VoicedGeminates (Ohala 1983; Kirchner 1998; Lombardi 1998; Hussain and Shinohara 2019).

- Coda /h/ is much rarer than onset /h/ cross-linguistically (Holmberg and Gibson 1979), and Davis and Cho (2003) argue that laryngeal gestures are only licensed in salient structural positions. While they make this argument for English, 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. Codas being of low salience is consistent with the general Spanish pattern of neutralizing voicing in coda position.
- In languages with aspirated stops, postaspiration is clearly the default. Preaspiration is typologically rare compared to postaspiration (Silverman 2003). Casserly (2012) uses this argument to motivate a constraint preferring that the [sg] gesture align with the right edge of a stop, resulting in postaspiration. Similarly, A *GLOTTAL-PLOSIVE constraint has also been used by Canfield (2015) and Takahashi (2019) to disprefer [h]-stop sequences.

The beginning of the Sevillian Spanish derivation is identical to that of Argentinian Spanish (tableau in (45)). In Step 1, prosodification results in a moraic coda /s/ and syllabification. In Step 2, the candidate with debuccalization (b) wins.

(45) Seville derivation

a. Step 1 | **Prosodification**: *APPEND ≫ *μ/OBS
/tjenes pali/ → [tje.nes$_\mu$.pa.li]

b. Step 2 | **Debuccalization**: CODACOND ≫ MAX[PLACE], HAVEPLACE
[tje.nes$_\mu$.pa.li] → [tje.neh$_\mu$.pa.li]

c. Step 3 | **Gemination**: MAX, *μ/OBS ≫ DEP
Metathesis: *OVERLAP ≫ *HC ≫ *CH
(Resyllabification)
[tje.neh$_\mu$.pa.li] → [tje.ne.pːh$_\mu$a.li]

d. Step 4 | **Convergence (Degemination)**
[tje.ne.pːh$_\mu$a.li] → [tje.ne.pha.li]

#2	**tje.nes$_\mu$.pa.li**	ONEPL	*μ/OBS	MAX	CCOND	DEP	*OVLP	MAX[PL]	HAVEPL	*HC	*CH
a.	tje.nes$_\mu$.pa.li		*		*!						
☞b.	tje.neh$_\mu$.pa.li		*					*	*	*	
c.	tje.ne(s$_\mu$.pː)a.li	*!			*	*	*				
#3	**tje.neh$_\mu$.pa.li**	ONEPL	*μ/OBS	MAX	CCOND	DEP	*OVLP	MAX[PL]	HAVEPL	*HC	*CH
d.	tje.neh$_\mu$.pa.li		*!						*	*	
e.	tje.ne.pa.li			*!							
f.	tje.ne(h$_\mu$.pː)a.li					*	*!				
g.	tje.neh$_\mu$.pːa.li					*				!*	
☞h.	tje.ne.pːh$_\mu$a.li					*					*
#4: Degemination		tje.ne.pha.li									

Output: tje.ne.pha.li **Derivation:** /tjenes pali/ → [tje.nes$_\mu$.pa.li] → [tje.neh$_\mu$.pa.li] → [tje.ne.pːh$_\mu$a.li] → [tje.ne.pha.li]

In Step 3, Seville diverges from Argentina. After debuccalization, available candidates include one with full deletion (e) and several with gemination, (f)-(h). Non-geminated candidates (d) and (e) are dispreferred by the ranking MAX, *μ/OBS ≫ DEP. Gemination is better than having

a moraic coda obstruent (d) or full deletion (e). Gemination is optimal at this step. Because [h] is placeless, there is no conflict with multiple sets of place features (no violation of higher-ranked ONEPLACE).

Among the candidates with gemination, markedness constraints determine the winner. The options are (f), which has unordered segments realized as overlapped; (g), which resolves overlap as [h.Cː]; and (h), which resolves overlap as metathesis [.Cːh]. The constraint ranking *OVERLAP ≫ *HC ≫ *CH chooses (h), with the metathesized order, as the winner. Linearizing (h_{μ}.pː) as [.pːh_{μ}] also involves resyllabification. I treat metathesized forms as onset sequences of two consonants that have resyllabified as an onset [V.pːh_{μ}] (see Section 4.1.1 for further justification).[12]

Degemination (Step 4) is the final step. Once [p] and [h] are syllabified as an onset cluster, the mora delinks. There is nothing left in coda position, and I assume that onset consonants are not moraic. Delinking the mora causes [p] to degeminate, leaving metathesized [ph] as an onset. Degemination can be seen as the ultimate step in coda /s/ weakening: all segmental and moraic material has been removed from coda position, but one identifiable feature has been retained: [sg], realized as [h]. While this solution differs from other dialects that delete /s/, it accomplishes the same goal. In Seville, recall that degemination is a change in progress among young Sevillians (Ruch and Harrington 2014; see also Section 3.1.3).

At Step 3, candidates with different syllabification would be possible. Geminate [p] could syllabify entirely as a coda and [h] as an onset ([tjenep:.h_{μ}ali]), but this candidate would be ruled out by constraints enforcing the language-wide dispreference for coda stops, which are almost nonexistent in the lexicon, and rarely surface unmodified when they do exist. While an onset cluster [ph] seems marked, it may be less marked than coda obstruents.

In short, because there is a path between underlying /tjenes pali/ and surface [tjene phali] in Sevillian Spanish, listeners in my experiment mapped [Ch] to /sC/ in perception too.

[12] I also assume that a high-ranked constraint against aspirated stops prevents a potential UR like /p^hali/ from mapping faithfully in Sevillian and other dialects of Spanish. Crucially, this constraint applies to underlyingly aspirated stops, but not to onset sequences like [Ch] derived from /sC/.

The ABX discrimination task in Chapter 5 follows up on Sevillian listeners' perception, to test whether they are better than other Spanish speakers at perceiving [Ch] sequences since they have experience with these surface forms. Results of that task show that they are not, suggesting that they are successful at the fill-in-the-blank task because their grammars map metathesized to unmetathesized forms, and they are able to connect metathesized forms to lexical representations with /s/. When given [Ch] in an underived context (the ABX task), they have no phonological category (aspirated stop representation) to map it back to, just like Spanish speakers of other varieties.

3.3.7 Analysis summary

In sum, Mexican, Argentinian, and Sevillian grammars map the same underlying /sC/ representations to different surface forms in production. Mexican Spanish is the most conservative, retaining coda /s/ as [s]. Argentinian Spanish goes one step further to debuccalization, and Sevillian Spanish goes all the way to gemination, metathesis, and degemination. The alternations needed to set the derivation in motion are driven by CODACOND. This constraint is ranked too low in Mexican Spanish to compel any alternations, leading the derivation to converge at coda /s/ retention. In Argentinian Spanish, this constraint *does* compel alternations (debuccalization), but DEP prevents gemination of the following consonant. In Sevillian Spanish, CODACOND compels debuccalization, and DEP is ranked low enough to allow gemination, which puts the segments in the same timing slot and thus allows markedness constraints to determine the surface order [Ch], which is then syllabified as an onset sequence. Finally, degemination represents the ultimate weakening of coda /s/, removing all material from coda position while retaining a single feature of /s/ in the onset of the following syllable. Crucial differences in constraint rankings between the dialects that lead to convergence at different steps are in Table 3.10.

Table 3.10: Crucial ranking differences between dialects

	Mexican	Argentinian	Sevillian
# 1 (prosodification)			
# 2	MAX[PL] ≫ CCOND Retention (Converges)	CCOND ≫ MAX[PL] Debuccalization	CCOND ≫ MAX[PL] Debuccalization
# 3		DEP ≫ *μ/OBS No Gemination (Converges)	*μ/OBS ≫ DEP Gemination *HC ≫ *CH Metathesis
# 4			Degemination

The differences in grammars between the dialects lead to differences in perception observed in my experiment. Sevillian listeners' grammars provide a path between the metathesized form used in the experiment [tjene phali] and an underlying 2SG form /tjenes pali/, but Mexican nor Argentinian listeners' grammars do not. When Mexican and Argentinian listeners hear a metathesized form like [tjene phali], they may not perceive [h], or they may perceive it but discard the information because they cannot map it to an underlying form.

3.4 Discussion and alternatives

In this section, I discuss two further reasons to believe that [Ch] sequences derive from /sC/ clusters, rather than underlying aspirated stops (Section 3.4.1). Then, I show what a system with underlying aspirated stops would look like, and argue that it is implausible (Section 3.4.2). Finally, I consider an alternative analysis in which [Ch] sequences arise through floating feature docking (Section 3.4.3).

3.4.1 Further arguments against [Ch] as an aspirated stop

The results from the perception experiment suggest that Sevillian listeners treat [h] in [Ch] sequences as belonging to the preceding word, not as part of an underlyingly aspirated stop. What

might prevent the aspirated stop analysis? Assuming that human learners construct phonological categories based on input that contains variable presence or realization of a feature, I speculate that the frequency of certain pronunciations affect the categories they create, and what they take to be the underlying form of that category. For example, an English-speaking child creates distinct categories for /s/ and /ʃ/, each with its own most canonical realization, even though phonetic realizations of the categories often overlap in acoustic space. Similarly, in Sevillian, if learners hear [Ch] sequences frequently enough and [hC] rarely enough, they could posit [Ch] as the underlying form of an aspirated stop (/C^h/) category.

[Ch] surface forms constitute a large proportion of the spoken forms children are likely exposed to. Ruch (2008) reports rates of surface variants of /st/ in Seville based on production data from 53 speaker in multiple speech styles (Table 3.11). She distinguishes a postaffricated variant of /st/ [t^s], which I group with [Ch] because both have the same (metathesized) linear order.

Table 3.11: 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, these percentages are high enough to suspect that learners *could* posit underlying /C^h/ based on their input. Providing a full learning model for this kind of learning is beyond the scope of this chapter, but the learning data contain enough metathesized

forms to make the aspirated stop analysis a possibility, and to suggest that there must be other factors preventing this representation. First, as discussed in Section 3.1.3, stop-h sequences arise in morphologically-derived environments, where the components belong to different morphemes. I believe additional factors preventing the aspirated stop analysis may include the presence of phonological alternations across boundaries (3.4.1.1) and sociolinguistic variation (3.4.1.2) .

3.4.1.1 Phonological alternations

Like 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/. The second question could be difficult, since the surface reflexes of /s/ vary in position.

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/ is produced as [s] in a single location corresponding to orthography ([Vs#V], [Vh#V]) (Table 3.12a). Preceding /ptk/, word- and morpheme-final /s/ can be realized before or after /ptk/ ([s#t], [h#t] vs. [th]) (Table 3.12b).

Table 3.12: Phonological alternations across morpheme boundaries

Variants	a. Prevocalic (/s#V/)	b. Preconsonantal (/s#ptk/)
Full [s]	[ˈmas ˈagwa]	[ˈmas ˈtorta]
Debuccalized [h]	[ˈmah ˈagwa]	[ˈmah ˈtorta]
Metathesized [th]	—	[ˈma ˈthorta]
Deleted	[ˈma ˈagwa]	[ˈma t(ː)orta]
	'more water'	'more cake'

I suggest the following reasoning as a plausible path for learners. Learners deduce that a word like /mas/ has an underlying final /s/, because it surfaces as [s] or [h] prevocalically and

sometimes preconsonantally. Now, 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/. This evidence for the /s#t/ cluster representation comes from multi-morphemic strings, but learners would transfer this knowledge to monomorphemic words like /pasta/. Although monomorphemic words lack alternations to provide clear evidence for /sC/, the evidence from sociolinguistic variation and orthography are compatible with this order, as I discuss in the next section. Phonologically-conditioned alternations provide evidence that stop-h sequences are clusters where /s/ precedes the stop, which may prevent learners from positing an aspirated stop representation.

To further investigate the hypothesis that socially-conditioned variation and phonological alternations could affect what learners posit as underlying representations, computational learning simulations—like those presented in Rasin and Katzir (2016)—could be helpful.

3.4.1.2 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 3.13.[13]

[13] 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 3.11).

Table 3.13: 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 these words produced with multiple variants, including coda [h] (Table 3.13b). When learners compare debuccalized forms (Table 3.13b) to faithful forms of the same word (Table 3.13a) they can deduce that [h] corresponds to [s p t]. From here, learners can conclude that underlying coda /s p t/ reduce to [h]. If they just considered /s/ and [h], /s/ could map to [h], or /h/ could map to [s]. By hearing words with different codas that surface as [h], learners see that the distribution of [h] is predictable: underlying /s p t/ map predictably to [h] in certain environments.

I suggest that variation gives 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 realizations of the same word—in other words, that they are surface realizations of the same underlying representation. Differently from phonological alternations, which require morpheme boundaries, these variation also shows learners that [Ch] forms are derived from /sC/ morpheme-internally.

Several existing approaches to learning posit that variability plays a crucial role (e.g. Dell 1981), and that the amount of surface variability can be a cue to differences in UR 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 (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; [h] is derivable.

3.4.2 Alternative: Aspirated stop analysis

In addition to variation and alternations, the phonological system resulting from an aspirated stop analysis is implausible. Before explaining why, I address the proposals that Sevillian [Ch] sequences are phonologizing into aspirated stops /pʰtʰkʰ/. I argue against this analysis throughout this chapter and dissertation.

Previous work has made several arguments, both phonetic and phonological, that [Ch] sequences 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*). However, I suggest that this argument does not hold: stop-h sequences may be bad onsets, but a heterosyllabic parse, where the stop is in the coda, would also result in sequences that are nearly-unattested in the language.

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/ → [ˈpa<u>th</u>a] → [ˈpa<u>ts</u>a]) (Moya Corral 2007; Ruch 2010; 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] (Harrington 2010: 104), and this high-frequency energy on [th] could easily be reinterpreted as [ts] without positing an underlying aspiration feature.

A third argument comes from duration. O'Neill (2009) argues for two representations for each word: one with an aspirated phoneme (/ˈpatʰa/ and another with an unaspirated phoneme (/ˈpasta/). In words with stop-h sequences, he finds that the preceding vowel is the same length

as when it is in an open syllable ([pat^{h}a] ∼ [ˈpata]). 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 sequence functions as a single segment that leaves the preceding syllable open, or a cluster in which the first segment closes the preceding syllable. To give one example, for NC sequences, preceding vowel duration is not a clear argument for segmenthood. Downing (2005) shows that some languages have vowel lengthening preceding NC sequences that are analyzed as prenasalized stops (e.g. Bantu languages), and this lengthening is argued to be compensatory. Once N vacates coda position and syllabifies as part of the following onset, the vowel lengthens. Other languages have vowel lengthening preceding NC sequences that are unambiguously clusters, where N is clearly a coda (e.g. Late Old English/Early Middle English, Jahore Malay) (Downing 2005: 186-8). This lengthening cannot be compensatory. These inconsistencies suggest that preceding vowel duration is not a good diagnostic of segmenthood.

A final argument put forth in favor of phonologization is that 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, but 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 believe this analysis is plausible for morpheme-internal sequences: [Ch] sequences could be analyzed as a single segment, and the distinction between /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 3.4.1), this kind of restructuring is plausible. In contrast, I believe that the aspirated stop analysis is 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 3.4.1.1).

I suggest that 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 /ptk/ and /p^{h}t^{h}k^h/. In this system, aspirated stops would contrast with unaspirated stops (46).

(46) 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 (46a), but is untenable word-initially (46b). Word-initially, the contrast would be neutralized, as illustrated in (47): unaspirated stops /ptk/ acquire [h] on the release allophonically, following words ending ending in /s/ (47b), so that they look just like underlying aspirated stops on the surface (47a). 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.

(47) Hypothetical aspiration contrast is neutralized word-initially

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

b. /tjenes paka/ → [ˈtjene p^{h}aka]

The distribution of aspirated stops illustrated in (47) would also be typologically odd, given how laryngeal contrasts usually work. Sevillian aspirated stops would be limited to morpheme-internal, word-medial position. This distribution for laryngeal contrasts is cross-linguistically uncommon (Lombardi 1991: Ch.3; Lombardi 2001). More frequently, a fuller range of laryngeal contrasts is available in word-initial position than in other positions.

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).

3.4.3 Alternative: Floating features

A very different possibility is to analyze the [h] in h-stop and stop-h sequences as a floating [+sg] feature, like in other floating feature analyses (Zoll 1996; Wolf 2007). Floating feature analyses are often proposed for affixation processes like Celtic mutation. In Irish, some words trigger lenition mutations on the first segments of following verbs, and some prepositions trigger lenition mutations on the initial consonant of following nouns (see Wolf 2007 for overview). In Wolf's (2007) analysis, a morpheme carries a floating feature that docks on the first consonant of the following morpheme. Docking is controlled by constraints that prefer floating features to dock (do not delete them or leave them floating), but disprefer docking on segments that already have that feature, or that are part of the same morpheme.

To apply this to Sevillian Spanish, let's say that [h] in [Ch] sequences is the realization of a floating [+sg] feature that docks on the following stop. For example, the present tense 2SG present tense morpheme is floating [spread glottis], and a verb <tienes> is represented as /tjene+$^{[sg]}$/. [sg] is realized as [h] in [Ch] preceding a /ptk/-initial word (48a), and as [s, h] preceding a vowel-initial word (48b). The choice between [s] and [h] preceding a vowel is sociolinguistically and stylistically conditioned.

(48) Floating feature analysis (argued against)

a. /tjene+$\underline{^{[sg]}}$ pata+$^{[sg]}$/ → [ˈtjene ˈp$\underline{^{h}}$ ata] 'you have paws'

b. /tjene+$\underline{^{[sg]}}$ ala+$^{[sg]}$/ → [ˈtjene$\underline{\text{h}}$ ˈala] 'you have wings'

The main conceptual issue with this style of analysis is the following: what would compel a floating [sg] feature to acquire place features and surface as [s] in some circumstances? Constraints also typically militate against floating features docking on a morpheme that already contains the feature. Sevillian [sg] cannot be subject to this restriction. If all morphemic [h] is represented as

/[sg]/, then many of the morphemes it docks on will contain the same feature because of number agreement.

While constraints typically used for floating features seem to rule out this analysis for Sevillian stop-h sequences, the biggest issue is a conceptual one. The change from h-stop to stop-h affects instances of [s] that belong to other morphemes (/mas torta/), as well as ones that constitute affixes (/tjene-s torta/). Under a floating feature analysis, words where /s/ belongs to another morpheme would have to be represented with floating [sg] instead of segmental /s/. While a specific morphological distinction like 2SG could plausibly be [sg], it seems less likely that multiple morphological distinctions are marked with the same floating [sg], and that all words ending in lexical /s/ instead have floating [sg]. Floating [sg] would have no consistent meaning—or even be consistent in whether it marks morphological distinctions.

3.5 Conclusion

In this chapter, I have argued that Sevillian surface [Ch] is the realization of an underlying /sC/ cluster. 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 make use of [h] in stop-h sequences. The formal analysis treats the mapping from /sC/ to [Ch] as gradual coda reduction, and establishes a connection between production and perception. The Sevillian grammar has a direct mapping between underlying /sC/ and metathesized [Ch] forms in production, so they do the same mapping backwards in perception. In contrast, Argentinian and Mexican grammars have no mapping in production from a UR to a metathesized form, and thus do not map the metathesized form to the intended UR in perception, essentially ignoring the presence of [h]. I also suggested that stop-h sequences may not be phonologizing into aspirated stops due to the presence of sociolinguistic variation and phonological alternations, which give learners evidence for the cluster representation. Finally, I argued that the phonological system implied by the aspirated stop analysis is implausible.

3.6 Fill-in-the-blank full list of stimuli items

Table 3.14: Full illustration of test, control, and filler items for *pali*. [p] = Step-0; [p^h] = Step-1; [ph] = Step-2/Naturally long.

Type	Phrase recorded in	Word	Answer 1	Answer 2	Listeners hear
Test	Juan tiene pali (3SG)	pali[Step-0]	Juan	tú	[* ˈtjene ˈpali]
(3)	Juan tiene pali (3SG)	pali[Step-1]	Juan	tú	[* ˈtjene ˈp^{h}ali]
	Juan tiene pali (3SG)	pali[Step-2]	Juan	tú	[* ˈtjene ˈphali]
Controls	Juan tiene pali (3SG)	pali[orig]	Juan	tú	[* ˈtjene ˈpali]
(2)	Tú tienes pali (2SG)	pali[orig]	Juan	tú	[* ˈtjene ˈphali]
Fillers	Juan tiene pali (3SG)	pali[orig]	Juan	nosotros	[* ˈtjene ˈpali]
(10)	Juan tiene pali (3SG)	pali[orig]	Juan	ellas	[* ˈtjene ˈpali]
	Tú tienes pali (2SG)	pali[orig]	tú	nosotros	[* ˈtjene pˈhali]
	Tú tienes pali (2SG)	pali[orig]	tú	ellas	[* ˈtjene ˈphali]
	Nosotros tenemos pali (1PL)	pali[orig]	nosotros	tú	[* teˈnemo ˈphali]
	Nosotros tenemos pali (1PL)	pali[orig]	nosotros	Juan	[* teˈnemo ˈphali]
	Nosotros tenemos pali (1PL)	pali[orig]	nosotros	ellas	[* teˈnemo ˈphali]
	Ellas tienen pali (3PL)	pali[orig]	ellas	tú	[* ˈtjenen ˈpali]
	Ellas tienen pali (3PL)	pali[orig]	ellas	Juan	[* ˈtjenen ˈpali]
	Ellas tienen pali (3PL)	pali[orig]	ellas	nosotros	[* ˈtjenen ˈpali]

3.7 Fill-in-the-blank full model results

Table 3.15: 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	***

CHAPTER 4

Opaque interaction between metathesis and stress

4.0 Introduction

The results of the fill-in-the-blank study in the previous chapter suggest that Sevillian listeners treat [Ch] sequences as separable, mapping them to underlying /sC/ clusters. How does the rest of the phonology treat [Ch] sequences, given that their underlying and surface forms differ? In this chapter, I investigate the interaction between stress and metathesis in Sevillian Spanish. Phonological patterns can be transparent or opaque. In transparent patterns, the environment conditioning the change is visible on the surface. In opaque patterns, a process has applied but its conditioning environment is not visible in the surface form. One commonly cited example is the interaction between Canadian Raising and flapping. In North American English, *writer* and *rider* surface with different initial vowels conditioned by the voicing of the following stop, even though the voicing distinction neutralizes to a flap on the surface. *Writer* (/rɑ̲ɪtər/) surfaces as [rʌ̲ɪɾər], while *rider* (/rɑ̲ɪdər/) surfaces as [rɑ̲ɪɾər]. Canadian Raising changes /ɑɪ/ to [ʌɪ] before voiceless stops *before* the flapping rule turns /t, d/ to [ɾ], which removes visible conditioning on the surface (McCarthy 2007). There is, however, debate over whether opacity is real and productive in this case, and more

broadly. For Canadian Raising, McCarthy (2007) and Idsardi (2006) argue that it is, but Mielke et al. (2003) argue that the interaction between flapping and raising is not actually opaque, and can be described through other mechanisms. Sanders (2003) argues that opaque patterns in several languages are unproductive, can be reanalyzed as transparent, or are only productive at morpheme boundaries. The experiment presented in this chapter shows that the interaction between stress and metathesis in Sevillian is opaque, and that listeners apply this opaque interaction to nonce words.

The opacity debate is particularly important for constraint-based frameworks. If opacity is real and productive, a phonological theory must be able to account for it. Parallel Optimality Theory (Prince and Smolensky 1993) cannot account for certain types of opacity, because it evaluates all candidates based only on their surface form and relationship to the underlying form.[1] Serial theories handle opacity better, because they allow intermediate derivational stages where a candidate can be selected as optimal before the conditioning environment disappears. While basic serial versions of OT like Harmonic Serialism (HS) were not designed to deal with opacity, slight modifications to HS can (McCarthy 2007; Jarosz 2016).

Metathesis and stress in Sevillian Spanish have the potential to interact in the following way. In Spanish, antepenultimate stress generally is not found in words with heavy penults (*[CV́.CVC.CV]) (Harris 1983; Roca 1991). If [Ch] sequences have been reanalyzed as single consonants, nonce words like [giˈ.na.ka.pho] should be acceptable, because the penult is light.[2] If, in contrast, [Ch] sequences have not been reanalyzed and are still represented as underlying /sC/ clusters, then nonce words with metathesis out of the penult ([giˈ.na.ka.pho]) should be as unacceptable as those with a surface-heavy penult ([giˈ.na.kam.po]).

The experiment thus has two goals: (1) to test the productivity of an opaque interaction; (2) test the representation of [Ch] sequences, by probing their representation at the moment of stress assignment. Regarding (1), the experimental results show that Sevillian listeners extend the opaque interaction to nonce words. As for (2), the experimental results suggest that the penult in

[1]Some types of opacity, like counterfeeding, can be modeled in parallel frameworks, but it requires a rich set of faithfulness constraints (McCarthy 2007).

[2]There are both phonotactic (Section 4.1.1) and analysis-internal (Section 4.5.4) reasons to believe that [Ch] sequences syllabify as onsets.

[gi'.na.ka.pho] words is closed at the level of representation where stress is assigned, supporting support the /sC/ cluster representation of [Ch] sequences. At that level, /s/ is still in the coda of the penult. The interaction between stress and metathesis is counterbleeding opacity (Kiparsky 1973): stress is constrained by a conditioning environment that is not present on the surface (McCarthy 2007: 24). The serial analysis corresponds to these results, assigning stress early in the derivation, before /s/ metathesizes out of the coda.

4.0.1 Roadmap

I begin by discussing my assumptions about the syllabification of [Ch] sequences, and Spanish stress from descriptive and experimental perspectives (Section 4.1). Then, I present the set-up of the current experiment (Section 4.2) and two preliminary studies confirming the validity of the stimuli (Section 4.3). Section 4.4 presents the main experiment. Finally, I analyze the interaction between stress and metathesis in Harmonic Serialism (Section 4.5). Section 4.6 concludes.

4.1 Background

4.1.1 Syllabification of [Ch]

The set-up of my experiment and interpretation of the results rely on several assumptions about syllabification. I will assume throughout this dissertation that Sevillian stop-h sequences syllabify as tautosyllabic onsets ([.Ch]), not heterosyllabic clusters (*[C.h]). This is for both phonotactic and analysis-internal reasons.

/s/-voiceless stop clusters are the historical forms corresponding to stop-h sequences. In terms of phonotactics, [s]-voiceless stop sequences are heterosyllabic ([Vs.CV]); these clusters are illicit onsets in Spanish. Coda [s] thus contributes weight to the preceding syllable.

For [Ch] sequences, the heterosyllabic parse (*[C.h]) is bad for several reasons. First, word-medial coda obstruents are dispreferred, which can be seen in lexical patterns and lenition. Some

obstruents are frequent in word-medial position (e.g. /s/), but are often reduced when they exist. Other obstruents are much less frequent (or non-existent) as word-medial codas (e.g. /ptkbdg/). These obstruents are also often reduced, and the voicing contrast is lost in this position (Campos-Astorkiza 2012; Colina 2012). This dispreference for coda obstruents is widespread in Romance languages (Vennemann 1988), and a preference for open syllables has been offered as a reason for metathesis (Moya Corral and Tejada Giráldez 2020; Moya Corral 2007) (see Section 2.1). Furthermore, voiceless coda stops are not followed by an onset [h] in the lexicon. Finally, the heterosyllabic parse *[C.h] cannot be right because this coda stop never reduces. In Andalusian Spanish, coda stops debuccalize and metathesize (see Chapter 2, Sections 2.1.2.4 and 2.1.2.4), so if the heterosyllabic parse were right, we would expect the coda stop to reduce.

Syllabifying [Ch] as an onset results in different syllable structures in [sC] and [Ch] forms (Table 4.1). [sC] clusters are heterosyllabic, and create a heavy syllable closed by [s]; [Ch] sequences are tautosyllabic, and are preceded by an open syllable on the surface.

Table 4.1: Metathesis changes surface syllable structure

SC		CH		
[pas.ta]	HL	[ˈpa.tha]	LL	‘pasta’
[es.pa.ɲa]	HLL	[eˈ.pha.ɲa]	LLL	‘Spain’
[es.pe.θja.li.ðað]	HLLLH	[e.phe.θja.liˈ.ðað]	LLLLH	‘speciality’

While it may seem odd to consider [ph th kh] as onset cluster, recall that these sequences are possible in other languages (e.g. Acehnese, Old Khmer, and Taba; Chapter 3, Section 3.1.1), without being aspirated stops. There are further, analysis-internal reasons to prefer this syllabification over the heterosyllabic one. I leave discussion of those until Section 4.5.4, where they are better illustrated with the data at hand.

4.1.2 Spanish stress

The change in syllable structure caused by metathesis could affect stress because Spanish stress is partially weight sensitive in non-verbs. I discuss only non-verbs, because verbal stress is morpho-

logically conditioned (Roca 1990b). Spanish primary stress falls in a right-aligned three-syllable window and can be penultimate, antepenultimate, or final (Table 4.2). Penultimate stress is most common, particularly for vowel-final words. Some words have antepenultimate stress, others have final stress. Words with final stress mostly end in consonants /d r n l s/, but can also be vowel-final (Harris 1983; Harris 1991; Roca 1991; Hualde 2012).

Table 4.2: Stress patterns in Spanish

# syllables	Pen.		Antepen.		Final	
3	kaˈðera	'hip'	ˈsaβaðo	'Saturday'	koliˈbɾi	'hummingbird'
4	esˈpaɲa	'Spain'	teˈlefono	'telephone'	aktiβiˈðað	'activity'
5	eleˈɣante	'elegant'	demoˈkɾatiko	'democratic'	puɾifikaˈsjon	'purification'

Orthographically, penultimate stress is unmarked (e.g. *cadera* 'hip'). Antepenultimate stress is always marked with an acute accent (e.g. *sábado* 'Saturday'), and final stress is marked on words ending in a vowel, <n> or <s> (e.g. *colibrí, purificación* 'hummingbird', 'purification').

Spanish stress is uncontroversially contrastive and controversially weight sensitive. Some words contrast only in the location of stress, indicating lexically-marked stress (e.g. /ˈsabana/ 'bed sheet' vs. /saˈbana/ 'savannah') (Harris 1991). In addition to lexically contrastive stress, several restrictions in the lexicon indicate weight sensitivity. For example, heavy syllables—those with a coda or diphthong—show a tendency to attract stress. These restrictions were noted early in Harris (1983), and have received extensive attention since. One of these restrictions is in (49):

(49) *ĹHσ: A heavy penult or ultima prevents antepenultimate stress (Harris 1983: p. 85, 88-90; Roca 1991; Harris 1992; Baković 2009; Baković 2016).

The *ĹHσ restriction means that a heavy penult or ultima narrows the stress window to the final two syllables by preventing antepenultimate stress. While there are exceptions to this generalization, many of these words are place names or loanwords (e.g. *Mánchester, Wáshington, Ámsterdam*; examples from Roca 1990a). These words arguably are not a direct result of the language's productive grammar: in addition to rare stress patterns, they violate Spanish phonotactics (e.g. Spanish disallows word-final /m/). The generalization in (49) is also violated by some native words (e.g.

régimen ‘regime’) (Harris 1983; Roca 1990a; Baković 2016). These words are difficult for many accounts of Spanish stress, but they are few and are clearly exceptions to a strong tendency.

The *ĹHσ restriction has been confirmed in corpus studies of the Spanish lexicon. Furthermore, the data show that, in preventing antepenultimate stress, heavy penults and ultimas attract stress to themselves. Bárkányi’s (2002) study of stress patterns in the *Diccionario de la Real Academia Española* makes several points, based on non-verbs of three or more syllables (Table 4.3). When the final two syllables are light, stress is mostly penultimate (85%) (Table 4.3a). When the penult or ultima is heavy and the other is light, stress is overwhelmingly on whichever is heavy (99.38% for a heavy penult; 81.21% for a heavy ultima) (Table 4.3b, c). Only .02% of words with heavy penults have antepenultimate stress (Table 4.3c, in gray). When both the penult and ultima are heavy, stress is mostly final (77%) (Table 4.3d). In short, a heavy penult or final almost always attracts stress to itself if the other is light. If both are heavy, the ultima is usually stressed.

Table 4.3: Stress patterns in Spanish by syllable type (adapted from Bárkányi 2002: 383).

Syllable structure			Antepen	Penultimate	Final	Total
a. ...CV.CV	...taka	LL	12.96%	**85.68%**	1.35%	2261
b. ...CV.CVC	...takan	LH	3.53%	15.26%	**81.21%**	5184
c. ...CVC.CV	...tanka	HL	.02%	**99.38%**	.59%	4550
d. ...CVC.CVC	...tankan	HH	.56%	22.4%	**77.04%**	710

Regardless of weight, antepenultimate stress is rare. Bárkányi’s (2002) numbers indicate that approximately 9% of nouns have antepenultimate stress, 75% have penultimate stress, and 15% have final stress (similar to numbers reported by Núñez-Cedeño et al. 2014). According to Bárkányi’s (2002) calculations, antepenultimate stress is most frequent when the final two syllables are light (12.96%), and rare when either the penult or ultima is heavy (<4%). Fuchs (2018: 9) finds even more extreme results in the EsPal corpus (Duchon et al. 2013): there were no words with heavy penults and antepenultimate stress (*CV.ĆV.<u>CVC</u>.CV; *CV.ĆV.<u>CVV</u>.CV). His search must exclude words like *Wáshington*, which tend to be loanwords. Neither of these studies look at the effect of weight of the *antepenultimate* syllable. For Portuguese, which is very closely related to

Spanish and obeys the same stress restrictions, Garcia (2017) found that heavy antepenults actually do not attract stress in the lexicon (in fact, they repel it).

While some theoretical proposals for Spanish argue against weight-sensitivity (Roca 1990a; Roca 1991 et seq., and Piñeros 2016), there is evidence that speakers generalize a weight-sensitive pattern in experiments, which I review next.

4.1.3 Experimental studies on Spanish stress

In multiple experimental tasks, Spanish-speaking participants apply weight-sensitive stress to nonce words, and show sensitivity to nonce words that violate these patterns. In a written production task, Waltermire (2004) finds that the weight of the final three syllables affects stress placement in nonce words. Weight had the strongest effect for penultimate and final syllables: heavy penults and ultimas prevented antepenultimate stress. Shelton and colleagues also find that weight sensitivty is an active restriction (Shelton 2007; Shelton et al. 2009; Shelton et al. 2012; Shelton and Grant 2018). For example, Shelton (2007) asks Spanish speaking participants to read nonce words one at a time off a computer screen. He finds that they perform worse with nonce words that violate legal weight-stress patterns. They make more mistakes in production with words like *dóvalda* (antepenultimate stress, heavy penult, violating *ĹHσ) in comparison to words like *dóvasa* (antepenultimate stress, light penult, phonotactically good) and *dováldа* (penultimate stress, heavy penult, phonotactically good).

Finally, Fuchs (2018) ran a stress judgment task that suggests the same conclusion as Shelton and colleagues. Spanish speakers extend weight sensitivity to nonce words, and a heavy penult makes antepenultimate stress worse. His participants judged the goodness of nonce words on a Likert scale. The nonce words had light penults ([da.ti.pe.βo]) and various types of heavy penults (e.g [da.ti.pem.bo]), and each word was presented with both penultimate and antepenultimate stress. Participants preferred penultimate stress over antepenultimate stress regardless of syllable profiles ([da.ti.ˈpe.βo] > [da.ˈti.pe.βo]), but the preference was stronger when the nonce word had a heavy penult ([da.ti.ˈpem.bo] ≫ *[da.ˈti.pem.bo]). This suggests that the restriction *ĹHσ present in the

lexicon—no antepenultimate stress when the penult is heavy—is active in Spanish speakers' synchronic grammars, and is applied productively to nonce words. These experimental results are strong evidence for weight sensitivity in the grammar of stress assignment.

My experiment, described in the next section, builds on Fuchs's (2018) findings and methodology to show that the weight effect extends to syllables whose codas have metathesized out.

4.2 Stress judgment experiment set-up

I use the *ĹHσ restriction to test the representation of underlying stop-h sequences in Sevillian Spanish. Using words with antepenultimate stress, the experiment tests whether penults with [h] metathesized out (/ginakaspo/ → [gi'.na.ka.pho], LLHL → LLLL) are treated as light or heavy. The aspirated stop analysis of [Ch] predicts that they should be treated as light, because [Ch] is a single segment at all levels of representation. The cluster analysis predicts that they should be treated as heavy, eliciting a strong dispreference due to the combination of antepenultimate stress and heavy penult. Although the penult is light on the surface, [Ch] derives from underlying /sC/, where /s/ closes the preceding syllable at early stages of the derivation. At the level of representation where stress is assigned, /s/ still closes that syllable. I test this by asking listeners to choosing between words with antepenultimate stress that are identical except for different types of light and heavy penults.

Now, I describe the stimuli used in the main stress judgment task and the comparisons listeners are asked to make. In the following section (Section 4.3), I briefly present acoustic and perceptual analyses of the stimuli, to confirm that stress is present in the intended locations. Having established that, I turn to the main experiment (Section 4.4).

4.2.1 Stimuli

Nonce words consist of four syllables and have antepenultimate stress. Several sets are illustrated in Table 4.4. The final onsets are /ptk/, which can host metathesized [h], and the vowels in both

the antepenult and the penult are /aiu/. These vowels were the same to ensure that listeners' stress judgments were based on syllable structure, not the ability of a vowel to host stress (/ginakapo/). Other consonants in the word were chosen to avoid repetition, [s], and [h] (since listeners may be prone to mishearing the linear position of sibilance and possibly [h] as well (Blevins 2004). The nonce words were designed in sets where members of the set differ only by the type of penult: CV (no coda), CVN (coda sonorant), CVS (coda [s]), CVH (coda [h]), CV.CH (metathesized [h]). There were 45 test nonce words (9 words x 5 penult types). The other CV base forms are listed at the bottom of Table 4.4.

Table 4.4: Penult types for stress acceptability experiment for a subset of stimuli words (those with /p/ final onset)

Final Onset	Penult V	CV	CVN	CVS	CVH	CV.CH
/p/	/a/	giˈnakapo	giˈnakampo	giˈnakaspo	giˈnakahpo	giˈnakapho
/p/	/i/	tʃuˈnifipo	tʃuˈnifimpo	tʃuˈnifispo	tʃuˈnifihpo	tʃuˈnifipho
/p/	/u/	naˈlufupo	naˈlufumpo	naˈlufuspo	naˈlufuhpo	naˈlufupho
Remaining CV base words: luˈmafato, kiˈpalako, laˈpikito, tʃaˈkiliko, taˈmukuto, tʃaˈlupuko						

36 filler words were constructed in sets, just like test words (9 sets x 4 per set). The filler conditions involved changes that should not affect stress judgments. Two fillers in each set differed minimally from the CV forms by (1) the onset of the final syllable from a stop to a nasal ([giˈnakapo] → [giˈnakamo]), and (2) the initial consonant by voicing, continuancy or nasality ([giˈnakapo] → [kiˈnakapo]). Another two fillers differed minimally from the CV.CH word by changing (1) the original voiceless stop final onset to another voiceless stop ([giˈnakapho] → [giˈnakatho], and (2) the first consonant ([giˈnakapho] → [kiˈnakapho]).

The nonce words were recorded by a male linguist who is a native speaker of Sevillian Spanish in his late 20s. The recording was done in a quiet room with a Zoom H4N Pro recorder. He was instructed to produce all words with antepenultimate stress, and different variants of coda /s/ ([s], [h], [Ch]), and did so easily, as I show in the stimuli verification study in Section 4.3.

4.2.1.1 Neighborhood density

Because antepenultimate stress is relatively marked, listeners might be more willing to accept it on nonce words that are more similar to existing words with antepenultimate stress. For this reason, I controlled for neighborhood density based on the CV forms using Levenshtein distance. The Levenshtein distance is the edit distance to turn one string into another, and calculations were based on orthography.[3] Using a subset of four-syllable words with antepenultimate stress from SUBTLEX-ESP (Cuetos et al. 2012)[4], I calculated the Levenshtein distance between each CV form and each word in the corpus subset (using the *stringdist* package for R, van der Loo 2014). Although the Levenshtein distance is a rough measure of neighborhood density, it has been found to explain much of listeners' judgments of word similarity (Vitevitch and Luce 2016: p. 78).

The CV nonce words used in the experiment (from which the rest of the word sets were built) have no lexical neighbors at a distance of three or less. The goal was to use nonce words that are similar—but not too similar—to real words. Dautriche (2017: fn. 13) suggests that word confusability effects tail off after an edit distance of one to two, so words at a distance of three should be equally susceptible to (lack of) effects of having lexical neighbors. The chosen CV nonce words also had no more than five neighbors at Levenshtein distance 4.

4.2.1.2 Experiment implementation

The preliminary and main experiments were implemented with PCIbex (Zehr and Schwarz 2018). The preliminary perception experiment was distributed through Prolific to participants from Spain. Participants for the main experiment were recruited through personal contacts, to restrict region of origin within Spain to Seville.

[3]There are potential downsides to using orthography for calculating neighborhood density, but the orthography-to-phoneme correlation in Spanish is almost one-to-one, unlike English, and was deemed sufficient for these purposes.

[4]SUBTLEX-ESP is a 41 million word Spanish corpus based on film subtitles.

4.3 Preliminary study: Analysis of stimuli

My study uses stimuli presented auditorily (unlike Fuchs 2018, who used orthographic stimuli). I thus conducted a stimulus verification study to ensure that the stimuli were produced with acoustic cues to stress in the intended antepenultimate location (Section 4.3.1), and are heard by listeners as having antepenultimate stress (Section 4.3.2).

4.3.1 Acoustic analysis of stimuli

Spanish primary stress is acoustically realized with a combination of f0, duration, and intensity (Llisteri et al. 2003; Ortega-Llebaria and Prieto 2010). Stressed vowels are longer and have slightly higher intensity (Ortega-Llebaria and Prieto 2010), although intensity alone is not a strong correlate. F0 is more complicated because it interacts with intonation, and many studies confound stress and pitch accents. In Castilian Spanish (North-Central Spain), stressed syllables have a rising pitch contour (Ortega-Llebaria and Prieto 2010; Vogel et al. 2016) that peaks on the following syllable (Ortega-Llebaria and Prieto 2010). I have not found studies of Sevillian intonation, but rising contours on stressed syllables are reported for nearby varieties (Henriksen and García-Amaya 2012).

My stimuli show acoustic evidence of stress in the intended antepenultimate location. The acoustic analysis compares antepenultimate vowels ([giˈna̲kapo]; intended to be stressed) to penultimate vowels ([giˈnaka̲po]; intended to be unstressed). Because these vowels are the same, vowel quality is not a confound. Antepenultimate vowels (stressed) are longer than penultimate vowels (unstressed) (Figure 4.1). Intensity is not a correlate of stress in these stimuli (plot not shown). F0 was measured at four points during the vowel (20ms, 30ms, 40ms, 50ms). For antepenultimate vowels ([gina̲kapo]), there was a fifth measurement point, which was the max f0 on the *following* vowel ([ginaka̲po]). This is because antepenultimate (stressed) vowels have a rising pitch contour that peaks on the following syllable (Figure 4.2). Unstressed penultimate vowels have a flat pitch contour (plot not shown).

Figure 4.1: Duration of vowels in stimuli nonce words by position

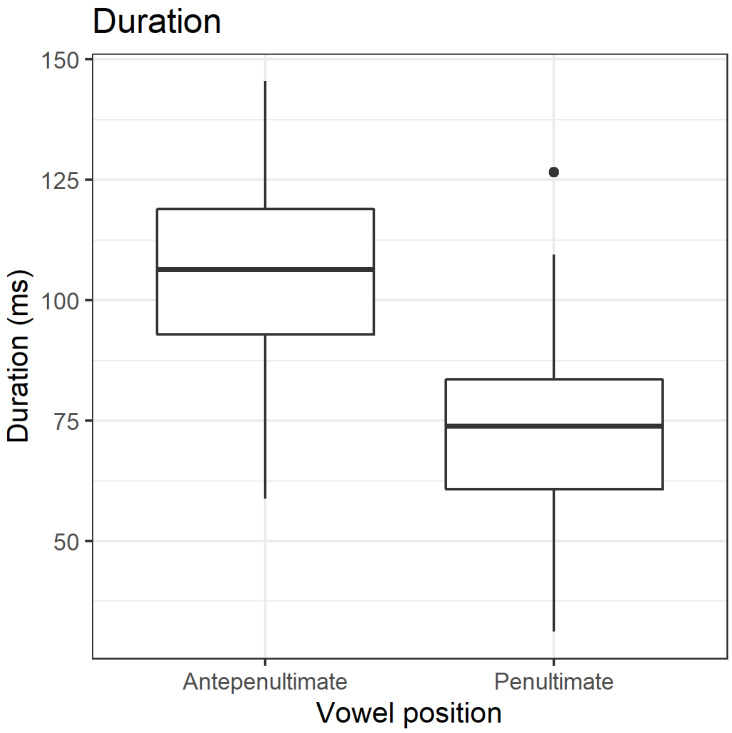
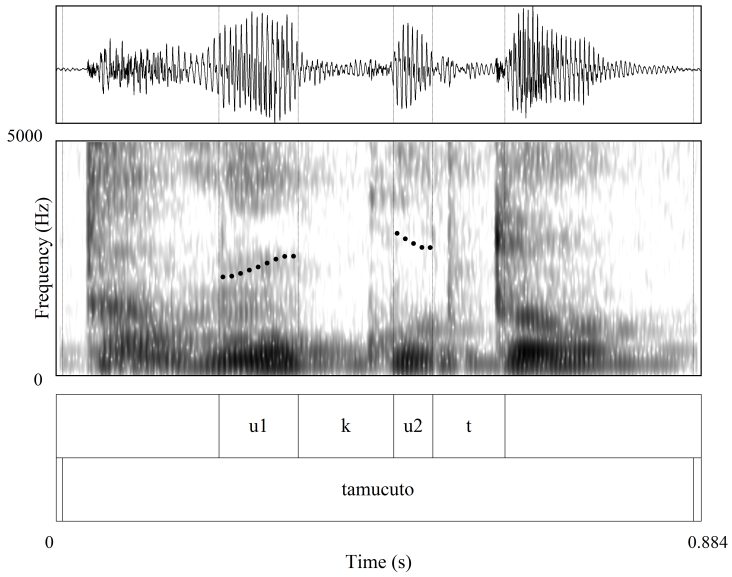


Figure 4.2: Sample f0 of antepenultimate vowels in stimuli nonce words



Linear regressions confirm these visual tendencies. The Duration and Intensity models contained an interaction between Condition (CV, CVN, CVS, CVH, CV.CH) and Vowel Position (antepenultimate, penultimate). Penultimate vowels (unstressed) are shorter than antepenultimate

syllables (stressed) (gi'nakapo) (β = -.031, p < .01), and there is no effect of Condition, either as a main effect or in the interaction. The intensity model has no significant effects, which is not unexpected since intensity is known to be a weak correlate of stress.

For f0, each vowel (/a, i, u/) was modeled separately in each position (antepenultimate, penultimate) to investigate contours in stressed and unstressed vowels. The models contained a main effect of Measurement Location (point in the vowel where measurement was taken). Results reported are from models and post-hoc tests with *emmeans* (with Tukey adjustments, Lenth 2020). For antepenultimate vowels of all qualities, the maximum pitch on the following vowel is higher than f0 at all locations within the antepenultimate vowel (p < .001 for all). For antepenultimate vowels /i u/, pitch at the end of this vowel trends higher than the pitch at the beginning of the vowel, indicating a rise over the course of the vowel. For penultimate vowels, measurement location was not a significant predictor of f0. Unstressed penultimate vowels show no pitch rise.

In sum, the antepenultimate vowels in my stimuli carry the acoustic correlates of stress. Now, I show that listeners hear stress as antepenultimate, too.

4.3.2 Preliminary perception study

4.3.2.1 Materials and task

The stimuli are four syllables long and have antepenultimate stress (see Table 4.4). Because all of the words were recorded with the same stress (antepenultimate), it was not possible to test perception of stress without modifying these words. To test perception of stress, I created two 3-syllable versions of each nonce word, one with antepenultimate stress and one with penultimate stress. To create the antepenultimate stress items, I removed the first syllable ([gi'nakapo] →

[ˈnakapo]). To create the penultimate stress items, I removed the final syllable (→ [giˈnaka]).[5] Table 4.5 illustrates the modified nonce words.

Table 4.5: Example stimuli for one word set for preliminary stress experiment

Original word	Condition	Penult	Answer Choices	AntePen	Answer Choices
[ginakapo]	CV	[giˈnaka]	guinaca, gúinaca	[ˈnakapo]	nácapo, nacapo
[ginakampo]	CVN	[giˈnakam]	guinacam, gúinacam	[ˈnakapo]	nácapo, nacapo
[ginakaspo]	CVS	[giˈnakas]	guinacas, gúinacas	[ˈnakaspo]	nácaspo, nacaspo
[ginakahpo]	CVH	[giˈnakah]	guinacas, gúinacas	[ˈnakahpo]	nácaspo, nacaspo
[ginakapho]	CV.CH	[giˈnaka]	guinacas, gúinacas	[ˈnakapho]	nácaspo, nacaspo

In each trial, listeners heard a single word and were asked to click on the orthographic representation of the word they heard. The answer choices differed only in the location of stress. Following orthographic conventions of Spanish, antepenultimate stress was marked with an acute accent, and penultimate stress was not marked. However, stressed syllables were underlined in both answer choices, to ensure that participants noticed the orthographic differences between options. CVH and CV.CH forms were represented orthographically with <s>, since orthographic <s> corresponds to the reduced [h] (e.g. [giˈnakah] as <guinacas>).

There were 162 words. Half were presented to listener Group A, and the other half were presented to Group B. Participants in a given group heard half the words in a set with penultimate stress (CV, CVS, StopH, Filler1, Filler3) and the half of the words in the set with antepenultimate stress (CVN, CVH, Filler2, Filler4). In the next word set, they heard each word type with the opposite stress pattern. This pattern repeated for all of the word sets. For example, looking at Table 4.5, Group A heard some of the test conditions of the /ginakapo/ set with penultimate stress (CV, CVS, CV.CH: [giˈnaka, giˈnakas, giˈnaka]), and the other test conditions with antepenultimate stress (CVN, CVH: [ˈnakapo, ˈnakah]) (plus fillers). Group B heard the opposite. Trials were

[5]This method produced some three-syllable words that are not phonotactically legal in Spanish—for example, the CVN condition of [giˈnakampo] is shortened to [giˈnakam], and [m] is not a legal coda. I do not believe this is an issue, since the answer choices are identical except for stress: *guinacam* has penultimate stress, and *guínacam* has antepenultimate stress. Following the orthographic convention, I only marked antepenultimate stress. Furthermore, in some conditions like Metathesis, the answer choices are presented orthographically with <s> even though there is no [s] in the acoustic signal. When cutting off the final syllable, evidence of /s/ disappears ([giˈnakapho] → [giˈnaka]). Again, this is not important, since the answer choices differed only in stress (*guinacas* vs. *guínacas*).

randomized. There were also two practice items and six attention checks to ensure participants were paying attention.

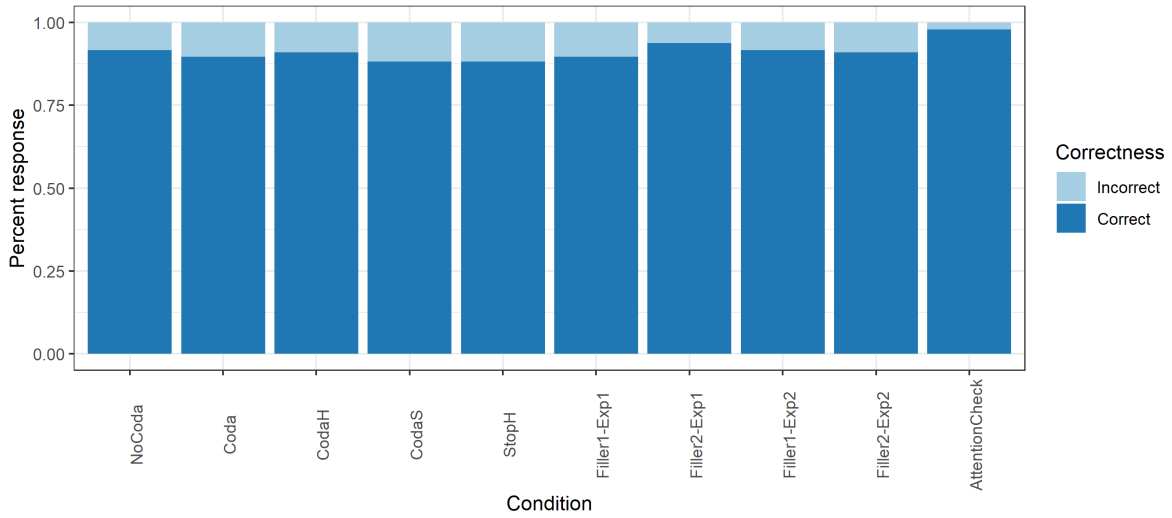
4.3.2.2 Participants

There were 20 participants from Spain, 10 in Group A, 10 in Group B. Region within Spain was not controlled, since the goal is to check if Spanish speakers broadly hear stress in the intended location. The experiment lasted 5-15 minutes and participants were paid. One listener from Group A and three from Group B were excluded for answering more than one attention check wrong, so the remainder of the analysis includes 16 listeners.

4.3.2.3 Results

Listeners heard stress where it was intended to be with 88%-98% accuracy across penult types (Figure 4.3). A logistic mixed-effect regression model (fixed effect of Condition, random intercepts for Participant and Item) finds that the only word type that differed in accuracy from baseline CV was the Attention Check (β = -1.58, $p < .05$). This is expected because the attention check items were unambiguous, and several listeners were excluded for failing them. The model and post-hoc tests with *emmeans* found no significant differences between other penult types. Listeners' accuracy was not significantly different across words with different types of penults.

Figure 4.3: Stimulus verification study: Accuracy in locating stress by penult type



To summarize, the acoustic cues to stress are present on the intended antepenultimate syllables in the stimuli, and listeners perceive them. This combined evidence ensures that listeners' judgments in the main experiment are based on a correct perception of stress.

4.4 Stress judgment task: Methods and results

I now turn to the main study, which tests listeners' perception of words with antepenultimate stress and different types of penults.

4.4.1 Methods

4.4.1.1 Task

The task was inspired by Fuchs (2018), but differs in several ways. First, his stimuli were presented orthographically. This was impossible in the current study, because some of the forms are allophonic realizations of the same word. Additionally, he asked participants to rate words on a goodness scale. The current study forces participants to make a binary choice between words. Some of the conditions in Fuchs' study had very similar goodness ratings, and forcing participants to make a binary choice makes their preferences clearer.

In each trial, participants heard two words paired as in Table 4.6 (e.g. [giˈnakapo]-[giˈnakahpo], CV-CVH). The CV comparisons compare all coda conditions to a CV base form. The CV.CH comparisons compare all coda conditions to a CV.CH base form. Two sets of comparisons were necessary because the CV comparisons were expected to show that words with light penults were preferred over those with heavy penults, but they do not test whether CV.CH penults differ from other types of heavy penults. The CV.CH comparisons test this difference directly.

Participants were told that the words were not real words of Spanish, and were asked to choose which word would be a better word of Spanish by clicking buttons labeled '1' or '2', corresponding to the first and second words they heard (e.g. [giˈnakapo] vs. [giˈnakampo].

Table 4.6: Condition pairings for CV and CV.CH comparisons (12 total comparisons)

	CV		CV.CH	
	Base	Comparison	Base	Comparison
Test	CV-	CVN	CV.CH-	CV
	CV-	CVS	CV.CH-	CVN
	CV-	CVH	CV.CH-	CVS
	CV-	CV.CH	CV.CH-	CVH
Filler	CV-	Filler1	CV.CH-	Filler1
	CV-	Filler2	CV.CH-	Filler2

There were a total of 108 trials (9 word sets (e.g Table 4.4) x 12 comparisons per word set (Table 4.6) = 108). The word pairs were only presented auditorily, and trial order was randomized for each participant. The audio presentation was counterbalanced, so Group A heard half of the item pairings with the base form first (e.g. CV-CVS) and the comparison form second, and the other half of the items with the sound files in the opposite order (e.g. CVH-CV). Group B heard the opposite orders. Participants were given one practice item before starting the actual experiment. They also filled out a demographic questionnaire. The experiment took most participants between 12-35 minutes.

4.4.1.2 Participants

The participants were 27 Sevillians (20 female, 7 male), with an average age of 37.9 years (range: 18-70). All participants had completed high school, thirteen had completed a technical or university degree, and twelve had done postgraduate studies (one did not report education). Participants reported knowledge of languages other than Spanish: English (20), French (7), German (2), Italian (2), Portuguese (2), Gallego (1) and 'Catalan/Valencian' (1). Three speakers reported almost native proficiency in another language (English, French, Gallego).

4.4.1.3 Statistical analysis

The results were analyzed in logistic mixed-effect regression models (*lme4*, Bates et al. 2015). Separate models were built for the CV comparisons (all forms compared to the CV form) and the CV.CH comparisons (all forms compared to the CV.CH form). How likely are participants to choose the base form (CV or CV.CH), given the alternative? The dependent variable was the response: CV vs. Alternative, and CV.CH vs. Alternative. The base word form was coded as 1 and the alternative was coded as 0, so positive effects indicate higher probability of the base form and negative effects indicate that listeners prefer the alternative more strongly. The models contained a fixed effect of Comparison Type, and random intercepts for Word Set and Participant, as illustrated in Table 4.7 (baseline levels are underlined). Post-hoc tests were used to test for differences between levels of Comparison type (*emmeans*, with Tukey adjustments for multiple comparisons; Lenth 2020).

Table 4.7: Predictors in models for stress experiment (baseline level underlined)

	Predictor	Levels (CV model)	Levels (CV.CH model)
Fixed effect	Comparison type	<u>CV-CVN</u>	<u>CV.CH-CVN</u>
		CV-CVS	CV.CH-CVS
		CV-CVH	CV.CH-CVH
		CV-CV.CH	CV.CH-CV
Random intercepts	Word set		
	Participant		

4.4.1.4 Hypotheses

The different possible representations of CV.CH make different predictions about how listeners will treat them. Table 4.8 and Table 4.9 lay out predictions for if stop-h is underlyingly an /sC/ cluster or an aspirated stop. The '>' symbol indicates that the first member of the pair should be preferred; '<' indicates that the second member of the pair should be preferred.

Table 4.8: Stress experiment predictions for CV comparisons

Condition	Both analyses		UR
CV-CVN	>	gi'nakapo > gi'nakampo	CV.'CV.<u>CV</u>.CV > CV.'CV.<u>CVN</u>.CV
CV-CVS	>	gi'nakapo > gi'nakaspo	CV.'CV.<u>CV</u>.CV > CV.'CV.<u>CVS</u>.CV
CV-CVH	>	gi'nakapo > gi'nakahpo	CV.'CV.<u>CV</u>.CV > CV.'CV.<u>CVS</u>.CV
	Cluster analysis (/sC/)		
CV-CV.CH	>	gi'nakapo > gi'nakapho	CV.'CV.<u>CV</u>.CV > CV.'CV.<u>CVS</u>.CV
	Aspirated stop analysis (/$p^h t^h k^h$/)		
CV-CV.CH	=	gi'nakapo = gi'nakapho	CV.'CV.<u>CV</u>.CV = CV.'CV.<u>CV.C^h</u>V

Table 4.9: Stress experiment predictions for CV.CH comparisons

Condition	Cluster analysis (/sC/)		UR
CV.CH-CV	<	gi'nakapho < gi'nakapo	CV.'CV.<u>CVS</u>.CV < CV.'CV.<u>CV</u>.CV
CV.CH-CVN	=	gi'nakapho = gi'nakampo	CV.'CV.<u>CVS</u>.CV = CV.'CV.<u>CVN</u>.CV
CV.CH-CVS	=	gi'nakapho = gi'nakaspo	CV.'CV.<u>CVS</u>.CV = CV.'CV.<u>CVS</u>.CV
CV.CH-CVH	=	gi'nakapho = gi'nakahpo	CV.'CV.<u>CVS</u>.CV = CV.'CV.<u>CVS</u>.CV
	Aspirated stop analysis (/$p^h t^h k^h$/)		
CV.CH-CV	=	gi'nakapho = gi'nakapo	CV.'CV.<u>CV.C^h</u>V = CV.'CV.<u>CV</u>.CV
CV.CH-CVN	>	gi'nakapho > gi'nakampo	CV.'CV.<u>CV.C^h</u>V > CV.'CV.<u>CVN</u>.CV
CV.CH-CVS	>	gi'nakapho > gi'nakaspo	CV.'CV.<u>CV.C^h</u>V > CV.'CV.<u>CVS</u>.CV
CV.CH-CVH	>	gi'nakapho > gi'nakahpo	CV.'CV.<u>CV.C^h</u>V > CV.'CV.<u>CVS</u>.CV

CV comparisons (Table 4.8): At the level of individual trials, both analyses of stop-h (Cluster and Aspirated stop), predict that Sevillian listeners should prefer the CV form over CVN, CVS, and CVH forms (these comparisons do not involve CV.CH forms). CVN, CVS, and CVH forms

have heavy penults. The CV-CV.CH comparison distinguishes the /sC/ cluster vs. aspirated stop underlying representation. If listeners treat CV.CH as a cluster, they should prefer CV over CV.CH forms (like the other CV-CVN, CVS, CVH comparisons), because CV.CH is represented with /s/ in the coda of the penult. If they treat CV.CH as an aspirated stop, they should treat CV.CH the same as CV forms since both have a light penult.

Another prediction is about the *rates* of preference of CV over forms with codas. Listeners should prefer CV over words with codas at approximately the same rate; there is no reason to expect that different coda types should be more or less preferable in comparison to CV. With CV.CH, the cluster analysis predicts that listeners should prefer CV over CV.CH at approximately the same rate as words with surface-present codas. The aspirated stop analysis, however, predicts that the rate of CV preference over CV.CH should be much lower than comparisons between CV and CVN, CVS, CVH.

CV.CH comparisons (Table 4.9): If listeners treat CV.CH forms as underlying /sC/ clusters, they should prefer CV forms over CV.CH forms. In terms of rate of preference, they should also not have strong preferences in CV.CH comparisons with CVN, CVS, and CVH, since all of these forms have heavy penults underlyingly. In contrast, if listeners treat CV.CH forms as underlying aspirated stops, they should have no strong preference between CV and CV.CH forms, since both have light penults. They should also prefer CV.CH forms over any form with a coda in the penult (CVN, CVS, CVH).

4.4.2 Results

Figures 4.4 and 4.5 show results for the CV and CV.CH comparisons, respectively. The y-axis is the percentage of choosing the base form (CV or CV.CH) over the alternative. The line at .5 represents 50% chance of the base form response. Each dot represents an individual speaker's response rate for the given comparison type. Dots above the line indicate that listeners preferred the base form over the alternative; dots below the line mean listeners preferred the alternative over the base form. Dots on line mean listeners did not have a preference for one form over the other.

The blue dot represents the mean percentage response of the base form, and the error bars represent a 95% confidence interval.

Figure 4.4: Stress judgment task: CV comparison results

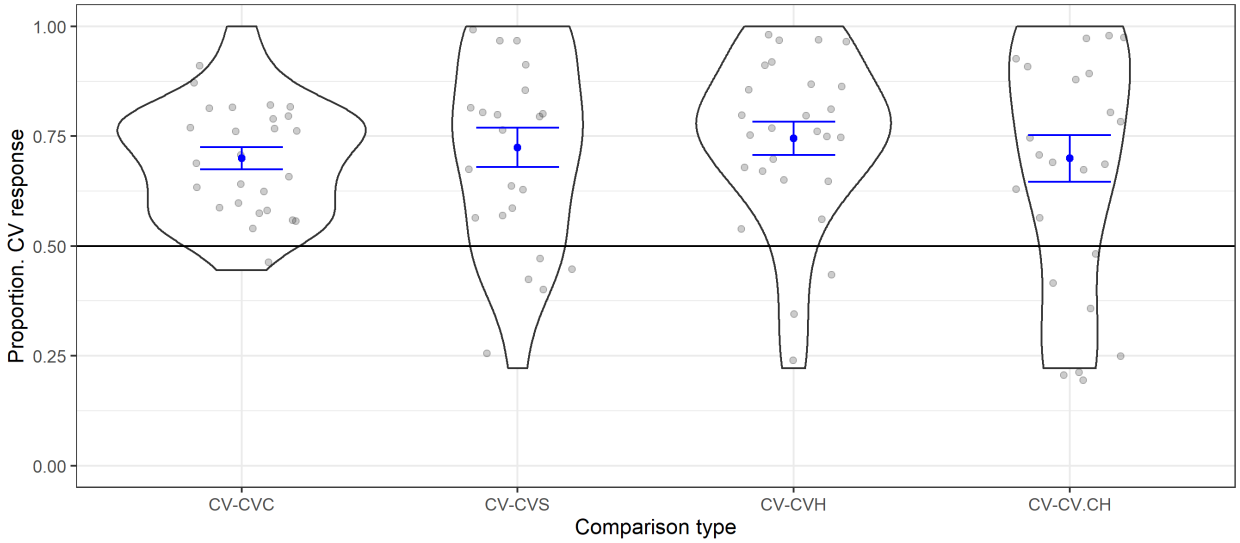
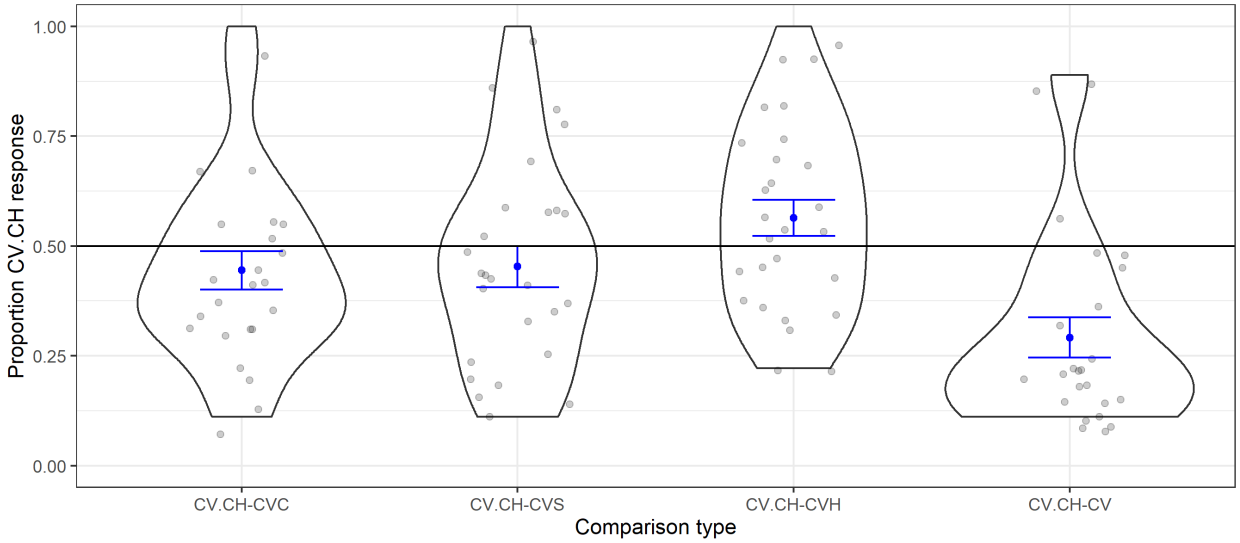


Figure 4.5: Stress judgment task: CV.CH comparison results



4.4.2.1 CV comparisons

In the CV comparisons, listeners compare CV forms to CVN, CVS, CVH, and CV.CH forms (Figure 4.4). In all comparison types, listeners chose the CV form at a rate of between 72%-77%. In words with antepenultimate stress:

- Listeners prefer light penults over heavy penults (CV > CVN, CVS, CVH, CV.CH; [luˈmafato] > [luˈmafanto], [luˈmafasto], [luˈmafahto], [luˈmafatho]).
- This preference for light > heavy penults applies at approximately the same rate regardless of which heavy penult it is (including CV.CH penults).

The logistic regression model confirms these results (Table 4.10). The model predicts the probability of CV response in the baseline comparison CV-CVN. Positive effects indicate a stronger probability of a CV response in the specified comparison than in the baseline CV-CVN comparison. Negative effects indicate a stronger probability for the alternative form in the specified comparison than in the baseline CV-CVN comparison. The effect of Comparison is not significant. The rates at which listeners prefer CV over the alternative are not statistically different from baseline CV-CVN for CV-CVS (β = -.15, p = .47), CV-CVH (β = .25, p = .24), CV-CV.CH (β = -.00, p = 1.00). Post-hoc pairwise comparisons indicate no significant differences between the other levels of Comparison Type.

Table 4.10: Model for CV comparisons predicting probability of CV response (1) vs. Alternative (0)

	Estimate	Std. Error	z value	Pr(> |z|)
Baseline: CV-CVN				
(Intercept)	0.974	0.219	4.450	0.000***
CV-CVS	-0.149	0.206	-0.725	0.468
CV-CVH	0.250	0.213	1.177	0.239
CV-CV.CH	-0.000	0.208	-0.000	1.000

Listeners treat CV-CV.CH comparisons similarly to other comparisons where the alternatives have heavy penults. However, the fact that the rate of CV preference is not statistically

different between comparison types should not be interpreted as confirmation that listeners prefer CV over CV.CH and other heavy penults at the same rate. A lack of statistical difference is not confirmation of similarity. It suggests that there is no robust evidence for a difference between conditions, but does not provide positive evidence for their similarity.

In order to test rate of preference for CV.CH against other penult types directly, the second set of comparisons pits CV.CH directly against other words with heavy penults. Are all coda consonants treated as contributing weight equally to the penult? The statistically significant results in the CV.CH comparisons are compatible with the lack of effect in the CV comparisons.

4.4.2.2 CV.CH comparisons

In the CV.CH comparisons, listeners compare CV.CH forms to CVN, CVS, CVH, and CV forms (Figure 4.5). In CV.CH-CVN and CV.CH-CVS comparisons, Sevillian listeners had no clear preference for one form over another, choosing each around 50% of the time ([luˈmafatho] ≈ [luˈmafanto]; [luˈmafatho] ≈ [luˈmafasto]). In CV.CH-CV comparisons, listeners disfavored CV.CH in comparison to CV ([luˈmafatho] < [luˈmafato]). In CV.CH-CVH comparisons, listeners have a slight preference for CV.CH over CVH ([luˈmafatho] > [lˈumafahto]). This preference may be due to sociophonetic rather than phonotactic factors: coda [h] as a realization of coda /s/ is not common in Sevillian Spanish (Ruch 2008). I discuss this effect further in Section 4.4.3.

The logistic regression (Table 4.11) confirms these results. CV.CH-CVN comparisons are the baseline for Comparison Type. Positive estimates indicate a stronger probability of a CV.CH response in the specified condition than in the baseline CV.CH-CVN comparison. Negative estimates indicate stronger probability of the alternative form in the specified comparison than in the baseline CV.CH-CVN comparison. The model shows that the rate of CV.CH responses is not significantly different in CV.CH-CVS comparisons than in the baseline CV.CH-CVN comparisons (β = .19, p = .33). In CV.CH-CV comparisons, listeners' preference for the alternative (CV) is stronger than in CV.CH-CVN (β = -.71, p < .01). In CV.CH-CVH comparisons, listeners' preference for CV.CH is stronger than in CV.CH-CVN comparisons (β = .77, p < .01).

Table 4.11: Model for CV.CH comparisons predicting probability of CV.CH response (1) vs. Alternative (0)

	Estimate	Std. Error	z value	Pr(> |z|)
Baseline: CV.CH-CVN				
(Intercept)	-0.462	0.219	-2.110	0.035*
CV.CH-CVS	0.193	0.196	0.984	0.325
CV.CH-CVH	0.774	0.197	3.940	0.000***
CV.CH-CV	-0.711	0.210	-3.386	0.001***

In words with antepenultimate stress, Sevillian listeners treat CV.CH forms the same as CVN and CVS forms, choosing between them at chance. Furthermore, the rate of choosing CV.CH is not statistically different between CV.CH-CVN and CV.CH-CVS comparisons. CVN and CVS forms are equally as bad as CV.CH forms, and equally bad as each other in comparison to CV.CH forms. CV forms, however, are better than CV.CH forms.

4.4.3 Discussion

The results are consistent with the predictions made for Sevillian listeners under the cluster analysis (/sC/) of stop-h sequences, for both the CV comparisons and the CV.CH comparisons (Table 4.12). Crucial predictions that distinguish the representational options are shaded gray.

Table 4.12: Comparison of predictions and results for CV (left) and CV.CH (right) comparisons

Condition	Predictions /sC/	Predictions /C^h/	Result
CV-CVN	>	>	>
CV-CVS	>	>	>
CV-CVH	>	>	>
CV-CV.CH	>	=	>

Condition	Predictions /sC/	Predictions /C^h/	Result
CV.CH-CV	<	=	<
CV.CH-CVN	=	>	=
CV.CH-CVS	=	>	=
CV.CH-CVH	=	>	=/>

In both sets of comparisons, CV.CH forms are treated like forms with surface-heavy penults. When given the choice, listeners prefer antepenultimate stress words with an open CV penult over antepenultimate stress words with a closed penult. Regardless of what consonant or acoustic man-

ifestation of /s/ closes the penult—even CV.CH forms, which have no acoustically present coda—Sevillian listeners evaluate stress as if these words have a heavy penult. Furthermore, CV.CH words do not pattern with CV words in either set of comparisons. This constellation of results suggests that [Ch] is underlyingly an /sC/ cluster, and that CV.CH forms have an underlying /s/ in the preceding syllable.

These results are, in principle, also consistent with syllabifying [Ch] sequences as heterosyllabic, with [C] closing the preceding syllable (*[gi'.na.kap.ho]). If this were the syllabification, the syllable structure would be the same pre-metathesis and post-metathesis, although the segments would have swapped ([gi'.na.kas.po] vs. *[gi'.na.kap.ho]). Earlier, I outlined reasons to assume that [Ch] should be syllabified as an onset, not as a heterosyllabic cluster (Section 4.1.1). In the analysis next, I also suggest there are analysis-internal reasons not to assume the heterosyllabic parse of [Ch] (Section 4.5.4).

One result does not fit clearly with the predictions: listeners prefer CV.CH forms over CVH forms. Under a cluster analysis of CV.CH forms, CV.CH and CVH forms share an underlying representation of /sC/ and should be treated equally. I suspect the preference for CV.CH over CVH words arises because this comparison differs from the others. When listeners judge e.g. CV vs. CVN forms, they are judging different words. When they judge CV.CH vs. CVH forms, they are judging different phonetic realizations of the *same* word. Between two non-standard pronunciations, they choose the one that is more frequent in their dialect (CV.CH). Furthermore, in conversational, spontaneous speech, Ruch (2008) found that forms with [ht] sequences are very infrequent (3.5% of all productions). Sevillians have little experience with these productions.

Furthermore, forms with metathesis ['patha], ['patsa]) are the most frequent (over 70% of productions), while forms with coda [h] and [s] ['pahta], ['pasta] are very infrequent (<4% each) (Ruch 2008). If this is true, we might expect listeners to prefer CV.CH over CVS, too, which is also a comparison between phonetic realizations of the same word. I speculate that listeners do not prefer CV.CH over CVS because the full sibilant is the national and international standard. They have no preference between the standard form and the form most common in their dialect (CVS ≈

CV.CH), but prefer the reduced form of their own dialect over reduced forms that are rare in their dialect (CV.CH > CVH). In other words, I believe this preference is due to their familiarity with surface phonetic forms, rather than to differing phonological structure of the forms.

4.5 Analysis

This analysis aims to capture how Sevillian listeners treated the interaction between stress and metathesis in the experiment. In the experiment, listeners had the following preferences. Among words with antepenultimate stress, they preferred those with surface-light penults over those with heavy penults, including penults that are underlyingly heavy but surface-light due to metathesis (50a). In comparing different types of codas in the penult, listeners dispreferred all equally (50b).

(50) Listeners' preferences in the stress judgment task

a. CV́.<u>CV</u>.CV > CV́.<u>CVC</u>.CV, CV́.<u>CV.CH</u>V

b. CV́.<u>CVC</u>.CV ∼ CV́.<u>CV.CH</u>V

Dispreference for one form in comparison to another does not necessarily indicate ungrammaticality. For the sake of simplicity, however, I treat the dispreferred forms as essentially ungrammatical, in the sense that the grammar does not produce them. Words with antepenultimate stress and heavy codas—which are always dispreferred in the experiment—are almost absent from the native Spanish lexicon. The analysis could be modified to account for dispreference (as opposed to complete ungrammaticality) by implementing some mechanism to introduce variability in the system (e.g. multiple grammars, weighted constraints, etc.).

The analysis focuses on two patterns that reflect listeners' preference in the stress experiments.

(51) Stress patterns to analyze:

a. Antepenultimate stress can surface when the penult is light.

b. Antepenultimate stress does not surface when the penult is heavy ([CVC] or [CV.CH]); instead, penultimate stress surfaces.

Together, the two analyses produce antepenultimate stress when the penult is light, but not when the penult is heavy. Instead, words with heavy penults receive penultimate stress.[6] Even through antepenultimate stress is lexically marked, a heavy penult will pull stress onto the penult. The analysis treats all types of heavy penults the same, reflecting the result that participants dispreferred them at roughly equal rates. This means that, even for CV.CH penults, which are light on the surface, stress will be pulled onto the penult. At the point when stress is assigned, that penult is still heavy.

In terms of mappings, my participants' preferences suggest that the grammar maps /CV́.CV.CV/ to [CV́.CV.CV], and vice versa in perception. The penult is light, so antepenultimate stress surfaces faithfully. In contrast, there is no derivational path in the grammar from /CV́.CVC.CV/ to [CV́.CVC.CV] or [CV́.CV.CHV]. The restrictions on stress and weight override lexical antepenultimate stress, which surfaces unfaithfully as penultimate.

The analysis builds on the analysis presented in Chapter 3. In that analysis, there were four steps: prosodification, debuccalization, gemination/metathesis, and degemination. In this chapter, I focus only on a step I did not analyze there: the stress step, which comes immediately after prosodification.

Following McCarthy and Pruitt (2013), I treat lexical stress as a diacritic present in the underlying representation. I use the diacritic * (as in Goldsmith 1976: 47) immediately preceding the marked vowel, instead of '$V^{h}_{\underline{\ }}$' used by McCarthy and Pruitt (2013). This is to avoid confusion with aspiration. This stress diacritic does not correspond to phonetic realization until it is mapped to a surface stress by means of violable constraints. Surface stress does not have to match underlying stress, but a mismatch incurs a violation. I simplify details by using a basic FAITHSTRESS constraint enforcing that stress be realized in the same location as the underlying diacritic (defined in (61)). I mark stress that has mapped to surface forms on the syllable to which it belongs, as traditionally done in the IPA (ˈ).

[6] My experiment did not include words with penultimate stress, but a strong dispreference for antepenultimate stress means that the preferred stress would be penultimate.

In order to get the right result, stress assignment must occur before /s/ metathesizes out of the coda (Table 4.13).[7] With this order of operations, CV and CV.CH penults are treated differently with regard to stress, reflecting my participants' differential treatment of them in the experiment. Lexical stress is mapped faithfully when the penult is light, and unfaithfully when the penult is heavy at the point where stress is mapped from underlying representation to a phonetic realization. In short, stress assignment feeds metathesis.[8]

Table 4.13: Stress: derivational steps with metathesized and CV penults in a word with assumed lexical antepenultimate stress

	Step	CVC penult	CV penult
		/gin*akaspo/	/gin*akapo/
Step 1	Prosodification	[gi.n*a.kas$_{\mu}$.po]	[gi.n*a.ka.po]
Step 2	**Stress**	**[gi.na$^{\prime}$.kas$_{\mu}$.po]**	**[gi$^{\prime}$.na.ka.po]**
Step 3	Debuccalization	[gi.na$^{\prime}$.kah$_{\mu}$.po]	—
Step 4	Gemination & Metathesis	[gi.na$^{\prime}$.ka.p:h$_{\mu}$o]	—
Step 5	Degemination	[gi.na$^{\prime}$.ka.pho]	—

If metathesis occurred before stress, the derivation would give the wrong result: CV.CH words behave exactly like words with CV penults, both allowing antepenultimate stress to surface faithfully (Table 4.14). If metathesis occurs before stress, there is no way to block antepenultimate stress in CV.CH: the penult is already light, so weight-to-stress allows antepenultimate stress to surface faithfully. Throughout the derivation, antepenultimate stress is present in the form of the diacritic, but is not mapped to a phonetic form until the stress step.

[7] I force stress to occur before all coda /s/ reduction (debuccalization and metathesis), but in theory it could be assigned on a form with a debuccalized coda as well.

[8] Recall that in Chapter 3, I frame the analysis in terms of features, but specify that I assume metathesis to manipulate gestures. What parts of phonology operate on what types of representations is a big question, and one I do not aim to answer. I only argue that metathesis operates on gestures, without making claims about the types of representations that stress operates on.

Table 4.14: Stress: Metathesis before stress gives the wrong result

	Step	CVC penult	CV penult
		/gin*akaspo/	/gin*akapo/
Step 1	Prosodification	[gi.*na.kas$_{\mu}$.po]	[gi.n*a.ka$_{\mu}$.po]
Step 2	Debuccalization	[gi.n*a.kah$_{\mu}$.po]	—
Step 3	Gemination & Metathesis	[gi.n*a.ka.p:ho]	—
Step 4	Degemination	[gi.n*a.ka.pho]	—
Step 5	**Stress**	**[giˈ.na.ka.pho]**	**[giˈ.na.ka.po]**

In the remainder of this section, I briefly review my assumptions about Harmonic Serialism from the analysis in Chapter 3 (Section 3.3.1), and extend the discussion to include how this framework has been used to account for interactions between stress and segmental processes (4.5.1). I also outline my assumptions about Spanish stress, and how I implement stress-weight restrictions in the analysis (4.5.2). Then, I analyze the two stress patterns that reflect my participants' behavior in the experimental task (Sections 4.5.3 and 4.5.4). I close the section with a summary of the analysis (Section 4.5.5).

4.5.1 Assumptions about Harmonic Serialism

Recall from Chapter 3 that one crucial property of Harmonic Serialism is that derivations proceed gradually, with candidates at each step differing from those at the previous step by a single change, which I define as a faithfulness violation. Gradualism is important in order to derive coda /s/ reduction and metathesis, because it allows for intermediate steps of reduction. Recall also that (re)syllabification can occur at any step, since it does not incur faithfulness violations. I limit my focus to viable syllabifications, without considering competing options.

Harmonic Serialism has also been used to analyze stress-segmental interactions. McCarthy (2008b) uses a version of HS to analyze metrically-conditioned syncope, which is similar to Sevillian metathesis in that stress feeds syncope (or, in my case, metathesis). He highlights two important results about the interaction between stress and syncope. First, stress assignment and syncope cannot occur at the same time, because they violate different faithfulness constraints (*forced seri-*

alism). Second, stress assignment must precede syncope (*intrinsic ordering*). Both of these results are also necessary for Sevillian stress and metathesis: the operations cannot happen simultaneously, and stress must precede metathesis.

McCarthy (2007) points out that vanilla Harmonic Serialism fares just as poorly as parallel OT in analyzing opaque interactions like counter-feeding and counter-bleeding, because the UR is not accessible to EVAL. His solution is to use candidate chains, in which all derivational steps are present for each candidate. This way, EVAL has access to the UR, all steps of the derivation, and the final output. Violations are assigned based on the UR, not the output of the most recent step. However, McCarthy (2008b: 503) notes that this manner of assigning violations is not necessary for his analysis of stress-syncope interactions, which require only a specific ordering. The Sevillian Spanish stress-metathesis interaction is the same, so I simplify my exposition by using plain Harmonic Serialism, although something like candidate chains is needed to analyze more complex opacity.

In my analysis, I show stress assignment in detail and use cover constraints for reduction and metathesis, which are analyzed in detail in Chapter 3 (Section 3.3.6). Ranking stress constraints above constraints compelling segmental reductions enforces stress assignment before reduction and metathesis.[9]

4.5.2 Spanish stress and implementing syllable weight restrictions

I make several assumptions to simplify the analysis. I assume Spanish stress falls within a right-aligned, three-syllable window and do not consider candidates with stress outside it (see Kager 2012 on deriving stress windows). I also assume that all stress is lexically-marked within the window (Harris 1983; Baković 2016). For preferences within the stress window, I use constraints preferring stem-final stress and non-final stress, following Baković (2016).

[9]For examples of accounts that derive ordering between stress and segmental processes, see Elfner (2009) on stress-epenthesis interactions and McCarthy (2008b) on stress-syncope interactions.

I also treat Spanish stress as weight sensitive, following previous theoretical (e.g. Harris 1983; den Os and Kager 1986) and experimental research (e.g. Shelton 2007 et seq.; Fuchs 2018). Recall that weight-sensitivity is also evident in the lexicon (Section 4.1.2, Table 4.3, Bárkányi 2002). Heavy penultimate and final syllables have a strong effect on stress location. To implement weight sensitivity within the stress window, I use the Weight-to-Stress Principle (52):

(52) WEIGHT-TO-STRESS: If heavy, then stressed. Assign a violation for a heavy unstressed syllable (Prince 1990).

A single WSP constraint does not suffice, because, as discussed in Section 4.1.2, not all syllables within the stress window have the same influence on stress. Final and penultimate syllables have a stronger effect, and antepenultimate syllables have a weak effect, or no effect at all. Following Garcia (2019) for Portuguese (in which the stress patterns are very similar), I use position-specific WSP constraints to penalize unstressed heavy syllables in different positions ((53)-(55)):

(53) WSP-FINAL: Assign a violation for an unstressed heavy syllable in final position.

(54) WSP-PENULTIMATE: Assign a violation for an unstressed heavy penult.

(55) WSP-ANTEPENULTIMATE: Assign a violation for an unstressed heavy antepenult.

In Portuguese—and apparently in Spanish too, given the lexicon patterns—these constraints are ranked as follows: WSP-FINAL $\gg$ WSP-PENULTIMATE $\gg$ WSP-ANTEPENULTIMATE. A heavy final syllable has the strongest requirement for WSP to hold, followed by a heavy penult. Heavy antepenults affect stress least.

My analysis only makes use of WSP-PEN, since my experimental results speak only to the effect of heavy penults. That said, splitting WSP into multiple parts would be crucial to account for other patterns. Interleaving WSP constraints with faithfulness constraints allows faithfulness to lexical stress to be overcome by heavy syllables in certain positions (penultimate and final) but not in others (antepenultimate). This allows for the differential strength of heavy syllables in antepenultimate, penultimate, and final positions.

4.5.3 Pattern 1: Antepenultimate stress allowed to surface with light penult

My participants preferred antepenultimate stressed words with CV penults as compared to CVC penults (56). This part of the analysis captures this preference by mapping stress faithfully in words with underlying antepenultimate stress and light penults (57).

(56) Preference in perception:
CV.CV́.<u>CV</u>.CV > CV.CV́.<u>CVC</u>.CV

(57) Mapping in production:
/C*V.CV.CV/ → [CV́.CV.CV] Faithful
*/C*V.CV.CV/ → [CV.CV́.CV] No unfaithful mapping

Recall from the analysis in Chapter 3 (Section 3.3.4) that prosodification is the first step, and includes mora assignment and syllabification. Prosodification works the same here. Syllabification occurs, giving the form in (58). As already discussed in Section 4.1.1, I assume that other syllabifications (e.g. [gi.na.kap.o], [gi.na.kap.ho]) are ruled out by by high-ranked markedness constraints that disallow these types of syllable structures.

(58) Syllabification of /gin*akapo/: [gi.n*a.ka.po]

No moras are assigned because the word has no coda consonants, and the constraints requiring mora assignment apply only to codas.

After prosodicification, stress assignment occurs. I use constraints militating for stem-final stress and non-final stress (Prince and Smolensky 1993), implemented following Baković's (2016) use for Spanish. The constraints are calculated based on stems and theme vowels.

(59) NONFINALITY: Assign a violation for stem-final stress; assign two violations for word-final stress.

(60) FINALSTRESS: Assign a violation for each syllable separating the stressed syllable in the stem from the end of the stem (not violated by stress falling outside the stem, e.g. word-final stress on a theme vowel).

Ranking FINALSTRESS ≫ NONFIN results in penultimate stress, the most common pattern: stress falls on the final vowel of the stem, even though this violates NONFINALITY.

Two further constraints are needed. FAITHSTRESS enforces mapping underlying stress to surface stress in the same location (61), and HAVESTRESS (62) ensures that the underlying stress diacritic is mapped to a surface form.

(61) FAITHSTRESS: Assign a violation if a vowel that is marked with a stress diacritic in the input is unstressed in the output.

(62) HAVESTRESS: Assign a violation for a word for which underlying stress has not been mapped to phonetic stress.

Only NONFIN, FINALSTRESS, and faithfulness constraints are relevant to assign stress to [gi.n*a.ka.po], since the word has no heavy syllables. However, I include WSP-PEN for the sake of consistency with the second part of the analysis.

The derivation for /gin*akapo/ is shown in the tableau in (63). In Step 1, the UR is prosodified. In Step 2, stress is assigned. The main comparison is between a form with faithful antepenultimate stress (a), and one with penultimate stress (b). The bar (|) separates the stem from the theme vowel [o]; recall that the constraints that determine the location of stress (NONFIN, FINALSTRESS) are calculated based on position in the *stem*. Candidate (a) wins over (b) because the ranking FAITHSTRESS ≫ FINALSTRESS favors faithful mapping of underlying to surface stress, as opposed to moving stress closer to the end of the stem. Candidate (c) is eliminated because the vowel marked for stress in the input does not have stress realized in the output. WSP-PEN is inactive, because the penult is light.

(63) Antepenultimate stress maps faithfully when the penult is light

a. Step 1 | **Prosodification**:
/gin*akapo/ → [gi.n*a.ka.po]

b. Step 2 | **Stress assignment**: HAVESTRESS, FAITHSTRESS ≫ FINALSTRESS
[gi.n*a.ka.po] → [giˈ.na.ka.po]

# 2	gi.n*a.ka.p|o	HAVESTR	WSP-PEN	FAITHSTR	FINSTR	NONFIN
☞a.	giˈ.na.ka.p|o				*	
b.	gi.naˈ.ka.p|o			*!		*
c.	gi.na.ka.p|o	*!		*		

Output: [giˈ.na.ka.po] **Derivation:** /gin*akapo/ → [gi.n*a.ka.po] → [giˈ.na.ka.po]

The derivation converges here. There is no coda to reduce. This is a straightforward case where antepenultimate stress surfaces in a word with a light penult, with no conflict between faithfulness to lexically-specified stress and WSP preferences.

4.5.4 Pattern 2 analysis: Antepenultimate stress does not surface with heavy penult

My participants dispreferred antepenultimate stress CVC and CV.CH words as compared to antepenultimate stress words with light CV penults (64). This second part of the analysis captures this dispreference by mapping lexically-marked antepenultimate stress unfaithfully to penultimate stress in words with heavy penults (65). I show how the derivation works similarly for both CVC and CV.CH penults, illustrating with [ginákampo] and [ginákapho]. The latter is the form of most interest, due to the mismatch in syllable structure between early and late derivational forms.

(64) Preference in perception (same as for Pattern 1):
CV.CV́.<u>CV</u>.CV > CV.CV́.<u>CVC</u>.CV

(65) Mapping in production:
/C*V.CVC.CV/ → [CV.CV́C.CV], [CV.CV́.CHV] Unfaithful
*/C*V.CVC.CV/ → [CV́.CVC.CV], [CV́.CV.CHV] No faithful mapping

Because all the words in my experiment had antepenultimate stress, I assume that listeners treated this stress as lexically marked. Additionally, although antepenultimate stress does not surface on existing lexical words with CVC penults, richness of the base means that the grammar must be able to take any input—including antepenultimate stress on a word shape LLHL—and map it to a viable surface form.

Starting with the CVN word /gin*akampo/, prosodification syllabifies the form and assigns a mora to coda /m/ (tableau in (66)), as was previously established in Chapter 3. At this point, stress still has not been mapped to a surface realization, so it is still marked with *.

(66) First derivational step: Assigning prosodic structure (mora and syllabification) to /gin*akampo/

	/gin*akampo/	*APPEND	*μ/OBS
a.	gi.n*a.kam.po	*!	
☞b.	gi.n*a.kam$_{\mu}$.po		*

The remainder of the derivation is in the tableau in (67). After prosodification, stress assignment occurs. Candidate (a) has antepenultimate stress, realizing stress in the same location as it is lexically marked. However, (a) also violates WSP-PEN, because the penult is heavy and unstressed, and this violation is fatal. Candidate (b) wins. This candidate has penultimate stress, violating faithfulness to lexical stress, but satisfies stress-weight requirements because the heavy penult is stressed. Candidate (c) is ruled out because underlying stress is not mapped to the surface form.

(67) Antepenultimate stress maps to penultimate stress in words with heavy CVN penults

a. Step 1 | **Prosodification**: (see (66))
/gin*akampo/ → [gi.n*a.kam$_{\mu}$.po]

b. Step 2 | **Stress assignment**: HAVESTRESS, WSP-PEN ≫ FAITHSTRESS
[gi.n*a.kam$_{\mu}$.po] → [gi.na'.kam$_{\mu}$.po]

# 2	gi.n*a.kam$_{\mu}$.p|o	HAVESTR	WSP-PEN	FAITHSTR	FINSTR	NONFIN
a.	gi'.na.kam$_{\mu}$.p|o		*!		*	
☞b.	gi.na'.kam$_{\mu}$.p|o			*		*
c.	gi.n*a.kam$_{\mu}$.p|o	*!		*		

Output: [gi.na'.kam$_{\mu}$.po] **Derivation:** /gin*akampo/ → [gi.n*a.kam$_{\mu}$.po] → [gi.na'.kam$_{\mu}$.po]

Coda /m/ does not reduce, so the derivation is finished. The heavy penult prevents underlying antepenultimate stress from surfacing faithfully, instead pulling it onto the heavy penult.

The derivation for /gin*akaspo/ → [gina'kapho] proceeds similarly: lexical antepenultimate stress maps to penultimate stress because the penult is heavy when stress assignment occurs. This derivation differs from the previous one (67) in that coda /s/ undergoes reduction after stress assignment, resulting in a light penult on the surface, and obliterating the context that blocked antepenultimate stress in the first place.

Prosodification works the same as for /gin*akampo/:/gin*akaspo/ is syllabified and a mora is assigned to coda /s/ (tableau in (68).

(68) First derivational step: Assigning prosodic structure (mora and syllabification) to /gin*akaspo/

	/gin*akaspo/	*APPEND	*μ/OBS
a.	gi.n*a.kas.po	*!	
☞b.	gi.n*a.kas$_{\mu}$.po		*

I use cover constraints for coda /s/ reduction and metathesis (69, 70). REDUCE is satisfied by debuccalization or deletion of coda /s/, and violated by retention of coda /s/ as sibilant [s]. METATHESIZE is violated by [hC], and satisfied by [Ch]; the mechanics of this constraint are not

important here. See Chapter 3 (Section 3.3.6) for the details of segmental processes. I focus on the crucial part to explain the results of the stress experiment: the sequencing of the stress step in relation to others.

(69) REDUCE: Assign a violation for an unreduced coda /s/.

(70) METATHESIZE: Assign a violation for a HC sequence.

The tableau in (71) illustrates stress assignment, coda reduction, and metathesis. I omit gemination and degemination for simplicity. Recall that I assume that only [h] is originally syllabified as a coda and the following (geminated) consonant is in onset position. Under this assumption, omitting gemination from this tableau does not affect the derivation, because the geminate does not affect the weight of the preceding syllable.

(71) Antepenultimate stress maps to penultimate stress in words with CV.CH penults

a. Step 1 | **Prosodification**: (see (68))
/gin*akaspo/ → [gi.n*a.kas$_{\mu}$.po]

b. Step 2 | **Stress assignment**: HAVESTRESS, WSP-PEN ≫ FAITHSTRESS
[gi.n*a.kas$_{\mu}$.po] → [gi.na'.kas$_{\mu}$.po]

c. Step 3 | **Coda Reduction**:
[gi.na'.kas$_{\mu}$.po] → [gi.na'.kah$_{\mu}$.po]

d. Step 4 | **Metathesis**
[gi.na'.kah$_{\mu}$.po] → [gi.na'.ka.ph$_{\mu}$o]

# 2	gi.n*a.kas$_{\mu}$.p|o	HAVESTR	WSP-PEN	FAITHSTR	FINSTR	NONFIN	RED.	MET.
a.	gi'.na.kas$_{\mu}$.p|o		*!		*		*	
☞b.	gi.na'.kas$_{\mu}$.p|o			*		*	*	
c.	gi.na.kas$_{\mu}$.p|o	*!		*			*	
# 3	gi.na'.kas$_{\mu}$.p|o	HAVESTR	WSP-PEN	FAITHSTR	FINSTR	NONFIN	RED.	MET.
d.	gi.na'.kas$_{\mu}$.p|o					*	*!	*
☞e.	gi.na'.kah$_{\mu}$.p|o					*		*
#4	gi.na'.kah$_{\mu}$.p|o	HAVESTR	WSP-PEN	FAITHSTR	FINSTR	NONFIN	RED.	MET.
f.	gi.na'.kah$_{\mu}$.p|o					*		*!
☞g.	gi.na'.ka.ph$_{\mu}$|o					*		

Output: [gi.na'.ka.pho] **Derivation:** /gin*akaspo/ → [gi.n*a.kas$_{\mu}$.po] → [gi.na'.kas$_{\mu}$.po] → [gi.na'.kah$_{\mu}$.po] → [gi.na'.ka.ph$_{\mu}$o]

In Step 2, penultimate stress is assigned. Lexically-marked antepenultiamte stress (candidate (a)) is ruled out by WSP-PEN, because the penult is heavy and unstressed. Candidate (b) is the winner: although stress is mapped unfaithfully to the penult, it satisfies WSP-PEN because the penult is heavy and stressed. Once stress is mapped, it is fixed and cannot move in subsequent steps (Pruitt 2010). In Step 3, coda /s/ reduces to [h], and in Step 4, metathesis occurs. (Gemination would also occur here, followed by degmination in the following step, as in the analysis in Chapter 3.)

The crucial part of this derivation is that stress assignment occurs before /s/ metathesizes out of the penult. The constraint that disprefers faithful mapping of antepenultimate stress is WSP-PEN, which can only force stress to surface unfaithfully on the penult if the penult is heavy. If metathesis were to happen first, stress would be assigned to the form [gi.na.ka.pho]. WSP-PEN would not be able to block faith to underlying stress location and /giˈ.na.kas.po/ would map to [giˈ.na.ka.pho]). That my listeners strongly dispreferred the faithful mapping indicates that the grammar does not produce this mapping. The interaction between syllable weight and stress assignment is opaque: although the resulting surface form has all light syllables, which should allow antepenultimate stress, this is blocked because stress is assigned on a form at an intermediate step of the derivation when the word still has a heavy penult.

Recall that I treat [Ch] sequences as tautosyllabic (onsets) rather than heterosyllabic. In Section 4.1.1, I gave phonotactic reasons. Now, I discuss a reason internal to the analysis itself. My analysis (and other discussions; e.g. Moya Corral and Tejada Giráldez 2020) treat metathesis as occurring in order to remove the consonant from coda position. If [Ch] syllabifies as a heterosyllabic cluster, metathesis fails to remove the consonant from coda position. Instead of /sC/, which syllabifies early as [s.C], the result would be [C.h]. [C.h] is no better than [s.C]. Actually, any word-medial coda stop is phonotactically *worse* than coda /s/. In an analysis that relies on harmonic improvement, the heterosyllabic syllabification *[C.h] is not possible because metathesis would not only fail to resolve the coda issue, but make it worse.

Although considering [Ch] as an onset sequence seems odd, since Spanish does not allow those clusters either, Steriade (1994) suggests that this may not be a problem. She proposes that the surface representation of a [Ch] sequence is the same, regardless of whether it is the realization of an aspirated stop or a cluster. She also argues that the more similar sequences of sounds are to segments, the less marked they are. Thus, a surface [Ch] sequence is more marked than an unaspirated stop, because it has an extra [h], but less marked than a cluster like [pr], which cannot be considered a single segment structurally. In the case of Sevillian, a [Ch] sequence may not be as bad as other clusters for this reason (see Section 6.2 in Chapter 6). Furthermore, there are languages

that have tautosyllabic stop-h sequences that are best analyzed as clusters (e.g. Acehnese, Old Khmer, and Taba; see Section 3.1.1)

4.5.5 Analysis summary

To summarize, listeners in the stress judgment task preferred words with antepenultimate stress that had light penults ([giˈ.na.ka.po]) over their identical counterparts that had heavy penults ([giˈ.na.kam.po], [giˈ.na.kas.po], [giˈ.na.kah.po]. Light penults derived from heavy penults (as in the case of metathesis, [giˈ.na.ka.pho]) patterned with words with underlyingly and surface-heavy penults. My analysis accounts for these experimental results. Listeners prefer [giˈ.na.ka.po] over [giˈ.na.kam.po] because their grammars can derive the first, but not the second. Antepenultimte stress is blocked by a heavy penult (*[giˈ.na.kam.po]), and will instead surface as penultimate. Listeners prefer [giˈ.na.ka.po] over [giˈ.na.ka.pho] for the same reason: their grammars can derive the former, but not the latter. At the point where stress applies, [giˈ.na.ka.pho] still has coda /s/ ([gi.na.kas.po]), and the heavy penult blocks antepenultimate stress. Ordering stress before metathesis is crucial for explaining why surface-heavy CVC penults are treated the same as the surface light CV.CH penults that are derived by metathesis.

4.6 Conclusion

In this chapter, I presented experimental results from a stress judgment task targeting the interaction between stress and metathesis in Sevillian Spanish. Based on the experimental results and following analysis, I made several points.

First, the results of the stress judgment task provide further evidence that surface [Ch] sequences derive from underlying /sC/ clusters. Sevillian listeners treat words with [CV.CHV] penults the same as words with surface-heavy penults [CVS.CHV], suggesting that [.Ch] has the same representation as [s.C] at the point when stress is assigned.

Second, stress assignment occurs before metathesis. The experimental results show that Sevillian listeners evaluate stress goodness based on a representation before metathesis has taken place, rather than on the surface form (✓ /ˈLLL/ → [ˈLLL] > */LHL/ → [ˈLLL]). Additionally, they judge antepenultimate stress words as equally bad when choosing between metathesis out of the penult (surface light, underlyingly heavy) and a closed penult (surface heavy, underlyingly heavy) (e.g. [giˈnakapho] ∼ [giˈnakampo]). This suggests that surface forms CV.CH and CVC.CV are representationally equivalent at the point when stress is assigned: they both have a consonant that is syllabified as a coda early in the derivation. The interaction between stress and metathesis is opaque.

Third, this opaque interaction between metathesis and stress is productive. Syllables with metathesis are heavy at early derivational stages, reflecting the underlying /sC/ representation, but light at later derivational stages, after metathesis has occurred. Listeners evaluate stress on nonce words as if metathesis had not applied, on a level of representation no longer visible on the surface. The avoidance of antepenultimate stress on words with heavy penults is apparent in the lexicon, and this experiment shows that listeners extend the pattern productively to nonce words that have derived surface-light penults.

CHAPTER 5

Cross-linguistic perception of [h]

5.0 Introduction

In the previous chapters, I have shown that Sevillian stop-h sequences are robust in both production and perception. Now, I investigate possible motivations for the change. Given that articulatory explanations and motivations do not seem to capture all aspects of the change (see Chapter 2), the current study investigates the role of perception in laryngeal metathesis—specifically, in the direction [hC] → [Ch]. I focus on proposals that metathesis is perceptually optimizing, and find no support for this hypothesis.

Metathesis has long been a challenge in phonology because it is difficult to capture in most phonological theories and does not always lend itself to phonetic explanations. Many scholars have considered metathesis to be rare, irregular, or non-existent (e.g. Webb 1974; Powell 1985). In extensive typological surveys, however, others have documented many systematic, synchronic metathesis patterns (Ultan 1978; Blevins and Garrett 1998; Blevins and Garrett 2004; Hume 2004; Canfield 2015).

Two main ways of analyzing metathesis have been proposed: teleological and non-teleological. In the teleological view, metathesis occurs for the purpose of optimizing the perception of one or both sounds involved (Grammont 1933; Steriade 2001; Hume 2004). In the non-teleological view, metathesis is not functionally motivated. Blevins and Garrett (1998) and Blevins and Garrett (2004: 118) argue that metathesis (and sound change more generally) may be phonetically natural, in the sense that properties of speech production and perception converge to give common sound changes, but optimization is not the cause or goal. Diachronically, they argue that metathesis may produce optimal phonetic forms by chance, through a mechanism of listener misperception.

Sevillian Spanish metathesis provides an test case for the role of perception because it is regular, synchronic, and productive. I have argued throughout this dissertation, with Torreira (2006) and Parrell (2012), that Sevillian [Ch] sequences are derived through gestural realignment of the [h] and stop gestures. While the articulatory description of this process is clear, the *motivation* is less so. Some accounts hold that the shift is driven by articulatory factors (Parrell 2012), while others point out that a purely articulatory motivation does not capture the phonetic details of metathesis, or the behavior of participants in perception tasks (Torreira 2012; Ruch and Harrington 2014: 14; Ruch and Peters 2016). This has led some researchers to propose that perceptual factors may also contribute.

Perceptual optimization has been proposed as a motivation for metathesis. Cross-linguistically, patterns of laryngeal metathesis, the rarity of preaspiration, and the alternations between preaspiration and postaspiration suggest that perception may be at play. From a functional perspective, the main explanation underlying both metathesis and alternations between preaspiration and postaspiration is that [h] is easier to perceive in some locations than in others. For example, Kingston (1990) and Silverman (2003) argue that stop releases are a salient place to dock [h], whereas coda [h] or [h] occurring after a sonorant may be less perceptually salient. Laryngeals frequently participate in metathesis (Yoon 2012), and many proposals argue that metathesis often puts [h] in a location where it is hypothetically easier to perceive (Flemming 1996; Cho 2012; Yoon

2012). Similarly for pre- and postaspiration, Steriade (1997) argues that distribution of aspiration often maximizes the perceptibility of [h].

From a non-functional perspective, [h] has been argued to have properties that make it particularly susceptible to undergoing metathesis. Blevins and Garrett (2004) and Hume (2004) both claim that [h] is a phonetically 'spread out' cue, which may make its location in the speech stream ambiguous, providing the pre-conditions for metathesis. If it is true that [h] is difficult to hear before a stop, and is difficult to place in linear order, then these factors could create the conditions necessary for metathesis.

If perceptual difficulties related to [h] facilitate and motivate metathesis at a cross-linguistic level, they should show up in experimental discrimination tasks. However, there is another possible explanation for listeners' behavior on these tasks: the perception of [h] may instead be controlled by native language experience, phonological inventories, and phonetics. Previous studies on naïve listener perception of a non-native language find that perception depends on how listeners map novel sounds to existing sound categories in their native language (e.g. Best and Tyler 2007). If metathesis is not driven by perceptual optimization, but rather results from language-specific perception, then perception of [h] in [hC] and [Ch] sequences should be driven largely by speakers' native language backgrounds. This account does not predict any universal perceptual property that would motivate metathesis, regardless of whether a language's grammar makes use of it.

In this chapter, I test the role of perceptual factors in laryngeal metathesis by investigating the perception of [h] in different locations. In a cross-linguistic ABX task, listeners are asked to distinguish the presence and linear order of [h] around different types of consonants. The experiment addresses the following broad questions:

- Is laryngeal metathesis [hC] $\rightarrow$ [Ch] perceptually optimizing?
- How are the patterns of [h] perception shaped by language-specific properties (e.g. phonemic categories)?

While I focus on possible perceptual motivations behind Sevillian metathesis, results from other language groups help distinguish universal vs. language-specific aspects of perception. The

experiment includes six listener groups: Arabic, English, French, Mexican Spanish, Argentinian Spanish, and Sevillian Spanish. These groups' native languages and varieties differ on two axes that could affect [h] perception: allowing coda [h], and having aspirated stops. To preview, results indicate that perception is likely not a motivating factor for Sevillian metathesis (or laryngeal metathesis more broadly), and that language-specific properties affect perception extensively.

The remainder of the chapter is structured as follows. Section 5.1 provides more information about why Sevillian metathesis is ideal to investigate motivations of metathesis, cross-linguistic approaches to perceptual optimization in the contexts of metathesis and pre/postaspiration, and the role of phonological categories and mappings in perception. Section 5.2 states the questions of the study and broad hypotheses. Section 5.3 describes the set-up of the experiment. Section 5.4 presents the results, which suggest that metathesis is not perceptually optimizing in Sevillian, and that perceptual factors would not facilitate metathesis for listeners of other languages, either. Instead, perception of [h] is controlled largely by native language background. Section 5.5 discusses the implications of these results for metathesis, and the crucial role of native language experience for the perception of [hC] and [Ch] sequences. I also discuss the implications of the results for pre/postaspiration. Section 5.6 concludes.

5.1 Background

5.1.1 Why study Sevillian metathesis and perceptual optimization?

Four factors make Sevillian a good test case for testing the difference between the perceptual optimization and phonological grammar account of metathesis: [Ch] sequences are frequent, derived by phonological process, and are limited to certain segments. Finally, the oft-proposed articulatory account of Sevillian metathesis has been argued to be insufficient, and perceptual hypotheses have been made, but not tested.

First, Sevillian listeners have extensive experience with [Ch] sequences. Metathesized sequences ([th] and an affricated variant [ts]) have been found to occur in over 70% of productions in conversational speech (Ruch 2008: 78; see Chapter 3, Section 3.4.1).

Second, Sevillian [Ch] sequences are derived, as I have argued in previous chapters. Sevillian listeners treat [Ch] sequences as separable, and map them onto underlying /sC/ clusters. Under this analysis, the [h] and [C] gestures metathesize, resulting in surface [Ch]. Furthermore, Sevillian speakers are exposed to extensive variability and phonological alternations, data which support the cluster analysis (see Chapter 3, Section 3.4.1).

Third, Sevillian metathesis is limited to clusters with voiceless stops as the second member, a limitation that could also be argued to be perceptually-based. Sequences with voiced stops (/sb sd sɡ/) and sonorants (/sm sn sl/) do not undergo full metathesis (see Chapter 2). But similar gestural changes may also affect these clusters, even if metathesis is not complete, as evidenced by phonetic changes in consonant duration and voicing (Hualde 1989a; Hualde 1989b; Morris 2000; Gerfen 2001; Campos-Astorkiza 2003; Martinez-Gil 2012; see also my phonetic results in Chapter 2). /s/-voiced stop sequences can be realized with the two segments coalesced (overlapped) into highly constricted fricatives that are voiced or voiceless, and geminate or singleton, depending on the dialect (/desde/: [deθːe] 'since'). Recall also that while /s/-sonorant sequences have been reported to be realized with partial devoicing (/isla/: [iḷ̥la] 'island'), I found no evidence of overlap; rather [h] deleted.

If metathesis is driven by perceptual optimization, Sevillian (and listeners of other languages too) should derive some perceptual benefit from [Ch] sequences over [hC] sequences. But there is a more nuanced hypothesis as well: accuracy may be modulated by the type of consonant adjacent to [h]. One hypothesis is that full metathesis is limited to /s/-voiceless stop clusters because metathesizing is only perceptually beneficial in these clusters. I further motivate this hypothesis in Section 5.1.2.2. To preview, the results of my perception experiment do not support this view, as metathesis around voiceless stops is *not* more beneficial than around other segments.

The results instead suggest that [h] perception in these sequences is driven by native language experience.

Finally, articulatory accounts of Sevillian metathesis fall short in several ways. In brief, the articulatory motivation proposed by Parrell (2012) is that gestural timing preferences *cause* Sevillian metathesis. The gestures in [Ch] sequences are in-phase (start together), whereas those in [hC] sequences are anti-phase ([h] starts before [C]) (Parrell 2012). Anti-phase timing is known to be unstable across human motor activities, including speech (Tuller and Kelso 1990), and Parrell (2012) argues that the pressure towards in-phase timing is a motivation for metathesis. Section 6.2 discusses this argument in more detail.

There are several reasons to believe that the cause is not purely articulatory. First, the temporal trade-off correlations are not as tight as would be expected if the change only involved gestural realignment (Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016). Several studies fail to find correlations between the duration of pre-posed and post-posed [h] (Torreira 2007; Ruch and Harrington 2014; Ruch and Peters 2016). Second, Parrell (2012) found that most speakers shift from anti-phase to in-phase coordination ([hc] $\rightarrow$ [Ch]) as speech rate increases. The pressure towards in-phase coordination increases at fast speech rates, so he takes this as evidence that the pressure for gestures to start in-phase is a motivation for metathesis. However, some participants produced in-phase coordination ([Ch]) even at slow speech rates, indicating that metathesis does not always require articulatory pressure in the moment of production, or that the pressure for in-phase coordination is not a causal mechanism for metathesis. Another possibility is that these speakers are most advanced in the metathesis change, so the retiming has stabilized, and is no longer driven by the original causal mechanism. Third, recall that full metathesis is limited to /s/-voiceless stop clusters. If the cause of metathesis were purely articulatory, we might expect it to apply to all clusters equally. Fourth, the purely articulatory account says nothing about the role of gemination in the change, since gesture length does not change (Torreira 2012; Ruch and Harrington 2014; Ruch and Peters 2016).

One more piece of evidence suggests that perceptual properties of [h] might facilitate metathesis. Ruch and Harrington (2014) found that Spanish-speaking listeners may parse [h] on a global level, without really localizing it in the speech stream. In their perception experiment, Argentinian listeners heard[V$\underline{h_1}$C$\underline{h_2}$] sequences. They perceived a lexical /s/ more frequently when the duration of *metathesized* [h_2] is long as compared to when it is short (the duration of [h_1] is held constant). These listeners have no native dialect familiarity with [Ch] sequences, and Ruch & Harrington hypothesize that they parse all of the [h]-like material as contributing to the same segment. Like the patterns of long-distance aspiration discussed in Section 5.1.2.2, this suggests that the location of [h] in the speech stream may be ambiguous: listeners may hear [h] on both sides of an intervening segment and parse it all together. Ambiguity in location makes reinterpretation in a non-historical location more likely. This finding motivates the condition in my experiment that tests how listeners perceive the location of [h] in relation to adjacent consonants. To preview, listeners are relatively good at perceiving the location of [h], going against the hypothesis that metathesis is facilitated because the location of [h] is difficult to perceive.

In the ABX discrimination task, I explore how perceptual factors could facilitate or motivate metathesis of [h]. The idea that metathesis is perceptually optimizing is supported by findings in two broad domains, which I discuss next: metathesis and pre/postaspirating languages (Section 5.1.2). Then, I discuss what a different approach—one based on mappings between surface and underlying forms—predicts for perception of [h] and metathesized sequences (Section 5.1.3).

5.1.2 Metathesis and perceptual optimization

5.1.2.1 Motivations and conditions for metathesis

Steriade (2001) argues that systematic, synchronic metathesis is perceptually optimizing. While she allows that metathesis could originate as simple listener misperception, she argues that assimilatory sound changes (including metathesis) that become systematic fulfill the properties in (72) (Steriade 2001: 12):

(72) Steriade's conditions for productive metathesis

a. The outcome is improved (articulatorily, perceptually, or in paradigm structure) over the original.

b. The outcome is 'tolerably similar' to the original.

She argues that metathesis patterns that do not provide perceptual improvement are not systematic, and that 'drastic dislocations' that are very perceptually different from the original are unattested (Steriade 2001: 14).

In the current study, I test Steriade's (2001) hypotheses focusing on Sevillian metathesis. I also test listeners of other languages, to distinguish what parts of [h] perception are universal, and which are determined by language experience.

5.1.2.2 Laryngeal metathesis and perception

Laryngeals are frequent participants in metathesis (Ultan 1978; Blevins and Garrett 1998; Blevins and Garrett 2004;Yoon 2012; Canfield 2015). Local laryngeal metathesis is common; reported cases include include Gujarati, Cherokee, Balangao, Cebuano, Yuman languages, Mandaic, Harari, Hungarian, Tübatulabal, Southern Estonian, Kiliwa, Hanunoo, Basaa, and Cayuga (Yoon 2012). Interested readers should refer to Yoon (2012) for typological overview and description of metathesis in these languages, and Canfield (2015: 271) for a list of attested synchronic metathesis patterns. Long-distance laryngeal metathesis is also attested. For example, Marathi had a diachronic change from Sanskrit, where aspiration shifted from the onset of the second syllable into word-initial position (Sanskrit: kaʈʰina-; Marathi: kʰaɖiːɳ 'hard/difficult') (Blevins and Garrett 2004: 133-4). Ultan (1978: 382) also reports for Marathi that, instead of being lost, the aspiration component on word-final stops was able to be retained through metathesis to word-initial position (kaːm̩kh̲ → kh̲aːm̩k 'armpit').

There are several reasons to believe that laryngeal metathesis might be perceptually motivated. One reason is that [h] is acoustically and perceptually weak, and metathesis can be argued to improve its perceptibility. A glottal [h] following vowel has weak cues: its fricative energy is

relatively dispersed throughout the spectrum and low intensity, and it has concentrations of energy in the same regions as the preceding vowel, so the vowel essentially 'masks' it (Bladon 1986; Ladefoged and Johnson 2011: 211). Metathesis provides a good solution for a weak glottal [h]: it preserves a sound that would otherwise be lost (Ultan 1978) by putting it in a more perceptible location.

Second, many languages have been argued to metathesize [h] only when doing so is perceptually advantageous. In these languages, metathesis depends on phonological context. For example, Korean has a process called *coalescence*, in which [h] merges with a following or preceding plain stop. When it merges with a following stop, the result is metathesis (73a) (Kim and Alderete 2008). However, when the adjacent sound is a sonorant or /s/, the [spread glottis] feature of /h/ is deleted. Instead, /h/ neutralizes with unreleased coda /t̚/, which then assimilates to the following nasal (73b) (Kim 1989; Cho 2012: 297).

(73) Korean coalescence

a.	/suh-talk/ → [su.tʰ ak]	'rooster'	(Kim and Alderete, 2008)
b.	/nah-ni/ → [nan:i]	'give birth to - Q'	(Cho, 2012: 297)

In Cherokee (Flemming 1996), laryngeal metathesis switches the order of a laryngeal feature and a vowel, so that the feature 'lands' on a preceding consonant (CVhX → ChVX).[1] Metathesis occurs only when the potential host is a voiceless stop (TVhT → ThT; 74a). When the potential host is a sonorant (NVhN), metathesis does not occur (74b).

(74) Cherokee laryngeal metathesis

a.	g-ahdiya → *k-diya*	'he is using it'	(Flemming, 1996: 30)
b.	d-uuna-hnuuwa	'they wear shirts'	(Flemming, 1996: 30)

For both Korean (Cho 2012) and Cherokee (Flemming 1996; Cho 2012), the authors argue that motivation for metathesis is perceptual: docking [h] on a voiceless stop improves the perceptibility of [h]. When it is not possible to dock a laryngeal feature in a perceptually optimal location

[1] Vowel deletion interacts with metathesis, such that the vowel preceding [h] deletes.

(e.g. when the available sound is a sonorant), the laryngeal feature deletes (Korean) or fails to metathesize (Cherokee).

[h] has also been argued to be susceptible to metathesis because it may affect the phonetics of surrounding segments. Blevins and Garrett (1998: 510-13) characterize laryngeal metathesis as a type of perceptual metathesis. For perceptual metathesis to occur, they argue that a phonetic feature must be long, affecting the phonetics of other segments in the word. They also argue that the source of the feature (in terms of which segment it belongs to or its location in the speech stream) must be ambiguous. Perceptual metathesis can occur in these cases because of the ambiguity in the speech signal that allows listeners to reinterpret it as originating from a different segmental source or from a different location (see also Ohala 1993). This view holds that metathesis results from a natural perceptual process, but is not inherently perceptually-optimizing.

5.1.2.3 Preaspiration and perception

Metathesis and pre/postaspirated stops differ in that the former is a phonological process, and the latter are segments whose aspiration feature can be realized on either side of the stop. In other words, pre- and postaspirated stops show metathesis within a segment. However, both metathesis and pre-/postaspiration involve similar surface sequences—[hC] or [Ch]— and both have been argued to have perceptual bases. Because of these parallels, I discuss preaspiration and postaspiration alongside metathesis.

In languages with preaspiration, sequences of a laryngeal and voiceless stop function as a contrastive unit. These include languages from the following families (Clayton 2010): Celtic (Scottish Gaelic, Irish), North Germanic (Icelandic, Faroese, Norwegian, Swedish), English, Finno-Ugric (Sami, Forest Nenets), Mongolic (Halh Mongolian), Uto-Aztecan (Toreva Hopi, Tohono O'odham), Algonquian (Central Algonquian, Fox), Tarascan, Romance (Sienese Italian), Arawakan (Chamicuro, Goajiro), Witotoan (Bora) (see also Silverman 2003).

Preaspiration is argued to be perceptually worse than postaspiration for reasons almost entirely parallel to those used to explain [hC] to [Ch] metathesis in the previous section. Postaspi-

ration is supposedly easier because the stop closure allows air to build up: when the closure is released, there is high intensity and volume of airflow, which provides a salient release (Bladon 1986; Kingston 1990; Silverman 2003). According to Bladon (1986), the stop closure also allows the auditory system to 'rest,' making it sensitive to whatever follows. In contrast, preaspiration occurs in the transition from the preceding vowel into a following stop, as the vowel is already ramping-down, and the fricative noise has similar shape as the preceding vowel. Furthermore, he considers [h] to be non-salient because of listeners' auditory adaptation to the preceding vowel. Additional support for the argument that preaspiration is perceptually difficult comes from the fact that languages either tend to lose preaspiration or strengthen it acoustically by adding an oral constriction (e.g. [hp ht hk] → [fp ɕt xk]; Silverman 2003; Clayton 2010). I return to this point in Section 5.5.2. The majority of work on pre/postaspiration argues that preaspiration is typologically uncommon because it is perceptually difficult (Bladon 1986; Silverman 2003).[2]

The allophonic distribution of preaspirated stops and postaspirated stops has been argued to optimize the perceptibility of [h] (Steriade 1997). In preaspirating languages, aspirated stops typically contrast with unaspirated stops, and preaspiration never contrasts with postaspiration (Ladefoged and Maddieson 1996: 73; Clayton 2010: 64). Rather, the location of aspiration has been argued to be conditioned by phonotactics (coda and cluster restrictions) and position in the word (Keer 1999; Suh 2001; Silverman 2003; Clayton 2010: 64), as well as the availability of acoustic cues (Steriade 1997). Icelandic is the most well-studied case. Icelandic plain voiceless stops contrast with aspirated voiceless (geminate) stops, which surface as pre- or postaspirated. Glossing over details, postaspirated stops occur word-initially while preaspirated stops occur in word-medial and final positions after short vowels (Ladefoged and Maddieson 1996: 70; Suh 2001). Preaspiration is expected to be perceptually difficult word-initially, because the cues for [h] cannot be fully realized without a preceding vowel, and in that position postaspiration occurs to make [h] more perceptible (Steriade 1997). Reported cases of word-initial preaspiration are extremely rare (Keer 1999).

[2]For a different perspective, see Clayton (2010), which I also further discuss in Section 5.5.1.1.

How does Spanish fit into this picture? Spanish dialects do not have contrastively pre- and postaspirated stops, but many do have [hC] sequences deriving from debuccalization in /sC/ clusters. Because the sequence is surface-similar to preaspirated stops, we might expect it to be subject to similar perceptual constraints. As Helgason (2002: 41) states, 'Two phonetically similar sound sequences that differ only in terms of phonological interpretation should not respond in different ways to the same auditory constraint. Such constraints must be applicable to sound patterns, irrespective of how these sound patterns are organised phonologically.'

Similarly to preaspiration and [hC] sequences that undergo metathesis in other languages, perception studies on Spanish speakers show that coda [h] is difficult to perceive. Listeners whose native varieties do not have coda [h] have low accuracy in identifying it (Schmidt 2013). Furthermore, even listeners whose native dialects *do* have regular debuccalization to [h] in coda position misperceive it more frequently than they misperceive other coda consonants (Schmidt 2013; Bedinghaus 2015). Finally, as mentioned earlier, Ruch and Harrington (2014) suggest that Spanish-speaking listeners may parse all [h] material together, without localizing it in the speech stream. Argentinian listeners, who have no experience with [Ch] sequences, are more likely to perceive coda /s/ when [h] following the stop is longer. This result cannot be explained by familiarity with metathesis, so Ruch and Harrington (2014) propose that listeners perceive [h], without caring about where it comes from in the speech stream.

5.1.3 An alternative: Language-specific perception

Studies presented in the previous section (Section 5.1.2) suggest that patterns of [h] perception may be universal, facilitating and and motivating metathesis at a cross-linguistic level. However, there is alternative: perception of [h] may be more controlled by listeners' native languages than by universal tendencies that happen to facilitate and motivate metathesis.

It is well known that properties of listeners' native languages influence perception (e.g. Werker and Tees 1984; Dupoux et al. 1999; Mielke 2003; Berent et al. 2007; Hallé and Best 2007; Davidson and Shaw 2012; Gallagher 2016). Mielke (2003) looks at the perception of [h] specif-

ically, testing how speakers of different languages perceive [h] in multiple consonantal contexts. His goal is to define the set of contexts in which Turkish [h] deletes—contexts that are difficult to unify phonologically—as being characterized by low perceptual salience. My experiment has a similar goal, while also considering the influence of native phonemic categories.

In the Perceptual Assimilation Model, as summarized in Best and Tyler (2007), naïve listeners categorize sounds in a non-native language based on both universal tendencies and properties of their own native language (contrastive categories and phonetics). Different categorization predicts differing degrees of discrimination between sounds. Some sounds are categorized as exemplars of a native category. If two sounds are assimilated to two different native categories, discrimination between those sounds is expected to be excellent, because the difference maps onto a native contrast. If two sounds map to the same native category, discrimination is expected to be poor, because the difference is not contrastive in the native language.

Additionally, listeners' perception of non-native sounds is influenced by the phonotactics of their native languages. Listeners are known to perceptually modify sequences that violate their own languages' phonotactics (e.g. Hallé et al. 1998; Kabak and Idsardi 2007; Dupoux et al. 2011), and repairs may be related to the native language inventory (e.g. Davidson 2010). Furthermore, listeners' accuracy in perception may be affected by what sounds are allowed in which positions in their language. For example, Whalen et al. (1997) finds that listeners disprefer inappropriate positional allophones in real words of their language, and have difficulty imitating them.

In relation to the current study, the existing literature suggests that listeners may perceive [h] based on their phonological inventories, as well as where [h] is allowed to surface (e.g. coda vs. onset, or both). For example, listeners whose native languages have [h] (e.g. Arabic listeners) may be better at perceiving [h] than listeners whose native languages do not have [h] (e.g. French listeners). Furthermore, listeners whose native languages have [h] in coda position may be better at perceiving [hC] sequences than listeners whose native languages do not, and listeners whose native languages have [Ch] sequences may be better at perceiving those. The boost may also depend on a listener's native phonemic categories. For example, English listeners may map [Ch] sequences to

an underlying phonemic category of aspirated stops, facilitating discrimination. Sevillian listeners, in contrast, have [Ch] but do not have an aspirated stop category. Their experience with surface derived [Ch] may not lead to higher accuracy in discrimination, since [Ch] does not correspond to a phonemic category for them.

5.2 The current study

The study probes two possible perceptual explanations for [h] metathesis:

(75) Two questions to be tested:

a. Perceptual optimization: Is metathesis of [h] driven by perceptual optimization?

i. Is [h] easier to perceive in [Ch] sequences than in [hC] sequences?

ii. Does the perceptibility of [h] in [Ch] sequences differ by adjacent consonant type (e.g. voiceless stop, voiced stop, etc.)?

iii. Are there differences in perceptibility of the linear order of [h] by the type of adjacent consonant?

b. Language-specific perception: Does native language experience affect the perception of [h]?

i. Does the presence of (a) aspirated stops and (b) coda [h] in listeners' native languages improve listeners' perception of [h] in [Ch] and [hC] sequences, respectively?

The first question (75a) encompasses Steriade's (2001) hypotheses that synchronic metathesis results in perceptual improvement (75a i, ii), but is not too different from the original form (75a iii). Metathesis with [h] may be facilitated and motivated on a universal level, due to general properties of perception. The second question (75b) considers an alternative hypothesis: that perception of [h] in [hC] and [Ch] sequences is driven by properties of listeners' native languages, instead of by cross-linguistic perceptual tendencies that could facilitate or cause metathesis. This question asks how listeners' ability to perceive [h] in [Ch] and [hC] sequences is modulated by the phonological categories of their native language. Listeners who have the relevant categories may find it easier to map the stimuli sequences to native phonological categories, leading to higher accuracy.

I address these questions in an ABX task run on listeners who are native speakers of of English, French, Arabic, and three varieties of Spanish (Mexican, Argentinian, Sevillian). These languages and varieties differ along two axes: having /h/ (or [h]) in the inventory, and having aspirated stops. The specific hypotheses tested are laid out in Section 5.3.4.

5.3 Experiment set-up

5.3.1 Task

Listeners carried out an ABX categorial discrimination task. Listeners heard three words and determined whether the final word (X) is the same as the first (A) or second (B) word. The ISI between words was 500ms (450ms between words, with 50ms padding at the beginning of each sound file).[3]

Listeners were given the following instructions, translated in to their native languages: 'You will hear three words presented in a row. In some of the triplets, the last word will be the same as the first. In other triplets, the last word will be the same as the second. Your task is to decide whether the LAST word is the same as the first or second.' They were also presented with the following illustration of the task:

(76) Example of task given to participants
'If you hear **pata** - bata- **pata**, the correct answer is '1', because the **last** word is the same as the **first**.
If you hear pata - **bata** - **bata**, the correct answer is '2', because the **last** word is the same as the **second**'.

Participants were asked to respond as quickly and accurately as possible after the three words finished playing. Instructions were printed in their native languages to encourage activation of that language.

[3]Davidson and Shaw (2012) found that manipulating ISI did not significantly affect results in an AX task, and used a single ISI of 500ms in their ABX task.

Before beginning, listeners had three practice trials that were the same kind as test and control trials, but with different items. After the experiment, participants filled out a demographic questionnaire.

The experiment was run in PCIbex (Zehr and Schwarz 2018) and distributed on Prolific, for all but Sevillian participants, who were recruited through personal contacts. The experiment took the participants, on average, between 25 and 35 minutes to complete. Participants were paid for their time, and Prolific participants were told they would earn $1 bonus for high accuracy to encourage attention to the task.

5.3.2 Materials

The stimuli were trisyllabic nonce words with penultimate stress. The medial consonants were /ptk, bdg, mnl, sf/, which were preceded and followed by a single vowel, /a/. /a/ was chosen because it exists in the languages of all listener groups. In total, there were 11 of these base items with the medial sequence /VCV/. To make the task less monotonous, nonce words also have onsets of a different phonological class than the medial target consonant (e.g. sonorants or fricatives for medial /ptk, bdg/). Sample stimuli are in Table 5.1. Within each word set, there were three words: NoH ([VCV]), HC ([VhCV]), and CH ([VChV]).

Table 5.1: Example stimuli items

Medial Consonant		NoH	HC	CH
Voiceless stops	/ptk/	famapa	famahpa	famapha
Voiced stops	/bdg/	namaba	namahba	namabha
Fricatives	/sf/	tanasa	tanahsa	tanasha
Sonorants	/mnl/	pakana	pakahna	pakanha
Corresponds to:			Sevillian-unmetathesized Preaspiration	Sevillian-metathesized Postaspiration

The ABX task included three types of comparisons (Table 5.2). The HC/CH comparison tests how listeners hear the linear order of [h] by asking them to make HC-CH-X comparisons.

The HC/C comparison tests how listeners distinguish [hC] from [C]. The CH/C comparison tests how listeners distinguish [Ch] from [C]. Within each comparison type, the stimuli were presented in four orders: AAB, ABA, BAB, BAA. Only one order of presentation is shown (ABA), but each comparison type appeared in all four orders for each word set and all listeners saw all orders for all word sets and comparisons.

Table 5.2: Sample trials for medial consonant /t/ in all three comparison conditions

Comparison	Order of Presentation		Example
HC/CH	HC-CH-HC	(ABA)	lanahta-lanatha-lanahta
HC/C	HC-None-HC	(ABA)	lanahta-lanata-lanahta
CH/C	CH-None-CH	(ABA)	lanatha-lanata-lanatha

There were a total of 132 test trials (11 medial consonants x 3 comparison types x 4 orders). There were also 72 control trials. Instead of manipulating the position of [h], the nonce words used in these trials manipulated the position of a nasal homorganic to one of six medial consonants (/ptk, bdg/). Besides replacing [h] with a nasal, these trials were identical to the test trials. For example, a control trial for HC/CH is *mafanta - mafatna - mafanta* (cf. the test trial in Table 5.2, *lanahta - lanatha - lanahta*).

The stimuli were recorded by a linguistically-trained, male native speaker of Turkish in a sound-proof booth, using a Shure KSM44 microphone with a pop filter.[4] Turkish was chosen as the stimuli language because it is not the native language of any participant groups, and because Turkish allows clusters with [h]. Although Turkish has [h] deletion in many contexts (Mielke 2003 and references therein), the speaker was instructed not to delete. The speaker read the stimuli words in a list, repeated five times. The two most natural-sounding productions of each word were chosen so that the X form would not be the same production of the same word as the comparison form in AB. This was done to ensure that listeners did not make their decision about X based on if it was the exact same sound file as one of the previous ones.

[4]We would like to thank the Department of Language and Literature at Norwegian University of Science and Technology for making their phonetics laboratory available to him to make the recordings.

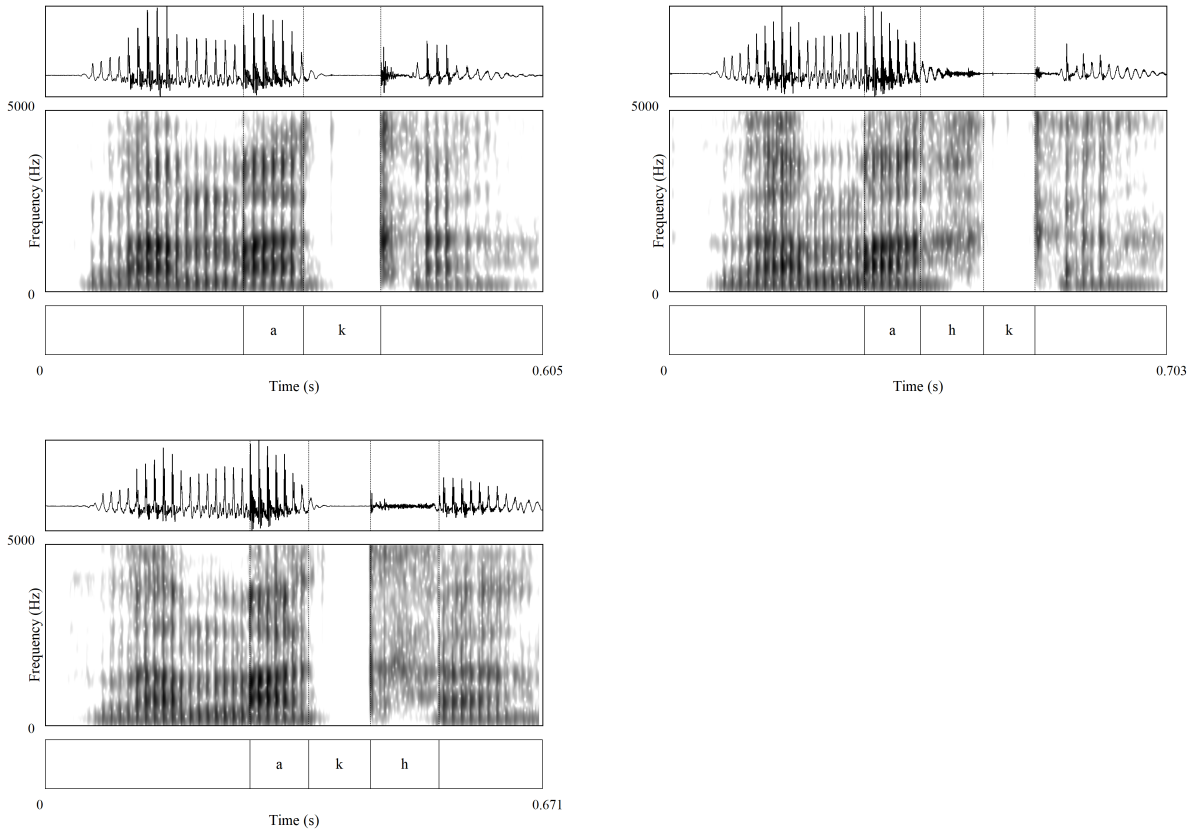
Test nonce words were manipulated. The manipulations were done within each set of words (e.g. *lanata, lanahta, lanatha*), because properties like vowel duration and [h] duration depend on the surrounding consonants.

- **Duration of the preceding vowel**: For each word in a given word set, I measured the vowel preceding [h] and [C] (e.g. *lanata*) and found the vowel that had the median duration. I then modified the vowels of the other words in the set to match that duration as closely as possible. Cuts and additions were made at zero-crossings in the waveform.
- **Duration of [h]**: The shortest [h] in the set was found. In the other members of the set with [h], the duration of [h] was shortened to match this value as closely as possible. The duration of [h] was set to the shortest rather than to the longest because pilot studies indicated that the task was too easy when set to the longest [h] duration. Even with [h] set to the shortest duration, accuracy was still high overall.
- **Pitch contour**: A single pitch contour was chosen from one of the words in each set, and overlaid over all the words in that set, replacing their pitch tiers. This was done so that listeners did not respond based on the similarity/difference of pitch contours in the set.
- **Intensity**: All words were normalized in intensity to 65 dB.

For items in control sets, the only manipulation was intensity normalization.

Spectrograms for a sample set of nonce words (*lamaka, lamahka, lamakha*) are shown in Figure 5.1.

Figure 5.1: Spectrograms of stimuli set *lamaka, lamahka, lamakha*



5.3.3 Listener groups

The listener groups were chosen so that the native languages and varieties differ in having HC sequences (coda [h], allophonic or phonemic) and CH sequences (aspirated stops or tautosyllabic CH sequences). This was done to test what factors of [h] perception are universal, and which are conditioned by language experience. The relevant properties of each language are outlined in Table 5.3.

Table 5.3: Properties of listener groups' native languages

Language	HC sequences (coda [h])?	CH sequences (aspirated stops or tautosyllabic sequence)
English	No	Yes - Aspirated stops
French	No	No
Arabic	Yes - /h/	No
Spanish - Mexico	No	No
Spanish - Argentina	Yes - [h] (← /sC/)	No
Spanish - Seville	Yes - [h] (← /sC/)	Yes - [Ch] sequence (← /sC/)

English has phonemic /h/, which is allowed in onset but not coda position, and aspirated stops that appear as contextual allophones of voiceless stops. Arabic has /h/, which is allowed in coda position, but lacks aspirated stops (Mustafawi 2018) and tautosyllabic [Ch] sequences ([h] is disallowed as C2 in an onset cluster). French does not have phonemic or allophonic [h], aspirated stops, or tautosyllabic [Ch] sequences. The varieties of Spanish differ in coda [h], which results from debuccalization of coda /s/. Mexican Spanish retains coda /s/ as [s] in many regions (Schmidt and Willis 2011; Hualde and Colina 2014: 289; Moreno-Fernández 2019), while Argentinian Spanish debuccalizes coda /s/ to [h] (Lipski 2011; Moreno-Fernández 2019). Sevillian Spanish also has coda /s/ debuccalization to [h]. None of the dialects have aspirated stops, but Sevillian has surface [Ch] sequences. While results from the fill-in-the-blank and stress judgment tasks suggest that Sevillian Spanish does not treat metathesized forms as contrastively aspirated stops, these listeners are familiar with surface [Ch] forms (phonetically very similar to aspirated stops).

5.3.4 Hypotheses

The broad hypotheses of the study are as follows, presented in conjunction with the predictions of two different accounts (repeated from (75)).

(77) **Perceptual optimization:** Is metathesis of [h] driven by perceptual optimization?

Hypothesis: Steriade's (2001) hypotheses about synchronic, productive metathesis, predicts metathesis to improve perceptibility and be similar to the original form. This hypothesis falls in line with other research arguing that the [Ch] order is perceptually better than [hC], both in the context of metathesis, and pre- and postaspiration.

Presence vs. Absence: In my ABX experiment, this account predicts listeners to perceive [h] better in the CH/C condition than in HC/C condition (HC and CH each in comparison to forms without aspiration), because [h] preceding a consonant is supposedly perceptually difficult on a universal level. Put differently, HC is expected to be be less distinct from C, than CH is from C. Additionally, the difference in accuracy between CH/C and HC/C conditions should be the largest for voiceless stops, if the reason metathesis often happens with those consonants but not others is because only they provide a sufficient perceptual boost.

Linear Order: Previous findings on perceiving [h] (see Section 5.1.2.2) predict that [h] will be difficult to locate in the speech stream, making the linear order comparison (HC/CH) difficult for all participants. The perceptual optimization account predicts that accuracy in the HC/CH condition should also be lower around voiceless stops than around voiced stops and sonorants, if the reason [h] metathesis occurs often in these contexts is because the outcome is more similar to the original than metathesis around other consonants.[5] These hypotheses should hold in all language groups, if the cause (or availability) of metathesis is due to general properties of the auditory system.

These predictions are schematized in Table 5.4.

[5]Linear order is expected to be most difficult around fricatives, because there is low modulation between [h] and fricatives (Ohala and Kawasaki-Fukumori 1997). Fricatives are included in the plots and statistical models, but I do not discuss them in detail.

Table 5.4: Predictions for perceptual optimization account of [h] metathesis

Hypothesis	Reason
CH/C > HC/C > HC/CH	H more salient in CH than in HC; hearing location of [h] is difficult
VcLessStop: CH/C > HC/C Other Cons: CH/C $\sim$ HC/C	Perceptual boost for HC $\rightarrow$ CH biggest for voiceless stops
HC/CH VcdStop, Sonorant > HC/CH VcLessStop	Linear order less distinct (more similar) around voiceless stops

(78) **Language-specific perception:** Does native language experience affect the perception of [h]?

Hypothesis: I limit my focus to the effect of native language on the perception of [h] around voiceless stops, putting aside other comparisons that are not directly relevant to the question at hand. The Perceptual Assimilation Model (see Section 5.1.3) predicts that listeners' accuracy in perceiving [h] will be directly related to the existence of sound categories in their native languages, and how the mapping between stimulus and this underlying category occurs. Perception may also depend on native language phonotactics: listeners are better at hearing sounds in locations that are phonotactically legal in their native languages. Hypotheses are schematized in Table 5.5. If the native language has neither [hC] or [Ch] sequences, the default prediction (based on the supposed perceptual inferiority of coda [h]) kicks in: [h] should be easier to perceive after a stop than before it (CH/C > HC/C).

Table 5.5: Hypotheses for effect of native language on perception of [h] around voiceless stops

Language	Hypothesis	Accuracy
English	CH/C > HC/C	Medium
	- Has aspirated stops; no coda [h]	
	English listeners > Other listeners on CH/C	
	- Only language with aspirated stops	
Arabic	HC/C ∼ CH/C	High
	- Has /h/; phonotactically permissive	
French	CH/C > HC/C	Low
	- No asp. stops; no /h/ in any position; general difficulty of HC/C	
Spanish-Mexico	CH/C > HC/C	Low
	- No asp. stops; no coda [h]; general difficulty of HC/C	
Spanish-Argentina	HC/C > CH/C	Medium
	- No asp. stops; coda /s/ debucc. to [h]	
Spanish-Seville	CH/C ∼ HC/C	Med-High
	- Experience with surface [hC] and [Ch] as allophonic realizations	
	Seville CH/C ∼ English CH/C	
	- Surface [Ch] seqs. in both, but differ in phonemic status	
	EITHER: CH/C > HC/C (like English)	
	- Apply experience with [Ch], even though not phonemic category	
	OR: HC/C > CH/C (like Argentina)	
	- Do not apply experience with [Ch] because not phonemic category	

If native language phonological inventories and phonotactics affect perception, English listeners should perform better on CH/C than on HC/C comparisons. English has aspirated stops, allowing them to map CH to a native category; [h] does not occur in coda position, so perceiving it in HC forms may be difficult. They should also outperform listeners of the other languages on CH/C because they are the only ones whose native language has aspirated stops to which CH can map. Arabic listeners are expected to have high accuracy in all conditions and no drastic difference between CH/C and HC/C, because their language has phonemic /h/ that occurs in many positions (although not in initial clusters). French listeners are expected to have low accuracy on both CH/C and HC/C, because French has no phonemic /h/ or allophonic [h] and no aspirated stops. In the

absence of available categories to map [h] onto, and phonetic experience with [h] from their native language, the general prediction of CH/C > HC/C kicks in. There is, however, another possibility: French listeners may have high accuracy in distinguishing [h] if it is uncategorizable, and thus judge the stimuli from a purely acoustic standpoint (see Section 5.5.1.2 for further discussion).

Listeners of different Spanish dialects have the same underlying representations, but different surface forms, so their perception behavior is expected to differ. Mexican Spanish listeners are expected to behave like French listeners: their dialect has no coda [h] or aspirated stops to map [h] onto.[6] In the absence of helpful experience, the general prediction of CH/C > HC/C takes over. Under an account where [h] perception is controlled more by native language experience than universal perceptual tendencies, Argentinian Spanish listeners should have higher accuracy for HC/C than CH/C. Coda [h] is familiar to them because their dialect regularly debuccalizes coda /s/ to [h]. In contrast, their accuracy on [Ch] should be low, because they do not have a phonemic category of aspirated stops to map [Ch] to.

Sevillian listeners could behave in several ways. They have experience with both [hC] and [Ch] surfaces sequences, but [Ch] sequences arise from gestural overlap in underlying /sC/ sequences. They are not realizations of an aspirated stop category. Sevillians could behave like English listeners, if they apply their experience with derived [Ch] sequences to the [Ch] sequences in the task. However, they might not apply this experience to this task—because the relevant piece may be having a *category* to map [Ch] to, not just familiarity with surface forms. In this case, they are expected to pattern with other Spanish listeners, specifically Argentinian listeners, whose variety also reduces coda /s/. A final possibility is that they are equally good (or bad) at both [Ch] and [hC], since they have experience with both as allophonic variants of the same underlying cluster.

[6]They may have some experience with dialects that debuccalize coda /s/ to [h], but in other studies, Mexican listeners have not been found to perceive coda [h] well (e.g. Schmidt 2013)

5.3.5 Participants

Arabic, English, French, Mexican Spanish, and Argentinian Spanish participants were recruited on Prolific (20 in each group). Sevillian participants (18) were recruited through personal contacts. Reported demographics for participant average age and sex (Table 5.6) exclude participants who were excluded for low accuracy (Section 5.3.6).

Table 5.6: Listener demographics (with exclusions)

			Sex		
Language	n	Age	Female	Male	Self-identified
English	18	31.61	15	2	1 (transgender male)
French	19	34.26	9	10	
Spanish-Mex	19	24.63	8	11	
Spanish-Arg	19	33.11	8	11	
Spanish-Seville	17	38.41	9	8	
Arabic	19	29.74	8	11	

English, French, Mexican Spanish, and Argentinian Spanish-speaking listeners were from the United States, France, Mexico, and Argentina, respectively. Arabic-speaking listeners were from Israel (3), Jordan (1), Lebanon (4), Palestinian Territory (3), and Syria (8); they are all speakers of Levantine Arabic varieties. English, French and Mexican Spanish-speaking listeners reside in their respective countries. Arabic and Argentinian Spanish-speaking listeners reside in other countries (mostly European). Sevillian participants were born, raised, and currently reside in Seville.

In terms of languages, most participants reported growing up only with their native languages, a filter that was applied in Prolific during recruitment and also held for Sevillian participants. However, two Arabic-speaking participants reported growing up with another language (English, and Hebrew and English). Nine Arabic-speaking listeners did not specify that they grew up with only their native languages, but did not report other native languages.

Participants also reported knowledge of other languages. Languages known at intermediate or higher proficiency are reported in Table 5.7. This language knowledge is not expected to help

listeners in the ABX task, since the known languages are mostly not more permissive regarding [h] than participants' native languages. Only one French speaker has potentially applicable knowledge from German. Argentinians might be the exception: their high English proficiency may have affected results, and this is further discussed in Section 5.5.1.

Table 5.7: Other languages known at intermediate+ proficiency

Listener group	Other languages
Arabic	English, Italian, French, German, Hebrew, Czech, Hungarian
English	French, Japanese, Spanish
French	English, German, Spanish
Spanish - Argentina	English, Italian, Catalan, French, Portuguese, Swedish
Spanish - Mexico	English
Spanish - Seville	English, Italian, Chinese

5.3.6 Statistical analysis

Statistical analyses were performed using linear mixed-effects models in R (RCoreTeam 2020) using the *lme4* package (Bates et al. 2015). For all models, the dependent variable is Accuracy (Incorrect, Correct), dummy coded as 0, 1. The relevant independent variables (baseline level underlined) are Language (Arabic, English, French, Spanish-Mexico, Spanish-Argentina, Spanish-Seville), Condition (HC/C, HC/CH, CH/C), and ConsonantType (Voiced Stop, Voiceless stop, Sonorant, Fricative). Models also had random intercepts of Participant and Word Set (e.g. the sets illustrated in Table 5.1).[7]

Based on these models, *emmeans* (Lenth 2020) was used to calculate estimated marginal means and contrasts between all levels of the relevant predictors. The results reported are from pairwise comparisons, (p-values adjusted using the Tukey method) and I show only comparisons with p-values less than or equal to .2. Model and *emmeans* coefficients are on the log-odds ratio scale (not the response scale).

[7] With one exception. The model run only on Arabic data could not fit a random intercept of Word Set.

A full model with a three-way interaction between Language, Condition, and ConsType did not run, and would have been too complex to interpret.[8] Instead, I ran models to address the specific hypotheses laid out in Section 5.3.4.[9] Results from the linear mixed-effects models are reported in Section 5.8. I report the relevant results from *emmeans* comparisons in the text.

To test question and hypothesis 1 (77), I ran the following models:

- **Condition model:** Accuracy $\sim$ Condition*Language

  This model tests how listeners of different languages perform on different Conditions (collapsing over ConsonantType).

- **ConsonantType models:** Accuracy $\sim$ Condition*ConsonantType (each language independently)

  These models test the effects of Condition and ConsonantType, within each language separately. The models test if the CH order is most beneficial for voiceless stops (if accuracy on CH/C for voiceless stops > CH/C for other consonants), and if linear order is most difficult around voiceless stops (if accuracy on HC/CH other consonants > HC/CH voiceless stops).

To test question and hypothesis 2 (78), I report the following model, *only on the data from voiceless stops*:

- **Voiceless Stop model:** Accuracy $\sim$ Language*Condition

  This model tests if the listener groups differ in how they perceive [h] around voiceless stops, differences which may be related to properties of their native languages.

Participants with accuracy less than 75% on control items were excluded from the results (Arabic = 1, English = 2, French = 0, Spanish-Mexico = 1, Spanish-Argentina = 1, Spanish-Seville = 1).

[8]The three-way interaction is plotted in Figure 5.7 in Section 5.7 for interested readers.

[9]Additionally, Order of Presentation of the pairs (ABB, BAA, ABA, BAB) has been found to affect results in previous studies (e.g. Davidson 2011). I included this predictor in one model, in interaction with Condition and Language. The interactions were significant. For all listener groups except Arabic, and for all conditions, the ABB and BAA orders had higher accuracy than ABA and BAB orders. This appears to be a recency effect. However, I omitted this predictor from the reported models because the results are basically parallel in the ABB/BAA and ABA/BAB conditions, except that accuracy in the ABB/BAA pairs is higher. Because all participants saw all items in all orders, both types of pairs contributed equally to each participant's overall results.

5.4 Results

I discuss the results in two parts, corresponding to the two questions and sets of hypotheses in Section 5.3.4.

5.4.1 Perceptual optimization motivation of metathesis

5.4.1.1 Broad results: Presence vs. Absence and Linear order

Figure 5.2 shows accuracy by Condition for each group of listeners. Model results (Section 5.8) show an interaction between Language and Condition, and selected pairwise comparisons (those with p-values <.2) from *emmeans* are shown in Table 5.8. For Arabic listeners, accuracy is overall high, and the difference between conditions is small. The only statistically significant difference is between HC/C and HC/CH: accuracy on HC/C is higher than on HC/CH. For English listeners, accuracy is similar in both HC/C and CH/C conditions, and statistically higher in these conditions than in HC/CH. The Spanish listeners all pattern similarly, regardless of dialect. All Spanish listener groups have higher accuracy in the HC/C condition than in HC/CH, and higher accuracy on HC/C than on CH/C. The difference between HC/CH and CH/C is not significant for any of them. French listeners differ from the other groups in that accuracy on HC/C and HC/CH conditions does not differ, but accuracy is higher in HC/C and HC/CH as compared to CH/C.

In other words, all listener groups except for French have the highest accuracy in HC/C, and substantially lower accuracy in the CH/C and HC/CH conditions.

Figure 5.2: Accuracy by Condition

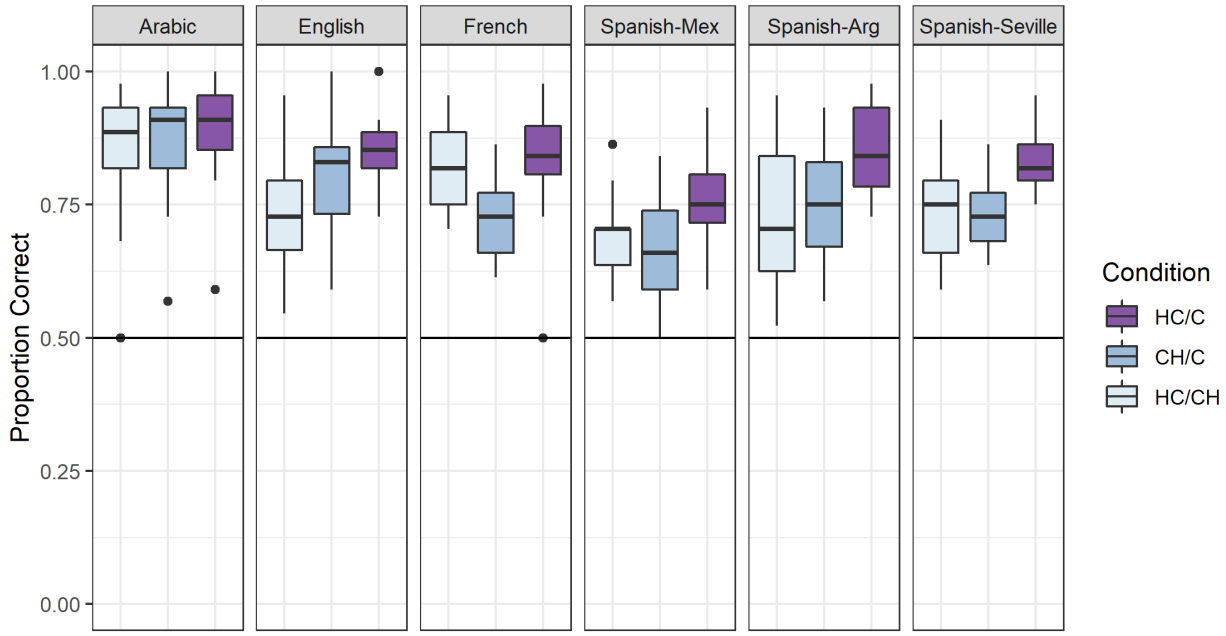


Table 5.8: Selected *emmeans* pairwise comparisons from the Condition model

contrast	Language	estimate	SE	df	z.ratio	p.value
(HC/C) - HC/CH	Arabic	0.42	0.15	Inf	2.74	0.02*
(HC/C) - HC/CH	English	0.77	0.13	Inf	5.95	0.00***
(HC/C) - (CH/C)		0.30	0.14	Inf	2.18	0.07.
HC/CH - (CH/C)		-0.48	0.12	Inf	-3.86	0.00***
(HC/C) - (CH/C)	French	0.75	0.12	Inf	6.06	0.00***
HC/CH - (CH/C)		0.66	0.12	Inf	5.42	0.00***
(HC/C) - HC/CH	Spanish-Mex	0.35	0.11	Inf	3.14	0.00**
(HC/C) - (CH/C)		0.50	0.11	Inf	4.47	0.00***
(HC/C) - HC/CH	Spanish-Arg	0.80	0.13	Inf	6.26	0.00***
(HC/C) - (CH/C)		0.64	0.13	Inf	4.98	0.00***
(HC/C) - HC/CH	Spanish-Seville	0.62	0.13	Inf	4.69	0.00***
(HC/C) - (CH/C)		0.65	0.13	Inf	4.98	0.00***

Table 5.9 summarizes the results in schematic form.

Table 5.9: Summary of comparisons for results collapsed across consonant type

Arabic	HC/C $\sim$ CH/C
	HC/C > HC/CH
	HC/CH $\sim$ CH/C
English	HC/C $\sim$ CH/C > HC/CH
French	HC/C $\sim$ HC/CH > CH/C
Spanish-Mexico	
Spanish-Argentina	HC/C > CH/C $\sim$ HC/CH
Spapnish-Seville	

5.4.1.2 HC/C and CH/C by Consonant

The hypothesis about the interaction between Condition and Consonant was that the CH order may provide a boost only for voiceless stops—in other words, that accuracy would be higher for CH/C than for HC/C around voiceless stops, but not around other consonants. Figure 5.3 shows accuracy for each language group by Consonant Type and Condition (only HC/C and CH/C shown). As described in Section 5.3.6, models run on each language independently tested the Condition*ConsonantType interaction. Pairwise comparisons for each language are reported in Table 5.10.

For Arabic listeners, accuracy on HC/C did not differ significantly from CH/C for any of the consonant types. For English listeners, HC/C had marginally higher accuracy than CH/C for VoicedStops. Numerically (but not statistically), voiceless stops and fricatives went in the same direction. For French listeners, HC/C had significantly higher accuracy than CH/C for Voiced Stops, Voiceless Stops, and Sonorants, but not for Fricatives. For Spanish, recall that the model included all Spanish dialects. There was a significant interaction between Dialect, Condition, and Consonant Type, but *emmeans* comparisons show that listeners of all dialects had higher accuracy on HC/C than on CH/C for Voiced and Voiceless Stops, but not for Sonorants and Fricatives.

Figure 5.3: Accuracy on HC/C and CH/C comparisons by Consonant Type

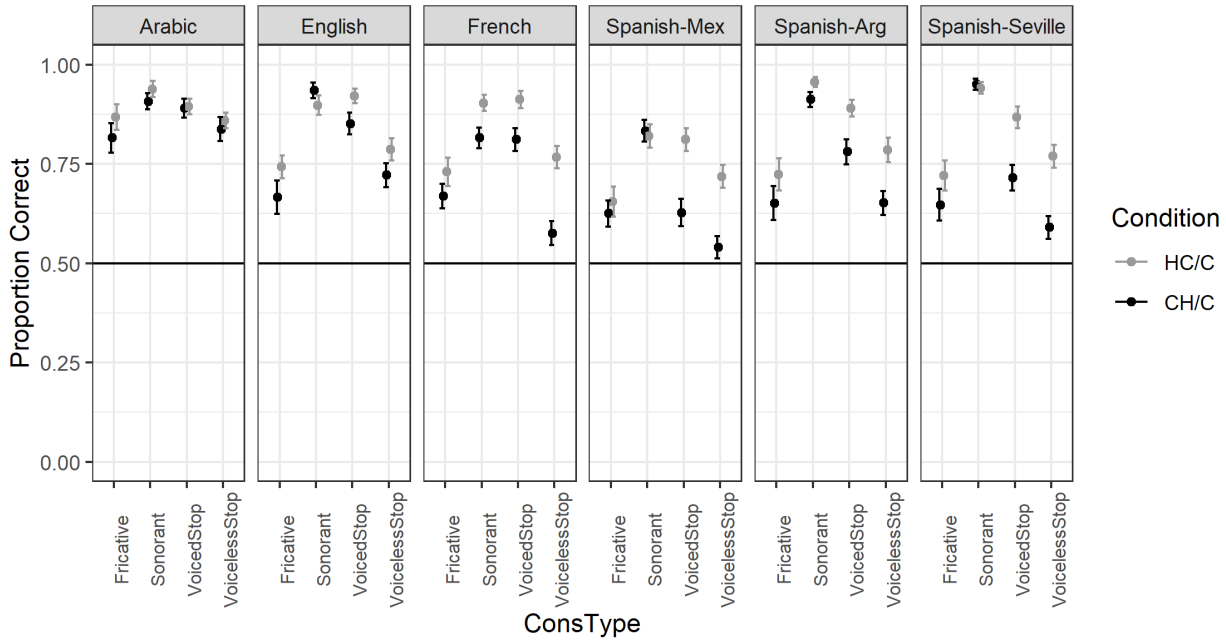


Table 5.10: Selected *emmeans* comparisons (accuracy on HC/C vs. CH/C) from a Condition*ConsType model on each language; Contrasts between Conditions within Consonant Types

	Contrast: HC/C vs. CH/C					
Language	ConsType	estimate	SE	df	z.ratio	p.value
English	VoicedStop	0.73	0.32	Inf	2.28	0.06.
French	VoicedStop	0.89	0.29	Inf	3.09	0.01**
	VoicelessStop	0.95	0.21	Inf	4.57	0.00***
	Sonorant	0.76	0.28	Inf	2.68	0.02*
Spanish-Mex	VoicedStop	0.97	0.22	Inf	4.39	0.00***
	VoicelessStop	0.78	0.20	Inf	3.89	0.00***
Spanish-Arg	VoicedStop	0.85	0.27	Inf	3.17	0.00**
	VoicelessStop	0.75	0.22	Inf	3.47	0.00**
	Sonorant	0.75	0.40	Inf	1.88	0.15
Spanish-Seville	VoicedStop	0.98	0.26	Inf	3.76	0.00***
	VoicelessStop	0.91	0.22	Inf	4.13	0.00***

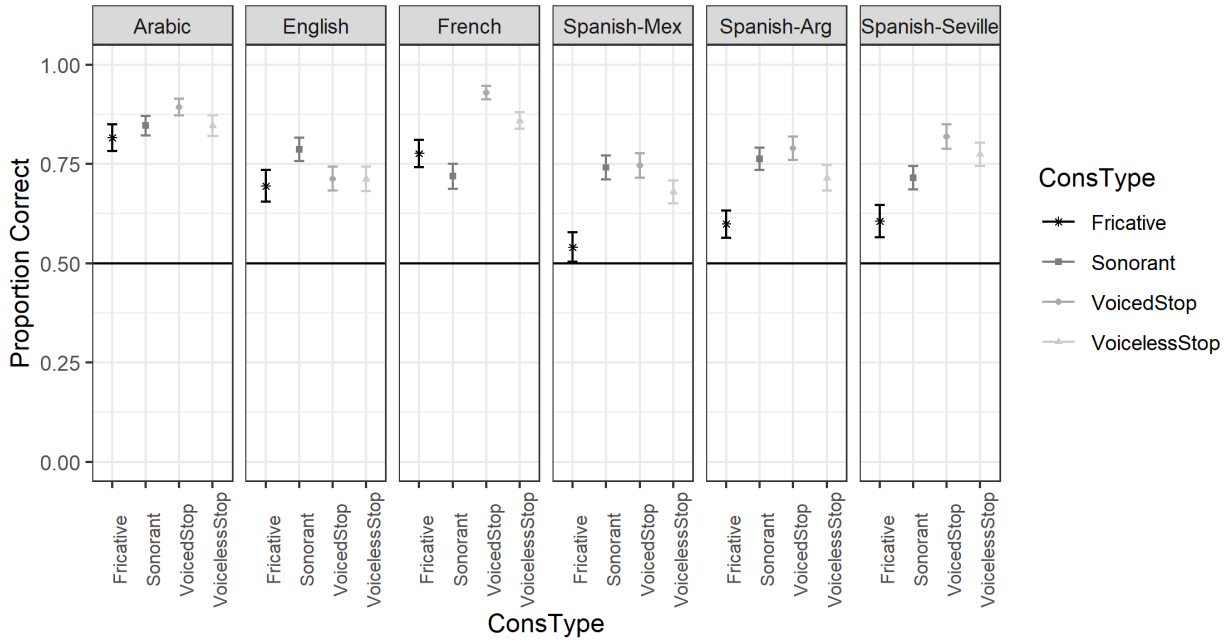
In sum, most listener groups (except for Arabic) are more accurate at perceiving [h] in HC/C than in CH/C around voiced and voiceless stops, but condition has much less of an effect around fricatives and sonorants. The difference between accuracy HC/C and CH/C conditions around stops is largest for Spanish listeners of all varieties and French listeners. Around fricatives and sonorants, most language groups have similar accuracy in both conditions. Against the perceptually optimizing account of metathesis, the CH order does *not* provide a clear perceptual benefit for [h]. Instead, the results show the opposite: HC is perceptually easier, and this benefit is most pronounced in voiced and voiceless stops. French listeners are again the anomaly: HC/C is easier than CH/C for them for all consonant types. I will later argue that this is due to the inability to map [h] to a phonemic category.

5.4.1.3 Linear Order by Consonant Type

The results just presented collapse over consonant type (Figure 5.2): HC/CH is either easier or equally difficult as CH/C. Crucially, the HC/CH comparison is *not* more difficult than CH/C. However, more specific hypotheses about linear order refer specifically to voiceless stops, so in this section I break down results by consonant type. The perceptual optimization hypothesis about linear order was that order might be more difficult to hear around voiceless stops than around other types of consonants (except fricatives, because of lack of modulation; Ohala and Kawasaki-Fukumori 1997).

Figure 5.4 shows accuracy on HC/CH by Consonant Type. Reported statistical differences come from the same models on individual languages discussed in the previous section (5.4.1.2) and reported in Section 5.8, which contain an interaction between Condition*ConsonantType. However, the calculated *emmeans* comparisons are different from those in Table 5.10. Instead of testing the contrast between conditions within each each consonant type, the set of comparisons in Table 5.11 tests the contrast between consonant types within each condition. Here, I only report on the linear order condition.

Figure 5.4: Accuracy on HC/CH by Consonant Type



For English and Arabic listeners, accuracy on HC/CH comparisons does not depend on consonant type (no comparisons are statistically significant). French listeners have higher accuracy on detecting linear order around Voiced Stops than around Fricatives and Sonorants, and higher accuracy around Voiceless Stops than around Sonorants. Mexican and Argentinian Spanish listeners show the same pattern as each other, with higher accuracy on linear order around Voiced Stops than around Fricatives, and lower accuracy around Voiceless Stops than around Sonorants. Sevillian listeners have higher accuracy around Voiced Stops than around Fricatives (like the other Spanish listeners), but also have higher accuracy around Voiceless Stops than around Fricatives.

Table 5.11: Selected *emmeans* comparisons from a Condition*ConsType model on each language; Contrasts between Consonant Types within Conditions (reported: only comparisons between consonant types within the HC/CH condition).

	Condition: HC/CH					
Language	contrast	estimate	SE	df	z.ratio	p.value
French	VoicedStop - VoicelessStop	0.78	0.33	Inf	2.36	0.08
	VoicedStop - Fricative	1.36	0.33	Inf	4.05	0.00***
	VoicedStop - Sonorant	1.67	0.31	Inf	5.42	0.00***
	VoicelessStop - Fricative	0.58	0.29	Inf	2.02	0.18
	VoicelessStop - Sonorant	0.89	0.25	Inf	3.51	0.00**
Spanish-Mex	VoicedStop - Fricative	1.00	0.28	Inf	3.57	0.00**
	VoicelessStop - Fricative	0.67	0.28	Inf	2.42	0.07.
	Fricative - Sonorant	-0.97	0.28	Inf	-3.48	0.00**
Spanish-Arg	VoicedStop - Fricative	0.97	0.29	Inf	3.34	0.00**
	Fricative - Sonorant	-0.81	0.29	Inf	-2.82	0.03*
Spanish-Seville	VoicedStop - Fricative	1.19	0.30	Inf	3.89	0.00**
	VoicedStop - Sonorant	0.60	0.28	Inf	2.13	0.14
	VoicelessStop - Fricative	0.90	0.30	Inf	3.03	0.01*
	Fricative - Sonorant	-0.58	0.29	Inf	-2.01	0.19

Table 5.12 provides a summary of the results. The crucial point is that linear order is *not* more difficult around voiceless stops than around other kinds of consonants, in contrast to the predictions of the perceptual optimization account of metathesis. This holds for listeners of all language backgrounds.

Table 5.12: Summary of significant differences in accuracy on HC/CH by ConsonantType

Language	Comparisons
Arabic English	VoicelessStop ∼ VoicedStop ∼ Sonorant ∼ Fricative
French	VoicedStop > Sonorant ∼ Fricative VoicelessStop > Sonorant
Spanish-Mexico	VoicedStop ∼ Sonorant > Fricative VoicedStop ∼ VoicelessStop ∼ Sonorant
Spanish-Argentina	VoicedStop > Fricative; Sonorant > Fricative VoicedStop ∼ VoicelessStop ∼ Sonorant
Spanish-Seville	VoicedStop ∼ VoicelessStop > Fricative VoicedStop ∼ VoicelessStop ∼ Sonorant; Sonorant ∼ Fricative

5.4.2 Effect of native language on [h] perception

This section reports results for voiceless stops only. I discuss each language group in separate sections, based on Figures 5.5 and 5.6. As a reminder, the model was run just on voiceless stops and tested the interaction between Condition*Language (full results in Section 5.8). As in the previous section, results reported are from two sets of *emmeans* pairwise comparisons (Table 5.13, Table 5.14), and will be referenced throughout this section.

Figure 5.5: Accuracy around Voiceless Stops by Language

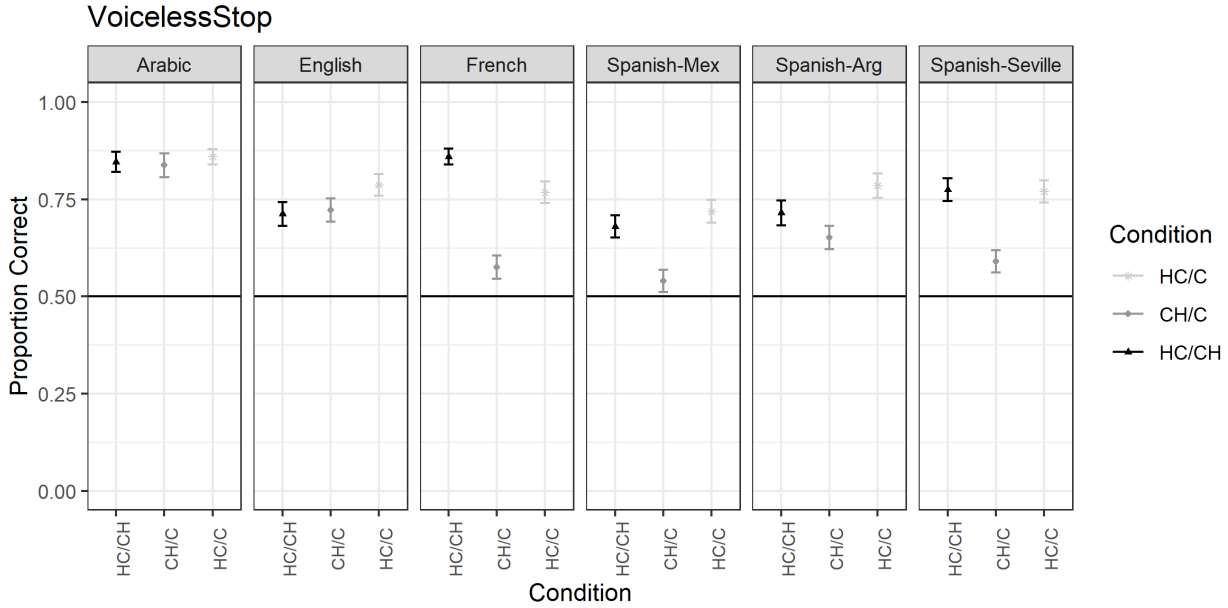


Figure 5.6: Accuracy around Voiceless Stops by Condition

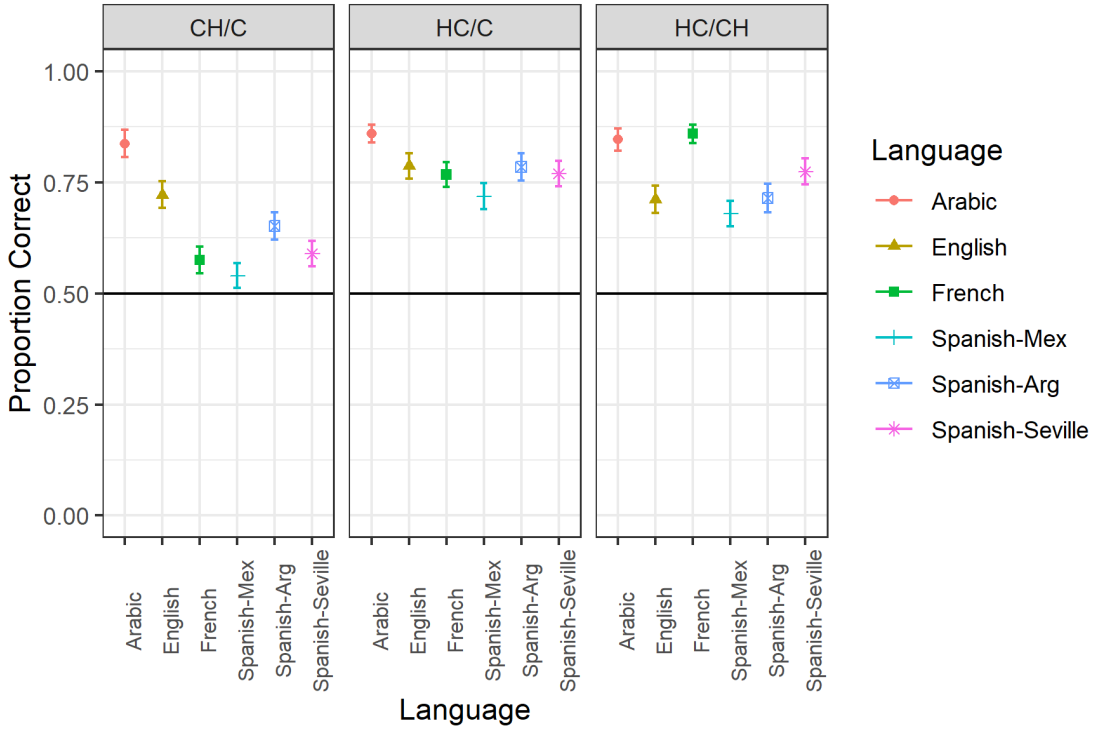


Table 5.13: Selected *emmeans* comparisons from VoicelessStop model: Contrasts of Condition within Language

Language	contrast	estimate	SE	df	z.ratio	p.value
English	(HC/C) - HC/CH	0.47	0.22	Inf	2.10	0.09.
French	(HC/C) - HC/CH	-0.62	0.25	Inf	-2.52	0.03*
	(HC/C) - (CH/C)	0.94	0.21	Inf	4.55	0.00***
	HC/CH - (CH/C)	1.57	0.23	Inf	6.71	0.00***
Spanish-Mex	(HC/C) - (CH/C)	0.77	0.20	Inf	3.86	0.00***
	HC/CH - (CH/C)	0.64	0.20	Inf	3.27	0.00**
Spanish-Arg	(HC/C) - HC/CH	0.38	0.22	Inf	1.74	0.19
	(HC/C) - (CH/C)	0.73	0.21	Inf	3.42	0.00**
	HC/CH - (CH/C)	0.35	0.20	Inf	1.72	0.20
Spanish-Seville	(HC/C) - (CH/C)	0.90	0.22	Inf	4.11	0.00***
	HC/CH - (CH/C)	0.93	0.22	Inf	4.22	0.00***

Table 5.14: Selected *emmeans* comparisons from voiceless stop model: Contrasts of Language within Condition

Condition	contrast	estimate	SE	df	z.ratio	p.value
HC/C	Arabic - French	0.63	0.26	Inf	2.43	0.15
	Arabic - (Spanish-Mex)	0.95	0.25	Inf	3.77	0.00**
	Arabic - (Spanish-Seville)	0.62	0.27	Inf	2.32	0.18
HC/CH	Arabic - English	0.88	0.25	Inf	3.51	0.01*
	Arabic - (Spanish-Mex)	0.97	0.25	Inf	3.95	0.00**
	Arabic - (Spanish-Arg)	0.80	0.25	Inf	3.22	0.02*
	English - French	-0.98	0.25	Inf	-3.84	0.00**
	French - (Spanish-Mex)	1.07	0.25	Inf	4.27	0.00***
	French - (Spanish-Arg)	0.90	0.25	Inf	3.55	0.01**
CH/C	Arabic - English	0.70	0.25	Inf	2.80	0.06.
	Arabic - French	1.40	0.24	Inf	5.88	0.00***
	Arabic - (Spanish-Mex)	1.55	0.24	Inf	6.50	0.00***
	Arabic - (Spanish-Arg)	1.08	0.24	Inf	4.50	0.00***
	Arabic - (Spanish-Seville)	1.35	0.24	Inf	5.53	0.00***
	English - French	0.70	0.22	Inf	3.22	0.02.
	English - (Spanish-Mex)	0.85	0.22	Inf	3.89	0.00**
	English - (Spanish-Seville)	0.65	0.22	Inf	2.90	0.04.

5.4.2.1 Arabic

We have already seen that Arabic listeners have the highest accuracy among groups, across conditions (Figure 5.2), and that their accuracy does not differ significantly in HC/C and CH/C for any consonant type (Figure 5.3). Considering just voiceless stops (Figure 5.5), Arabic listeners have equally high accuracy for voiceless stops in all conditions (*emmeans* comparisons between all levels non-significant; Arabic is therefore absent from Table 5.13).

Arabic listeners also have higher accuracy than all other language groups for voiceless stops in the CH/C condition (marginally, in comparison to English listeners). For HC/CH, they have higher accuracy than English, Mexican, and Argentinian listeners, and for HC/C they have higher accuracy than Mexican listeners (Figure 5.6, Table 5.14). These results are largely as predicted by the phonological mapping hypothesis: Arabic has /h/ and is phonotactically permissive, which likely facilitated [h] perception because [h] could be mapped onto a native category.

5.4.2.2 English

English listeners fall in the middle in terms of overall accuracy (Figure 5.2). Figure 5.5 shows that, around voiceless stops, English listeners perceive HC/C and CH/C equally well (*emmeans* comparison not significant; thus absent from Table 5.13). Furthermore, the difference between HC/C and CH/C appears to be smaller than for Spanish dialects and French, where CH/C is much harder than HC/C. Finally, Figure 5.6 (middle panel) shows that English listeners have higher accuracy on CH/C for voiceless stops than French, Mexican, Sevillian listeners, and marginally lower accuracy than Arabic listeners (selected comparisons in Table 5.14).

These results are mostly consistent with the phonological mapping hypothesis, but show some effect of the apparent universal goodness of [hC] sequences as well. English does not allow coda [h], but listeners are relatively good at the HC/C comparison anyway. English-speaking listeners' experience with aspirated stops is enough to boost accuracy on CH/C to the same level as HC/C, but not to exceed it. Still, it is notable that native-language background helped them in

this way: they are the only group of listeners (except for Arabic listeners), that have no difference between conditions CH/C above HC/C.

The accuracy of French and Argentinian listeners does not differ statistically from that of English-speaking listeners on CH/C, but, unlike for English speakers, these groups do show a difference between the HC/C and CH/C conditions. French and Arabic speakers' relatively high accuracy on CH/C is discussed further in Section 5.5.1.

5.4.2.3 French

French listeners were expected to have low accuracy overall, since French lacks [h] as a segment and feature, but this prediction does not hold. French listeners had much higher accuracy than anticipated overall (Figure 5.2), and in specific conditions (e.g. Figures 5.5 and 5.6). The only part of the hypothesis that holds is that French listeners have lower accuracy around voiceless stops on CH/C than on HC/C and Linear Order (Figure 5.5; Table 5.13). French listeners also have lower accuracy on CH/C than English and Arabic listeners, and similar accuracy as the Spanish-speaking listener groups (Figure 5.6; Table 5.14). This result is somewhat unexpected, given that French has neither aspirated stops nor /h/. However, I later suggest that the inability to map [h] to a category may have actually led to better perception.

5.4.2.4 Spanish dialects

Listeners of Spanish varieties were predicted to perform differently based on properties of their native dialects. Although the dialects share phonological inventories and underlying representations, the phonological grammars produce different surface mappings.

For all dialect groups, overall accuracy on HC/C is higher than CH/C and HC/CH, which are similar to each other (Figure 5.2). For voiceless stops, Spanish listeners have significantly higher accuracy in HC/C than in CH/C (Figure 5.5; Table 5.13). Mexican and Sevillian Spanish listeners also have significantly lower accuracy on CH/C than English and Arabic listeners (Figure

5.6; Table 5.14). Argentinian listeners do not differ significantly from English listeners on CH/C comparisons.

Listeners of Spanish dialects were predicted to differ in several ways, because their phonological grammars are clearly different, but they behaved surprisingly similarly. First, around voiceless stops, the listener groups behave similarly on accuracy for HC/C comparisons(Figure 5.6; Table 5.14). Argentinian and Sevillian listeners have extensive experience with coda [h] resulting from debuccalization and were expected to perform well. That Mexican listeners performed on par with them was unexpected, since they do not have experience with coda [h].

Second, Sevillian listeners behaved very similarly to Argentinian listeners in terms of CH/C: both groups showed higher accuracy in HC/C than in CH/C, and the difference between the conditions was large for voiceless stops (Figure 5.3).

A comparison between Sevillian and English listeners is also relevant. Both Sevillian Spanish and English have [Ch] forms, although they correspond to different underlying representations. In the ABX task, Sevillian listeners' accuracy on CH/C was lower than English listeners', reflecting this difference in representation and mapping between surface and underlying forms. Furthermore, in contrast to Sevillian listeners, recall that English listeners were equally accurate in both the HC/C and CH/C conditions around voiceless stops (Figure 5.3). Their experience with aspirated stops boosted accuracy in the CH/C comparison to be similar to their accuracy in the HC/C category, whereas Sevillians did not get the same boost in accuracy from experience with metathesized [Ch] sequences.

Third, listeners of all Spanish dialects behaved similarly on HC/CH comparisons. Sevillian listeners had similar accuracy to Mexican and Argentinian listeners on HC/CH around all consonants (Figure 5.4) and around stops specifically (Figure 5.6, left panel; Table 5.14). Even though their dialect has both orders, Sevillians were neither better nor worse at distinguishing linear order than listeners without experience with both orders.

5.4.3 Results summary

The results from the ABX task do not support the hypothesis that metathesis occurs in order to optimize the perception of [h], and largely support the hypothesis that [h] perception is mostly driven by native language perception.

Perceptual optimization:

Recall the two hypotheses about conditions of systematic, synchronic metathesis: that it improves perceptibility and that the result is not too perceptually different from the original (see Section 5.1.2.1). The results of my study do not support either of these conditions, for Sevillian Spanish or cross-linguistically. Listeners of all groups had higher accuracy in perceiving [h] in HC/C comparisons than in CH/C comparisons; the CH order did not improve perceptibility, even for voiceless stops. Additionally, the linear order of [h] is not more confusable around voiceless stops than around other consonants. Metathesis with [h] occurs mostly around voiceless stops cross-linguistically, so we might have expected [hC] and [Ch] to be more confusable when C is a voiceless stop, if the result of synchronic metathesis must sound similar to the original. The experimental results did not support this hypothesis.

The results do not support [hC] sequences as perceptually worse than [Ch] sequences, and do not support perceptual optimization as the cause for laryngeal metathesis from [hC] $\rightarrow$ [Ch] in Sevillian Spanish or cross-linguistically. Instead, the opposite was true: in my experiment [h] is *easier* to perceive in [hC] sequences than in [Ch] sequences.

Language-specific perception:

The results largely support the language-specific perception hypothesis: perception of [h] is controlled more by native language background than by general perceptual patterns that would facilitate or motivate metathesis. As predicted, Arabic listeners had high accuracy in all conditions. Also as predicted, English listeners got a boost in perceiving [h] in CH contexts in comparison to

other listener groups, although this boost merely raised accuracy on CH/C to the level of HC/C, without surpassing it.

Several other results also support the phonological mapping hypothesis, although less straightforwardly. The results for Spanish listeners are partially in line with the hypotheses. As predicted, Mexican Spanish listeners, who lack coda [h], had lower accuracy overall than Argentinian and Sevillian listeners, whose dialects do have coda [h]. More surprisingly, however, Sevillian listeners pattern most closely with Argentinian listeners—there is nothing special about them. Contrary to predictions, they do not get a boost in perceiving [h] in HC or CH contexts, and their accuracy on linear order is similar to that of Argentinians'. This is striking because the surface forms of Sevillian Spanish differ drastically from those of other Spanish dialects. Finally, the results for French listeners seem to be contrary to predictions: these listeners had high accuracy for many conditions and consonant types, but French lacks aspirated stops and /h/. However, I will suggest that the complete inability to map the forms in the experiment to native categories leads them to treat the forms in the experiment differently, resulting in high accuracy.

In the remainder of this section, I discuss why listeners of different languages performed differently than straightforwardly expected (Section 5.5.1). The results point to the importance of mappings from surface forms to underlying forms (not just phonetics) and a more nuanced look at the details of mappings suggests that the results are in line with the phonological mappings hypothesis. I also discuss the implications of the results for our understanding of the perceptual factors in preaspirating languages and metathesis (Section 5.5.2). Finally, I discuss some limitations of the current study (Section 5.5.3).

5.5 Discussion

5.5.1 Perception is about more than phonetics: it's about native categories

The effect of native language is likely driven not only by experience with surface phonetic forms, but by how these forms are mapped to underlying representations. Here, I take a more nuanced

look at the expected mappings for each language group, schematized in Table 5.15 . In this section, I focus only on /s/-voiceless stop sequences.

Table 5.15: Expected mappings given the ABX experiment input

Heard	Arabic	English	Spanish-Mex & Arg	Seville	French
[hC]-[C]	/hC-C/	/hC/ - /C/	/sC/ - /C/	/sC/ - /C/	/C/ - /C/
[Ch]-[C]	/Ch-C/	/C^h/ - /C/	/C/ - /C/	/sC/ - /C/	/C/ - /C/
[hC]-[Ch]	/hC-Ch/	/hC/ - /C^h/	/sC/ - /C/	/sC/ - /sC/	/C/ - /C/

Arabic listeners likely map [h] onto their native-category phonemic /h/, regardless of position. This mapping is straightforward, and these listeners would have been contrasting hypothetical words with contrasts like /hC/-/C/, /Ch/-/C/, /Ch/-/hC/. In contrast, for English listeners, [h] could be either the aspiration component of a voiceless stop, or the realization of phonemic /h/. In [hC] sequences [h] must be the realization of phoneme /h/; it appears in a phonotactically illegal position though, which likely hampers perception. In [Ch] sequences, English listeners likely map [h] as an aspiration feature on a voiceless stop /C^h/.[10] These listeners were thus comparing phonetic forms that mapped onto consonant sequences ([hC]) to those that mapped onto aspirated stops. The fact that [h] could either be a segment or a feature may have contributed to lower accuracy for this group than for Arabic listeners, because English listeners had to consider multiple possible mappings. Although English listeners' accuracy on CH/C was higher than other groups', as expected, it was still relatively low considering that English was the only language with aspirated stops.

Spanish listeners of all dialects behaved similarly in the ABX task, despite drastically different surface forms. Sevillians' experience with [Ch] surface forms did not help them, because the task depends on more than simple perception: it depends on the mappings from surface to underlying forms and accessing lexical representations. These mappings were expected to differ for Sevillian listeners, but apparently did not.

[10]English does have heterosyllabic [C.h] sequences, e.g. *top-hat*, but I think it is unlikely that they interpreted the stimuli in this way. To a native English speaker, the stimuli do not sound heterosyllabic.

Mexican and Argentinian Spanish listeners should have roughly the same mappings, although with a difference in accuracy. [hC] should map to an underlying /sC/ sequence for both groups, but perceiving [h] in [hC] sequences was expected to be easier for Argentinian listeners—whose dialect has coda /s/ debuccalization to [h]—than for Mexican listeners. This result held, and is consistent with results found in Schmidt (2013) and Bedinghaus (2015). [Ch] forms are expected to map to underlying plain consonants ([h] should be disregarded). In Seville, [hC] should map to /sC/ as well. However, [Ch] should *also* map to /sC/. In the ABX task, Spanish listeners are comparing allophonic realizations of /sC/ sequence to a plain consonant /C/.

French listeners' had unexpectedly high accuracy, indicating that the expected mappings in Table 5.15 did not hold. These listeners were able to distinguish the presence of [h] and its linear order at rates on par with other listeners of other language groups. It appears that, since they do not have experience with phonemic categories (or even experience with phonetic surface forms with [h]) to guide them, they may have taken a more purely phonetic approach to the task, a possibility I take up in Section 5.5.1.2.

In short, although all listeners heard the same stimuli, they were doing quite different comparisons based on the properties of their native languages. Arabic listeners were comparing the presence and location of a phoneme /h/, and English listeners were comparing the presence and location of what could be either a feature or a phoneme. Mexican and Argentinian Spanish listeners were comparing allophonic realizations of underlying coda /s/, some of which are not present in their dialects, to plain voiceless stops. In other words, they are not even comparing different stops to each other—they are comparing two different segments (/s/ and a voiceless stop). Sevillian listeners were expected to map [Ch] sequences onto underlying /sC/ as well, differently from the other Spanish groups. This expected mapping from [Ch] $\rightarrow$ /sC/ did not hold. If it had, Sevillians should have behaved differently from the other Spanish groups on CH/C and HC/CH comparisons.

The Argentinian listeners had surprisingly high accuracy on the CH/C comparison, and this is possibly due to L2 proficiency in English. Only two of these listeners currently reside in an English speaking country (the UK), and all others reside in countries where the ambient languages

do not have aspirated stops (e.g. Spain, Italy, Hungary), although some of these languages do have /h/ (e.g. Hungarian, some varieties in Spain). However, all Argentinian participants reported advanced to near-native fluency in English, which could have boosted their accuracy somewhat on CH/C comparisons. Recall that their accuracy on CH/C around voiceless stops was lower than that of English listeners (although not much), and was not significantly different than Spanish listeners from Mexico and Seville, who did not report advanced knowledge of English en masse.

5.5.1.1 Why are Sevillians the same as other Spanish listeners?

Sevillian listeners' accuracy on the CH/C comparison is no higher than other Spanish listeners' accuracy, who have no experience with [Ch] sequences. In fact, Sevillians have numerically—although not significantly—*lower* accuracy than Argentinian listeners. In the ABX experiment, both [Ch] and [C] must have mapped to [C], leading to low accuracy in discrimination.

Like other Spanish speakers, Sevillians may have mapped [Ch] sequences onto underlying /C/, not /sC/. This seems incompatible with evidence from the fill-in-the-blank experiment in Chapter 3, and from other studies that show that Sevillians use the length of [h] following [C] to distinguish /sC/-/C/ contrasts (e.g. Ruch and Peters 2016). Although Sevillians have extensive experience with [Ch] sequences acoustically, they did not extend this knowledge to the ABX context. Why?

Previous results show that Sevillians use their knowledge of post-posed [h] in three contexts: in distinguishing a minimal pair of their language (e.g. *pata-pasta*, [pata-patha], 'paw-pasta'; Ruch and Harrington 2014), in distinguishing a morphological contrast that relies on it (e.g. *tiene pali - tjenes pali*, [tjene pali - tjene phali], '3SG has *pali* - 2SG has *pali*'; Chapter 3), and in judging stress patterns on nonce words (Chapter 4). What these contexts have in common is that there is an explicit possibility that [Ch] sequences arise from /sC/ clusters in one member of the comparison. The first two tasks involve lexical representations and ask listeners to distinguish between /C/ and /sC/ representations. The stress judgment task uses nonce words, which have no lexical representation. However, the task presents multiple surface forms of each word ([gi'nakaspo, gi'nakahpo,

giˈnakapho]), essentially showing participants that the metathesized form [giˈnakapho] is an allophonic realization of a word with an underlying /sC/ cluster. This experimental context reflects the variability present in daily speech, where speakers are exposed to multiple surface variants of /sC/. Exposure to multiple variants may reinforce the analysis of [Ch] as a realization of /sC/ (see Section 3.4.1). Sevillian listeners never hear words produced only with metathesis, so they always have evidence to link [h] in [Ch] back to underlying /s/.

The ABX task differs from the previously mentioned experimental tasks because it does not access lexical representations, or give a plausible /s/ source. The tasks presents [Ch] sequences in nonce words and out of context. That the experiment uses nonce words may not be the issue, since the stress judgment task also used nonce words. When Sevillians hear a sequence like [laˈnatha]-[laˈnata]- [laˈnatha] in the ABX task, the nonce words have no lexical representation *and* there is no explicit possibility that [Ch] derives from /sC/. When given surface [Ch] without the possibility that [Ch] derives from /s/ (context that would be given either by lexical representations or multiple forms shown in the experiment), they do not map [Ch] to /sC/. For them, [h] in [Ch] is a surface realization of /s/, not /h/. If they are unable to map [h] to the phonological category /s/, then a comparison like [laˈnatha]-[laˈnata]- [laˈnatha] is comparing three repetitions of the same word and accuracy is expected to be low. In short, Sevillians can distinguish [Ch] from [C] in derived contexts only.

5.5.1.2 French listeners

French listeners had unexpectedly high accuracy on my ABX task. Recall that they were chosen specifically because French lacks a phonemic category /h/ and aspirated stops, so these listeners had little experience with the forms presented. In a similar study testing the perceptibility of [h], Mielke (2003) also used French listeners and found that they did indeed have low accuracy. A different take of categorization and perception predicted by the Perceptual Assimilation Model actually can explain these results.

One possible reason for French listeners' high accuracy could be that [h] was uncategorizable for them. The Perceptual Assimilation Model (Section 5.1.3) predicts that listeners' ability to discriminate sounds depends on how these sounds are mapped to native categories. If one sound maps onto a native category but the other is unable to be mapped—it is uncategorized—discrimination is also expected to be good because one sound assimilates to a native phoneme category, and the other is decidedly *not* part of that category. In terms of the current ABX task, French listeners may have performed well precisely because plain consonants (/ptk bdg mnl sf/) were assimilated to native categories, and [h] was not mapped to a phonemic category. The distinction between plain consonants [C] and consonants with accompanying [h] ([hC, Ch]) would have then been straightforward. However, this does not explain why French listeners had such high accuracy on linear order comparisons as well ([hC-Ch]). According to Best et al. (1988: 347), non-assimilated contrasts (uncategorizable), 'should be perceived in terms of their auditory (acoustic or nonspeech) or phonetic (phonologically neutral articulatory) characteristics.' While other listeners struggled to map (and thus discriminate) [h], French listeners may have only attended to acoustic information, leading to higher accuracy.

The unexpectedly good performance of listeners whose language lacks the relevant sounds is not actually unprecedented in AX/ABX categorization tasks. In an AX task testing the perception of [h] in HC and CH sequences, Clayton (2010) used Polish listeners as a group lacking both [hC] and [Ch] sequences, like my French listeners. Contrary to expectations, he found that Polish listeners were surprisingly good at perceiving [h] in HC in word-medial position. He hypothesizes that Polish listeners' high accuracy may have been because they assimilated [hT] to native [xT] clusters. Ejectives present another case of good discrimination with a non-native phoneme category. Gallagher (2012) tests English listeners on Quechua ejectives, and they are surprisingly good at distinguishing between roots with 1 vs. 0 ejectives (k'api/kap'i vs. kapi; approx. 78%). Given that English listeners lack experience with ejectives, these numbers are surprisingly high. The effect makes more sense, however, if ejectives are uncategorizable for English listeners, who may

rely on acoustics alone to distinguish uncategorizeable ejective stops from categorizeable plain stops.

5.5.2 Relevance to the typology of metathesis and preaspiration

The results from my ABX task do not support a perceptual optimization account of metathesis, or a perceptual inferiority account of the typological rarity of preaspiration.

Many metathesis processes involving laryngeals, specifically with voiceless stops (see Section 5.1.2.2), have been argued to improve perceptibility. My study does not test listeners of these languages specifically, but the HC/C contrast is more perceptually distinct than the CH/C contrast for all of my listener groups and in the majority of consonantal contexts. [h] is easier to perceive before a stop than after a stop. It is possible that languages with regular [hC] $\rightarrow$ [Ch] metathesis have a less phonetically salient coda [h] than languages that do not metathesize, and that metathesis might actually provide perceptual benefits for [h] in these languages. It is also possible that the explanation is more along articulatory lines.

In terms of pre- and postaspiration, my results are largely consistent with one of the few experimental studies on the perception of pre- and postaspiration. In an AX categorization task, Clayton (2010) also found that preaspirated stops are easier to discriminate than postaspirated stops (HC > CH), but position in the word (initial, medial, final) had a bigger effect on accuracy than aspiration type. He also found that native language affected accuracy. Gaelic listeners (who have native language experience with pre- and postaspiration), perceived medial HC as well as initial CH. English listeners, who have experience with aspirated stops but not [hC] sequences, were the only group that had more difficulty with HC than CH. He claims that this is likely not because [h] in [hC] sequences is 'intrinsic[ally] defect[ive]', but rather that only initial [Ch] sequences are L1-like for them. Polish listeners, who have neither [hC] nor [Ch] sequences performed surprisingly well on medial HC. Clayton concludes that preaspiration is not cross-linguistically rare because [h] in HC sequences is perceptually difficult, but rather because preaspiration on stops is rarely innovated. In his dissertation, he argues that preaspirated stops are rarely innovated for several

reasons, including the fact that the structures that give rise to them are rare (see Clayton 2010: 70 for further discussion). My perception results mirror his: perceiving [h] in HC sequences is not overly perceptually difficult, even for listeners whose languages do not have these sequences.

Furthermore, preaspiration may be quite salient in the languages where it occurs as a realization of aspirated stops. Some of these languages realize preaspiration with an additional oral constriction, which can be homorganic (or not) to the following stop (e.g. [hp ht hk] → [fp ɕt xk]; Silverman 2003). Clayton (2010: 116) points out that some languages have glottal and oralized forms in variation (e.g. Faroese, Central Swedish, Tarascan, Lewis dialect of Scottish Gaelic, Goajiro), while others use exclusively oralized variants (e.g. Scottish Gaelic (non-Lewis dialects), Bora, Fox). Silverman (2003) argues that this additional oral constriction enhances perceptibility.

Listeners of languages that have preaspiration are also highly accurate at perceiving it, so it must be salient. For example, Icelandic listeners have high accuracy at perceiving preaspiration, and can do so even when there is no voiceless glottal portion. They rely mostly on the breathy voiced transition between the preceding vowel and [h], showing that the coarticulation is crucial for perception (Ni Chasaide 1985: 367). Although my stimuli were naturally produced, there may have been less of this coarticulation that is important for perceiving preaspiration in languages that have it.

5.5.3 Limitations of the current study

My experiment uses the same set of stimuli for listeners of different language backgrounds, and these stimuli were recorded by a Turkish speaker. The benefits of this are that all listeners responded to exactly the same phonetic properties, and that no group had a native-listener advantage. However, the stimuli are non-native sounding for all of my listeners, which may introduce certain artifacts into the results.

Several characteristics of the stimuli could have resulted in higher accuracy for HC/C as compared to CH/C. First, [h] in my [hC] sequences may have been more salient than in natural speech. My stimuli were naturally produced in careful speech, and the segments were likely less

overlapped than in casual speech. [h] overlapping with an adjacent consonant (especially a nasal or fricative) may make [h] hard to perceive. Devoiced nasals are acoustically and perceptually similar to glottal fricatives (Herbert 1986: 245; Ohala 1993: 240).[11] Despite these considerations, there is no acoustic reason that [h] in my stimuli should have been overly salient in [hC] contexts. It was purely glottal, and it was short. Recall that [h] in all members of a word set was shortened to the length of the shortest [h] in that set (Section 5.3.2), in order to lower almost-ceiling accuracy.

English listeners' accuracy on CH/C was not as high as expected. It is possible that [h] in [Ch] sequences was too short to qualify as 'aspiration' on a stop. If the difference in aspiration in [C] and [Ch] was not large enough to cue the voiced/voiceless distinction, both forms may have been classified as allophones of a single category, leading to less accurate discrimination (e.g. Best and Strange 1992; Best and Tyler 2007). But this explanation does not hold. In the ABX stimuli, the duration of [h] in [Ch] sequences was around 70-80ms, and Lisker and Abramson (1964) reports VOTs for English voiceless stops between 58-80ms. Another possibility is that English listeners did not perceive [h] in CH sequences well because [h] occurred in the onset to an unstressed syllable (e.g. [laˈnatha]), and English aspirated stops occur only as the onset to stressed syllables. The [Ch] sequence in a phonotactically illegal environment may have lowered their accuracy, but recall that English listeners were one of the only groups (in addition to Arabic listeners) that perceived [h] equally well in CH and HC contexts. English listeners got *some* benefit from familiarity with aspirated stops, even if it was not as strong as expected.

Sevillians may have had similar accuracy on CH/C as other Spanish listeners due to similar surface-underlying form mappings. One objection may be that [h] in [Ch] sequences was not long enough for Sevillians to map to underlying /sC/, but this does not explain their behavior. The duration of [h] in the ABX stimuli was between 70-80ms, and in the fill-in-the-blank task Sevillians responded strongly to [h] in [Ch] that was between 65-95ms.

[11] Indeed, aspiration on stops sometimes has its diachronic origins in devoiced nasals. Clayton (2010) reports that some preaspirated stops derive from historical nasal-voiceless stop clusters, where the nasal underwent anticipatory devoicing before the voiceless stop, and was reinterpreted as a glottal [h]. For postaspirated stops, nasal devoicing (again, by assimilation) led to pre- and then postaspiration (Givón 1974: 110; e.g. Swahili, Hinnebusch 1975; Ikalanga (Bangu), Mathangwane 1996; Bantu languages more broadly, Downing and Hamann 2018).

5.6 Conclusion

This study tested two questions: that (a) laryngeal metathesis from [hC] → [Ch] is facilitated and driven by a need to optimize the perception of [h]; (b) in contrast, the perception of [h] is based on properties of listeners' native languages, and the perceptual properties tested do not universally favor metathesis. The ABX task tested listeners' accuracy on perceiving the presence and absence of [h], the location of [h], and the effect of the following consonant on [h] perceptibility. To investigate the influence of native language on perception, I tested listeners of different language backgrounds (Arabic, English, French, Mexican Spanish, Argentinian Spanish, and Sevillian Spanish), whose languages differ in the presence of aspirated stops, and of phonemic /h/ or allophonic [h].

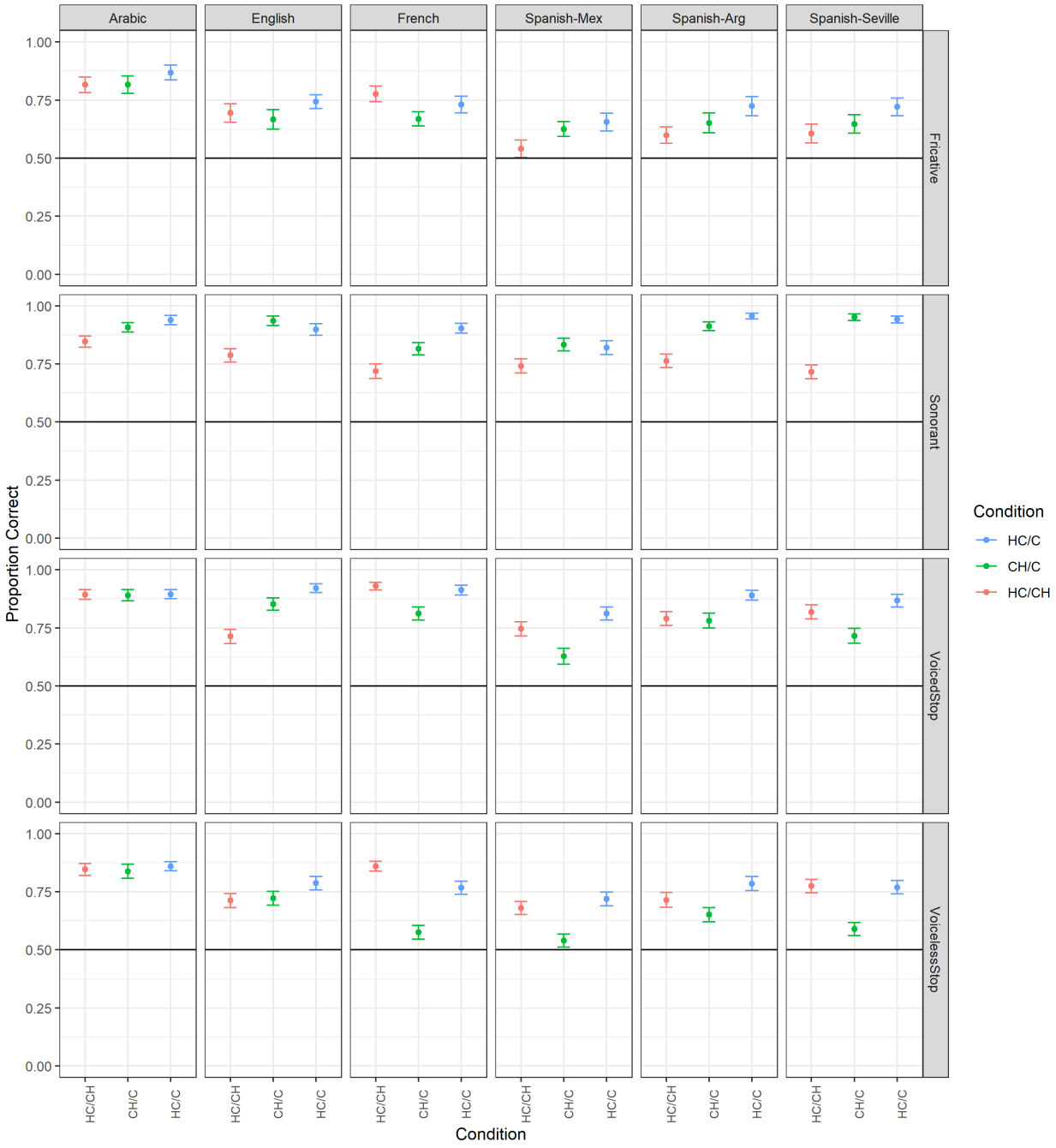
Results do not support the hypothesis that metathesis is driven by perceptual optimization. The results actually show the opposite: [h] is easier to perceive in HC sequences than in CH sequences for listeners of most language groups. For no group is [h] easier to perceive in CH sequences than in HC sequences. Furthermore, the location of [h] was not more confusable around voiceless stops than around other consonant types, so metathesis around voiceless stops cannot be due to the fact that linear order is particularly difficult (= the two forms are perceptually similar) in this context. Instead, the results largely support the influence of native language on perception. Native language had a large effect on accuracy. Arabic listeners had the highest accuracy, in most conditions and consonantal contexts. This was expected because Arabic has phonemic /h/ and allows it in a many phonological contexts. English listeners had higher accuracy on CH/C comparisons than other groups, likely due to the fact that English has a category of aspirated stops. French listeners performed surprisingly well, given that French lacks [h] as a feature or segment. Their high accuracy may reflect a different strategy: because they had no phonological category to map [h] to, it was acoustically salient to them. Listeners of all Spanish dialects patterned together, likely because their dialects have the same phonemic categories. That surface forms in each dialect differ appears not to matter. In particular, Sevillian Spanish listeners behaved like listeners of other dialects—who lack familiarity with [Ch] sequences—suggesting that they were unable to apply their knowledge of derived [Ch] (←) to a new context, where [Ch] could not have derived from

/sC/. Like other Spanish-speaking listeners, Sevillians do not have phonemic aspirated stops, and were thus unable to map [Ch] to a contrastive category.

In sum, laryngeal metathesis is not perceptually optimizing, for Sevillian Spanish or listeners of other language backgrounds. For Sevillian, a more likely explanation is articulatory or structural; I elaborate further on this in Chapter 6, Section 6.2. The results also do not support arguments that preaspiration is rare because it is perceptually difficult. In contrast, the current study suggests that [h] in [hC] sequences is quite perceptible, and motivations for both metathesis and the rarity of preaspiration should be sought elsewhere.

5.7 Extra ABX figures

Figure 5.7: Effect of Condition by Language and Consonant Type



5.8 Extra ABX models

Table 5.16: Model: Condition (Baseline: HC/C) x Language (Baseline: Arabic)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	2.340	0.205	11.396	0.000
Language English	-0.501	0.217	-2.315	0.021
Language French	-0.595	0.213	-2.797	0.005
Language Spanish-Mex	-1.146	0.207	-5.534	0.000
Language Spanish-Arg	-0.465	0.214	-2.171	0.030
Language Spanish-Seville	-0.626	0.218	-2.877	0.004
Condition HC/CH	-0.418	0.152	-2.744	0.006
Condition CH/C	-0.257	0.156	-1.653	0.098
Language English:Condition HC/CH	-0.356	0.200	-1.777	0.076
Language French:Condition HC/CH	0.329	0.202	1.625	0.104
Language Spanish-Mex:Condition HC/CH	0.064	0.189	0.338	0.735
Language Spanish-Arg:Condition HC/CH	-0.380	0.199	-1.915	0.055
Language Spanish-Seville:Condition HC/CH	-0.198	0.201	-0.986	0.324
Language English:Condition CH/C	-0.040	0.207	-0.192	0.847
Language French:Condition CH/C	-0.493	0.199	-2.482	0.013
Language Spanish-Mex:Condition CH/C	-0.241	0.191	-1.261	0.207
Language Spanish-Arg:Condition CH/C	-0.386	0.202	-1.910	0.056
Language Spanish-Seville:Condition CH/C	-0.395	0.203	-1.943	0.052

Table 5.17: Model: Arabic, Condition (Baseline: HC/C) x Consonant Type (Baseline: VoicedStop)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	2.350	0.280	8.404	0.000
Condition HC/CH	-0.187	0.304	-0.614	0.539
Condition CH/C	-0.049	0.310	-0.158	0.875
ConsType VoicelessStop	-0.352	0.296	-1.189	0.234
ConsType Fricative	-0.272	0.332	-0.818	0.413
ConsType Sonorant	0.620	0.357	1.738	0.082
Condition HC/CH:ConsType VoicelessStop	0.073	0.409	0.178	0.859
Condition CH/C:ConsType VoicelessStop	-0.137	0.412	-0.332	0.740
Condition HC/CH:ConsType Fricative	-0.247	0.448	-0.552	0.581
Condition CH/C:ConsType Fricative	-0.385	0.453	-0.850	0.395
Condition HC/CH:ConsType Sonorant	-0.900	0.455	-1.979	0.048
Condition CH/C:ConsType Sonorant	-0.414	0.479	-0.863	0.388

Table 5.18: Model: English, Condition (Baseline: HC/C) x Consonant Type (Baseline: VoicedStop)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	2.576	0.315	8.171	0.000
Condition HC/CH	-1.601	0.296	-5.401	0.000
Condition CH/C	-0.728	0.319	-2.284	0.022
ConsType VoicelessStop	-1.183	0.363	-3.262	0.001
ConsType Fricative	-1.429	0.388	-3.679	0.000
ConsType Sonorant	-0.293	0.392	-0.747	0.455
Condition HC/CH:ConsType VoicelessStop	1.114	0.373	2.984	0.003
Condition CH/C:ConsType VoicelessStop	0.359	0.393	0.915	0.360
Condition HC/CH:ConsType Fricative	1.345	0.401	3.357	0.001
Condition CH/C:ConsType Fricative	0.336	0.416	0.808	0.419
Condition HC/CH:ConsType Sonorant	0.706	0.409	1.726	0.084
Condition CH/C:ConsType Sonorant	1.228	0.478	2.568	0.010

Table 5.19: Model: French, Condition (Baseline: HC/C) x Consonant Type (Baseline: VoicedStop)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	2.387	0.251	9.496	0.000
Condition HC/CH	0.244	0.349	0.699	0.485
Condition CH/C	-0.894	0.290	-3.086	0.002
ConsType VoicelessStop	-1.164	0.292	-3.992	0.000
ConsType Fricative	-1.366	0.309	-4.426	0.000
ConsType Sonorant	-0.107	0.332	-0.321	0.748
Condition HC/CH:ConsType VoicelessStop	0.384	0.429	0.895	0.371
Condition CH/C:ConsType VoicelessStop	-0.058	0.357	-0.162	0.872
Condition HC/CH:ConsType Fricative	0.009	0.441	0.021	0.983
Condition CH/C:ConsType Fricative	0.514	0.384	1.339	0.181
Condition HC/CH:ConsType Sonorant	-1.560	0.441	-3.535	0.000
Condition CH/C:ConsType Sonorant	0.136	0.405	0.335	0.738

Table 5.20: Model: Spanish dialects, Language (Baseline: Spanish-Mexico) x Condition (Baseline: HC/C) x Consonant Type (Baseline: VoicedStop)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	1.509	0.220	6.869	0.000
Spanish-Arg	0.686	0.301	2.281	0.023
Spanish-Seville	0.427	0.298	1.434	0.152
ConsType VoicelessStop	-0.605	0.271	-2.236	0.025
ConsType Fricative	-0.922	0.294	-3.133	0.002
ConsType Sonorant	0.052	0.286	0.180	0.857
Cond. HC/CH	-0.394	0.230	-1.715	0.086
Cond. CH/C	-0.968	0.220	-4.391	0.000
Spanish-Arg:ConsType VoicelessStop	-0.228	0.351	-0.650	0.516
Spanish-Seville:ConsType VoicelessStop	-0.092	0.349	-0.263	0.793
Spanish-Arg:ConsType Fricative	-0.255	0.372	-0.687	0.492
Spanish-Seville:ConsType Fricative	-0.037	0.372	-0.100	0.920
Spanish-Arg:ConsType Sonorant	0.940	0.456	2.060	0.039
Spanish-Seville:ConsType Sonorant	0.840	0.436	1.927	0.054
Spanish-Arg:Cond. HC/CH	-0.403	0.354	-1.138	0.255
Spanish-Seville:Cond. HC/CH	0.014	0.360	0.038	0.969
Spanish-Arg:Cond. CH/C	0.116	0.347	0.335	0.738
Spanish-Seville:Cond. CH/C	-0.011	0.341	-0.034	0.973
ConsType VoicelessStop:Cond. HC/CH	0.267	0.308	0.866	0.386
ConsType Fricative:Cond. HC/CH	-0.083	0.331	-0.251	0.802
ConsType Sonorant:Cond. HC/CH	-0.081	0.326	-0.250	0.803
ConsType VoicelessStop:Cond. CH/C	0.192	0.297	0.647	0.518
ConsType Fricative:Cond. CH/C	0.910	0.327	2.784	0.005
ConsType Sonorant:Cond. CH/C	1.062	0.332	3.194	0.001

Table 5.21: Model: Spanish dialects (continued)

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	1.509	0.220	6.869	0.000
Span.-Seville:ConsType VoicelessStop: Cond. HC/CH	0.142	0.477	0.297	0.767
Span.-Arg:ConsType Fricative: Cond. HC/CH	0.289	0.495	0.584	0.559
Span.-Seville:ConsType Fricative: Cond. HC/CH	-0.144	0.504	-0.285	0.776
Span.-Arg:ConsType Sonorant: Cond. HC/CH	-1.075	0.554	-1.939	0.052
Span.-Seville:ConsType Sonorant: Cond. HC/CH	-1.414	0.543	-2.605	0.009
Span.-Arg:ConsType VoicelessStop: Cond. CH/C	-0.091	0.455	-0.199	0.842
Span.-Seville:ConsType VoicelessStop: Cond. CH/C	-0.124	0.452	-0.275	0.783
Span.-Arg:ConsType Fricative: Cond. CH/C	-0.412	0.493	-0.836	0.403
Span.-Seville:ConsType Fricative: Cond. CH/C	-0.281	0.494	-0.568	0.570
Span.-Arg:ConsType Sonorant: Cond. CH/C	-0.960	0.584	-1.644	0.100
Span.-Seville:ConsType Sonorant: Cond. CH/C	0.112	0.608	0.184	0.854

Table 5.22: Model: Voiceless stops

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	1.844	0.215	8.557	0.000
Language English	-0.514	0.266	-1.933	0.053
Language French	-0.631	0.260	-2.431	0.015
Language Spanish-Mex	-0.953	0.253	-3.766	0.000
Language Spanish-Arg	-0.524	0.262	-1.998	0.046
Language Spanish-Seville	-0.618	0.266	-2.323	0.020
Condition HC/CH	-0.106	0.266	-0.400	0.689
Condition CH/C	-0.173	0.263	-0.659	0.510
Language English:Condition HC/CH	-0.365	0.347	-1.051	0.293
Language French:Condition HC/CH	0.730	0.363	2.012	0.044
Language Spanish-Mex:Condition HC/CH	-0.019	0.335	-0.058	0.954
Language Spanish-Arg:Condition HC/CH	-0.276	0.344	-0.801	0.423
Language Spanish-Seville:Condition HC/CH	0.134	0.356	0.378	0.706
Language English:Condition CH/C	-0.184	0.347	-0.530	0.596
Language French:Condition CH/C	-0.770	0.335	-2.303	0.021
Language Spanish-Mex:Condition CH/C	-0.593	0.329	-1.801	0.072
Language Spanish-Arg:Condition CH/C	-0.558	0.339	-1.647	0.099
Language Spanish-Seville:Condition CH/C	-0.730	0.342	-2.133	0.033

CHAPTER 6

Discussion

6.0 Summary of results

The experiments presented in this dissertation converge on several main results: Sevillian stop-h are still represented as /sC/ clusters, metathesis occurs by gestural overlap, stress operates on non-surface representations, and metathesis is not motivated by perceptual optimization. While these findings are mostly specific to Sevillian Spanish, the extent to which they generalize is an important topic for future research.

The production study (Chapter 2) provided evidence for gestural overlap in /s/-voiceless and /s/-voiced stop clusters, through a cluster of phonetic features. Metathesis appears to re-time the gestures associated with the stop in relation to those from /s/, so that they overlap. For /s/-voiceless stop clusters, metathesis results in a stop closure followed by a long stop release ([Ch]). These stop closures resist intervocalic voicing more than underlyingly intervocalic stops, suggesting that the [h] gesture 'protects' the closure. Furthermore, the presence of realizations where [h] is split across a voiceless stop represents an intermediate stage of metathesis. Realignment has partially—but not fully—occurred, leaving parts of the [h] gesture on both sides of the stop. For

/s/-voiced stop clusters, metathesis results in increased constriction degree, increased noise during the constriction, and partial voicelessness. These cues are mostly explainable by overlap of gestures from /s/ and the stop. It is not possible to tell if the 'metathesis' in /s/-voiced stop clusters is the same as in /s/-voiceless stop clusters because voiced stops spirantize in this context. The process might be similar, but there is no stop release past which [h] could extend, so the physical manifestation of the same process may be different. /s/-sonorant clusters show no evidence of overlap or metathesis, and [h] often deletes. I suggested that overlap may be prevented by gestural (and thus featural) incompatibility between [h] and a sonorant: the spread glottis gesture required to produce [h] cannot be realized at the same time as the gesture producing voicing in sonorants. It also may be that noise superimposed on a sonorant is difficult to perceive, so overlap would be perceptually equivalent to deleting [h]. Finally, there is also evidence that metathesized /s/-voiceless stop clusters are advanced in the change, and may be integrating into the phonological system by participating in other phonological processes in the language.

The fill-in-the-blank experiment (Chapter 3) tested the representational status of Sevillian [Ch] sequences by probing the mapping from surface forms to underlying representations. Results showed that Sevillian listeners mapped surface [Ch] sequences to underlying /sC/ clusters, treating [Ch] as a derived sequence where /s/ and /C/ belong to different morphemes. In contrast to Sevillian listeners, Mexican and Argentinian listeners did not 'undo' metathesis of [Ch] to get back to the intended /sC/ representation. Assuming that all listeners have the same underlying representations for the stimuli, I argued that their differences in behavior arise from differences in how their phonological grammars map /sC/ to surface forms, and thus how they map [Ch] to underlying forms.

Building on the fill-in-the-blank task, the stress judgment task (Chapter 4) tested how [Ch] sequences interact with the broader phonological system. Results showed that, for the purposes of stress, Sevillian listeners treated surface [Ch] sequences as underlying /sC/ sequences. They evaluated stress on an intermediate representation where /s/ is still present in the coda of the preceding syllable (CV.<u>CVS</u>.CV, LHL), even though the surface form has all open syllables (CV.<u>CV</u>.CHV,

LLL). The mismatch between surface forms and forms on which stress is evaluated highlights the role of the phonological grammar. The grammar evaluates stress on a form that is not surface-visible. In doing this perception task, Sevillian listeners first undid metathesis to get back to a representation where /s/ is in the preceding syllable (its original location), and then evaluated stress. In production, then, the phonological grammar assigns stress before metathesis. Again, the phonological grammar is crucial in explaining this result: it governs the order in which metathesis and stress interact.

Finally, the ABX task (Chapter 5) tested potential perceptual motivations behind metathesis. Native speakers of Arabic, English, French, Mexican Spanish, Argentinian Spanish, and Sevillian Spanish were asked to discriminate pairs of words differing in the presence and linear order of [h]. The results provided no evidence that [hC] → [Ch] metathesis is perceptually optimizing. For all language groups and around almost all types of consonants, [h] was easier to perceive before a consonant (*lanahta*) than after a consonant (*lanatha*). Additionally, the location of [h] was *not* more difficult to perceive around voiceless stops—the environment where metathesis occurs—than around other types of consonants. These two findings provide no support for the claim that synchronic, productive metathesis is either perceptually optimizing or similar to the original form, in contrast to the hypotheses put forward in Steriade (2001). The findings also fail to support the argument (most clearly made by Silverman 2003) that preaspiration is rare because it is difficult to perceive.

The results of the ABX task did not indicate that there are universal tendencies of [h] perception that would motivate metathesis; instead, perception was driven by native language background and phonemic categories. Listeners' accuracy perceiving [h] is largely explainable by the presence or absence of relevant phonemic categories in their native language, to which they could map the stimuli. Strikingly, all three groups of Spanish speakers behaved the same. This is surprising because, although the dialects share phonemic categories, their surface forms are drastically different—specifically, Sevillians have experience with surface [Ch] sequences. These results differ from the fill-in-the-blank task, where Sevillians' behavior diverged from that of other Spanish

speakers. I hypothesize that this is due to a derivedness effect. In the fill-in-the-blank task, Sevillians were able to 'undo' metathesis because there was an explicit possibility that [h] was derived from an /s/ on the preceding verb. In contrast, the ABX task presented [Ch] sequences in nonce words, in a context where they did not appear to be derived. Outside a derived context, Sevillian listeners did not perceive [Ch] sequences. In sum, in this experiment, listeners' perception was governed by a combination of one seemingly universal tendency (that [h] is easier to hear before a consonant than after a consonant), and their native language phonological systems. Perception was governed largely by how listeners' phonological systems mapped mapped the stimuli to phonemic categories in their own languages.

In the remainder of this chapter, I broaden the discussion of the results in relation to other areas. First, I discuss how my results fit into previous work on metathesis (Section 6.1). While most theories treat metathesis as an atomic operation that fully reorders segments, other proposals suggest that at least some types of metathesis can occur via overlap, and Sevillian supports this analysis. Next, given that the results of the ABX task do not support a perceptual explanation for metathesis, I discuss possible articulatory factors that could lead to metathesis, drawing from previous proposals on Sevillian. I also draw on proposals about how the internal structure of stops provides multiple options for laryngeal timing (Section 6.2). I also discuss Sevillian metathesis and phonologization (Section 6.3). Based on criteria used for other cases of phonologization and a brief investigation of how aspirated stops arose in other languages, I speculate about what conditions might be necessary for phonologization, and why Sevillian [Ch] sequences may not be phonologizing. Finally, I compare stress-metathesis interactions to another, better-studied interaction: that between stress and epenthesis (Section 6.4.1). I suggest that the two types interactions may differ in ways that shed light on how and where in the phonological grammar metathesis occurs. This section is speculative: I aim to is to put forth hypotheses that motivate future research on the role of metathesis in the phonological grammar.

6.1 Metathesis and gestural overlap

While most analyses treat metathesis as transposition of two segments, others treat it as resulting from gestural overlap between two segments. Furthermore, different types of metathesis might operate differently. For example, Mooney (To appear) argues that phonological metathesis occurs by gestural overlap, while true segmental metathesis is a morphological operation that is morpheme-specific.

Grammont (1933) is one of the earliest studies to treat metathesis as gestural overlap. While he allows for some metatheses to occur by segmental transposition, he describes other metatheses as occurring when one gesture passes through another, rather than around it (Grammont 1933: 244-9). He illustrates with the case of *w* and *r* in Indo-European, where *wr* became *rw/ru* in some contexts. He argues that there is first a stage of assimilation, where [r] assimilates to [w] in raising the dorsum and lip rounding. Then, [w] 'springs up again' on the other side (translation from Hall 2003: 4). He also gives an example from Zend to illustrate the intermediate stage, where [r] is submerged in [w], and parts of [w] appear on both sides of [r]: *ᵘ̲rv̲āta* 'dogma' (cf. Sanskrit *v̲ratám* 'precept'; Grammont 1933: 244). Hall (2003: 4) interprets Grammont's proposal in gestural terms, illustrating his assumption about a diachronic state in which the [r] gesture is completely overlapped by the [w] gesture, which extends beyond both sides of [r].

More recent analyses of metathesis also imply that both gestural overlap and perceptual factors play a role. Blevins and Garrett (1998) propose that perceptual metathesis occurs 'when features extending across a CV or VC domain, or perceived as extending across such a domain, are reinterpreted as originating in nonhistorical positions' (Blevins and Garrett 1998: 510-11). They illustrate with laryngeal metathesis in Cayuga (/Vh/ → [hV]), which occurs in specific syllables. They posit several stages to this change: /CVh/ → [CVV̥] → [CV̥]. First, a vowel is followed by [h]. They then propose that this vowel shortened, resulting in a mostly devoiced vowel. At this intermediate step, features of [h] are realized on the vowel, setting the stage for the final step, where [CV̥] could be reinterpreted as /ChV/, as opposed to the original /CVh/. While Cayuga has not gone to the third step, they point out that related languages have (e.g. Cherokee).

A crucial part of their argument for Cayuga is that metathesis can occur *because* laryngeal feature is realized across a domain (the vowel, in this case), which provides an opportunity for the feature to be reinterpreted in a non-historical position. While they do not frame it explicitly in terms of gestural overlap, the intermediate stage of vowel devoicing implies that the laryngeal gesture of [h] overlaps with the vocalic one. However, as I will discuss later in this section, results from my ABX task provide no evidence that the linear order of [h] is highly confusable, which undermines the step in Blevins and Garrett's (1998) proposed chain of events where the location of [h] is reinterpreted in a non-historical position.

Blevins and Garrett (1998: 527) also propose compensatory metathesis, which is prosodically conditioned and leads to a high degree of coarticulation in VCV sequences. A domain-final vowel shortens, and its gesture overlaps the consonant to align more with the preceding stressed vowel. The result is reduction or loss of the domain-final vowel, and a stressed, diphthongized vowel (e.g. [tíko] → [tíok] 'flesh'). These metatheses are common in Austronesian languages like Rotuman and Kwara'ae. The high degree of coarticulation that gives rise to compensatory metathesis implies gestural overlap.

A crucial point of Blevins and colleagues is that perceptual metathesis is limited to certain types of segments whose cues are realized over a long domain, providing an opportunity for reinterpretation. Specifically, Blevins and Garrett (1998) propose that these segments include liquids, laryngeals, pharyngeals, and glides/vowels; this is similar to the argument made by Ohala (1981b: 193) for dissimilation. Dissimilation and metathesis are similar in that the perceptual motivations are argued to occur in similar ways. For metathesis, Blevins and colleagues argue that spread-out cues create ambiguity as to the source of the phonetic cue in the acoustic signal, and listeners can reinterpret it in a non-historical position. For dissimilation, Ohala (1981b) and Ohala (1993) propose that dissimilation occurs when a listener hears a word with two of the same consonant, and interprets one of them to be a phonetic effect of the other. Under this account, spread-out cues make it possible for listeners to hear one segment as the phonetic effect of another, failing to attribute the coarticulatory effect appropriately and positing only one instance of the segment instead

of two. Metathesis and dissimilation are thus similar in that they both involve misinterpretation of the source of a phonetic effect.

Most of the claims about metathesis are not supported by evidence from perceptual experiments. Specifically, there is little experimental evidence (a) that the cues to these segments are realized over a long domain, or that there is phonetic evidence of overlap (production), or (b) listeners misperceive the location of segments due to these spread-out cues. In terms of phonetics, sequences that undergo metathesis may show phonetic evidence of gestural overlap, as these theories hypothesize. Indeed, my own phonetic evidence from Sevillian—presented in Chapter 2—suggests that metathesis occurs via gestural overlap, like Grammont's [rw] sequences and laryngeal metathesis in Cayuga. There are even intermediate forms, [hCh], where the laryngeal gesture completely overlaps a voiceless stop, extending beyond it on both sides (see (11) in Chapter 2). Other phonetic evidence for metathesis-via-overlap comes from Uab Meto, where metathesis can strand a portion of the metathesized vowel in its original location (Mooney To appear; Gilbert and Mooney To appear). For example, in (79b), metathesis of /u/ across /n/ results in a diphthong [a͡ʊ], but part of /u/ remains on the right side of /n/. See the spectrogram corresponding to this word in Gilbert and Mooney (To appear: 6).

(79) Uab Meto metathesis

a.	/manus/	[manus]	'betel vine'
b.	/manus-es/	[ma͡ʊnᵘses]	'a betel vine'

Phonetic evidence like this suggests that at least some kinds of metathesis occur through a mechanism of gestural realignment and overlap. To determine how common this is, future work needs to examine the phonetic properties of metathesized sequences at a broad, cross-linguistic level.

Some studies on dissimilation also finds evidence that some of the segments hypothesized to have spread-out cues actually do have long-distance coarticulatory effects on other segments. For example, for Standard Southern British English, Tunley (1999) found that /r/ has measurable coarticulatory effects on the formants of vowels multiple syllables away. Other studies have found similar results for varieties of British English (see Hall et al. 2017 for overview). However, Hall

et al. (2017) find that /r/ and /l/ do not have meaningful long-range coarticulatory effects in American English. This suggests that, if long-range coarticulatory effects are argued to be the source of ambiguity in metathesis, that they must be established experimentally for the language or dialect in question.

In terms of perception, there is little evidence that listeners misperceive the source of segments due to elongated cues, in a way that could lead to metathesis or dissimilation. Most of the experimental work on this topic comes from dissimilation. Several perception studies have failed to find evidence that listeners misperceive the source of a phonetic cue, leading to dissimilation (e.g. Abrego-Collier 2013; Harrington et al. 2016). However, Hall et al. (2019) point out that these studies test patterns that do not exist in the listeners' native languages. In their study, they do find evidence of dissimilation: American English listeners are less likely to write *r* if it is followed by *r* later in the word. They interpret these results as suggesting that listeners perceive the first *r* as a phonetic effect deriving from the second, not a segment in and of itself. Hall et al. (2020) find similar results for /l/ but not for /n/ in American English, which they argue mirrors actual dissimilation patterns: /l/ dissimilation is occasional, and /n/ dissimilation is rare.

Results from my ABX task in Chapter 5 suggest that Sevillian metathesis is not perceptually motivated in the Ohala sense of misperceiving the source location of a phonetic effect: Sevillian listeners (and other listeners) have relatively high accuracy at hearing the linear location of [h]. They do not seem to be misattributing it to the wrong location, even—or especially—around voiceless stops. Despite the fact that the motivation for metathesis does not appear to be misperception, it is appropriate to characterize metathesis as involving one gesture sliding around another.[1]

This collection of results about dissimilation and metathesis suggests that proposals arguing that spread out cues can lead listeners to misattribute the source of those cues need to be substantiated by language-specific experimental studies. For these accounts to hold, it must be shown that

[1]It is possible that my stimuli do not have long-distance coarticulatory effects, and that this is the reason for listeners' high accuracy in determining the location of [h]. Many proposals arguing for metathesis based on perceptual grounds do not consider whether these effects are actually present or not.

the specific language actually implements the cue over a long domain, and that listeners are prone to misattributing the cue.

Instead of attributing a sound to a non-historical position for perceptual reasons, there could be other reasons that metathesis is most common with certain segments (articulatory compatibility). One reason may be that some gestures are more able to 'slide' over others, resulting in a change that looks like metathesis. Laryngeals, for example, have no supralaryngeal constriction and do not interfere with the production of consonants with constrictions at other places in the vocal tract. A laryngeal gesture can overlap other consonantal and vocalic gestures because the demands of the articulations are compatible.

6.1.1 Directions for future work on metathesis as overlap

The topic of gestural overlap in metathesis would benefit from future research in several areas. First, there is little work on the phonetics of metathesized sequences. Metathesis patterns that occur by overlap should show some evidence of this in the acoustics. Additionally, given the popularity of hypotheses that spread-out phonetic cues are necessary for metathesis, the presence of these cues needs to be established. In terms of perception, the misattribution hypothesis—that metathesis occurs due to misattribution of phonetic cues to a non-historical position—needs to be tested with different kinds of segments and in different languages. Finally, the articulatory compatibility hypothesis makes testable predictions about what kinds of segments should be able to undergo metathesis, and what segments might block metathesis because the articulations involved are incompatible.

6.2 An articulatory hypothesis for Sevillian metathesis

The results of my ABX experiment do not support a perceptual story for metathesis in Sevillian, or language-universal tendencies in perception that could give rise to it. In this section, I review the explanations for metathesis based on articulatory timing preferences that have been put forward

in the literature (Section 6.2.1), and explore how articulatory binding (Section 6.2.2) and aperture theory (Section 6.2.3) accounts could be applied to Sevillian metathesis. I leave developing formal accounts along these lines to future work.

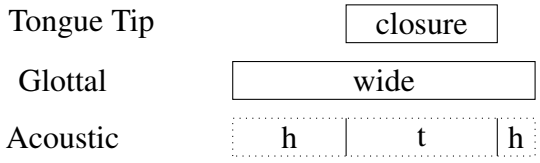
6.2.1 Articulatory timing pressures and Sevillian metathesis

Torreira (2006) and Parrell (2012) argue that Sevillian metathesis results from the pressure for gestures to start in-phase as opposed to anti-phase. In-phase coordination is when gestures begin simultaneously; anti-phase coordination is when the gestures are produced sequentially (e.g. Browman and Goldstein 1986). The two types of coordination show differences in stability across a range of human motor skills tasks. For example, Kelso (1984) found that peoples' hand movements are less stable when coordinated in an anti-phase relationship than when coordinated in an in-phase relationship. Furthermore, he found that people tend to switch from anti-phase to in-phase coordination when performing the task quickly. Movements that are coordinated in-phase are stable at both slow and fast rates of production. In speech, similar results have been found. When forced to increase speech rate, participants switch from VC syllables (anti-phase coordination) to CV (in-phase coordination), suggesting that the anti-phase coordination is less stable in speech as well (e.g. Tuller and Kelso 1990).

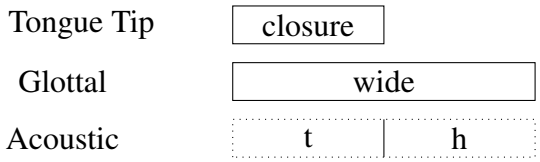
Parrell (2012) hypothesizes that Sevillian metathesis can derive from the generalized preference for in-phase gestural coordination. I repeat the gestural diagram from (11) in Chapter 2 here for convenience (80):

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

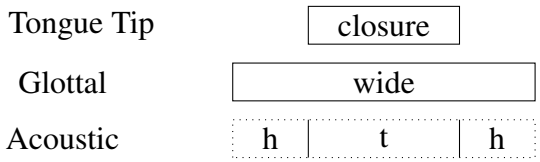
a. Sample gestural score for [ht] sequence



b. Sample gestural score for [th] sequence



c. Sample gestural score for [hth] sequence



In Parrell's (2012) experiment, Western Andalusian participants (from Seville and Cadiz) produced target words at increasingly fast speech rates. He found that most participants switched from anti-phase productions ([hC]) to in-phase productions ([Ch]) when they were forced to increase speech rate, supporting the hypothesis that timing preferences drive metathesis. However, some participants produced in-phase coordination [Ch], regardless of speech rate. This suggests that additional factors must be at play. Whereas the original pressure for metathesis may have been to achieve more stable coordination, acting as an in-the-moment constraint during production, the in-phase coordination has stabilized for these speakers. This timing is clearly a target, but does not mean that [Ch] is necessarily an aspirated stop.

One challenge to this account is that it is not clear that stop-h sequences are actually coordinated in-phase, in the sense that their onsets are synchronized. Gafos (2002), following Browman and Goldstein (2001), proposes that consonants in clusters have competing coordination relationships. Under this theory, each onset consonant of the cluster is coordinated with the following vowel, but they are also coordinated with each other. The consonant-consonant phasing must be stronger than the consonant-vowel phasing to ensure recoverability of both consonants. If the consonant-vowel phasing were stronger, each consonant would be coordinated with the vowel and the consonants would end up completely overlapped. Instead, the consonant-consonant phasing ensures that the consonantal gestures coordinate with each other so that they are offset. These competing coordination relationships result in the c-center effect, whereby the midpoint of the en-

tire cluster is in a stable timing relationship with the vowel (e.g. Browman and Goldstein 1988). Some languages do not show the c-center effect, and instead time only the final consonant of the onset sequence with the following vowel (e.g. Moroccan Arabic, Shaw et al. 2009; Jazani Arabic, Durvasula et al. 2021; Italian, Hermes et al. 2013). In either case, though, the onset consonants are not necessarily coordinated in-phase with each other, unlike Parrell's (2012) proposal for Seville.

Determining if gestural coordination preferences drive Sevillian metathesis would benefit from articulatory studies on these sequences. How are they coordinated? How stable is this coordination, and how does it compare to other clusters and sequences in the language?

6.2.2 An articulatory binding approach

Another possible explanation for metathesis in Sevillian involves articulatory binding: an affinity between glottal articulations and the structure and articulation of stops (Kingston 1990). This explanation could address why metathesis appears to apply fully in /sp st sk/ clusters, but not others. Kingston (1990) posits that stops and continuants are structurally different, with different components that result in distinct possibilities for coordination with glottal articulations.

Kingston (1990) points out that, cross-linguistically, stops are more likely to contrast for glottal articulations than fricatives or sonorants, and that glottal articulations on stops are more frequently realized at the release than at the beginning.

He argues that stops are more likely to contrast for glottal articulations than continuants or sonorants because of the aerodynamics involved in stop production. The quality of the glottal release is determined by factors like the size of the glottal aperture and the degree of constriction. He states that,

> 'The larger the glottal aperture and the lower the fold tension, the more glottal resistance will be reduced, the more air will flow through the glottis, and the more rapidly air pressure will rise in the oral cavity behind the obstruent articulation. This aerodynamic interaction between glottal and supraglottal articulations will be most dramatic in stops since the downstream obstruction of air flow is complete. The amount that

> intraoral air pressure is elevated behind the stop closure together with the size of the glottal aperture determine the acoustic character of the explosive burst of noise that occurs when the stop is released. Glottal aperture affects the acoustics of the burst both by the aerodynamic means just described and by determining what acoustic coupling there is between supra- and sub-glottal cavities. The acoustic character of the burst therefore at once depends on and cues the state of the glottis.' (Kingston 1990: 408)

In short, his argument is that the aerodynamics involved in stop production make stops good hosts to contrasts involving glottal features. This approach is indirectly perceptual, since his assumption seems to be that stops can host contrasts more easily because the glottal articulation is most perceptible on them, but it is also aerodynamic and articulatory.

Kingston (1990) also proposes that the structure of stops is responsible for the asymmetry between where glottal articulations tend to occur. Stops have a closure and a release, and glottal articulations they tend to occur much more frequently on stop releases than at their onsets. He argues that this is because 'it is the acoustic character of the burst which the glottal articulation is intended to modify' (Kingston 1990: 425).

An additional point of Kingston's argument makes it clear that his point is in part structural and articulatory. Kingston argues that stops and continuants differ in structure, and thus in how glottal articulations align to them. His proposal goes as follows. Stops constrain the timing of the glottal articulation (aligned to the release) more than continuants because stops have releases, and are thus asymmetric. Glottal articulations are tightly 'bound' in stops because, during the stop portion, a glottal articulation cannot be realized because the closure obstruents airflow (Kingston 1990: 409). The glottal cannot be realized during the stop, and must be realized on one side of it—most frequently, the release.

In contrast, he argues that continuants are more symmetric, with almost mirror-image onsets and offsets, and glottal gestures can be realized simultaneously because airflow is not obstructed. Glottal abduction 'is simply intended to produce a sufficient and continuous flow of air', so there is no clear preference for binding site (Kingston 1990: 424).

While part of binding theory is perceptual, it is largely structural and articulatory as well. Only stops have a release burst that a glottal gesture can bind to, and the complete closure means that the glottal cannot be realized simultaneously with the closure. Continuants lack a release burst, and, articulatorily, allow glottals to be realized simultaneously because airflow is continuous.

Applied to Sevillian metathesis, this account could explain the difference in phonetic outcome for Sevillian /s/-voiceless stop sequences vs. /s/-voiced stop sequences. In /s/-voiceless stop sequences, there is a stop release for [h] to bind to, and [h] is realized at the release. However, this account does not seem compatible articulatory descriptions (or motivations) of the change, in which the glottal gesture aligns with the onset of the stop and simply happens to extend past its release (Section 6.2). Determining whether the gestures are aligning in-phase vs. whether [h] is binding to the release would require articulatory research.[2] For /s/-voiced stop sequences though, the binding account predicts the attested outcome. Recall that voiced stops are spirantized in most phonological environments. In comparison to those stops, /s/-voiced stop sequences are realized with more frication and higher constriction degree, although they are still continuants that lack releases. The binding account correctly predicts that, because these are continuants, the laryngeal is realized during the constriction rather than after it.

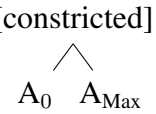
6.2.3 Aperture theory

An account related to binding theory would be one framed in Aperture Theory (Steriade 1993; Steriade 1994). In this theory, segments differ in number and type of aperture position nodes. Released stops have a closure node (A_0) followed by a release node (A_{Max}) and continuants have a single position node (A_{Max}). Features can associate with the nodes. Steriade (1994) develops an account of laryngeal timing within aperture theory, and my discussion here largely follows hers. She argues that, because stops have two aperture positions, laryngeal features can associate to one or both of them, giving rise to features realized during, before, or after the closure. She

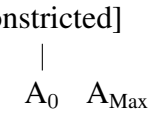
[2]Note also that if [h] is binding to the release instead of aligning to the onset of the voiceless stop, then Parrell's (2012) hypothesized motivation for metathesis—the pressure for in-phase coordination between the stop and [h] gestures—no longer applies.

illustrates with glottalization: the structure of the stop provides multiple docking options, resulting in different the different surface possibilities in (81)-(83) (examples adapted from Steriade 1994: 207).

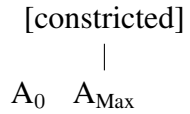
(81) Fully glottalized



(82) Pre-glottalized

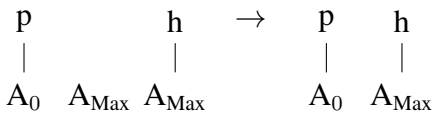


(83) Post-glottalized



Another crucial aspect of her proposal is that A_{Max} nodes merge if they are of the same type, adjacent, and are 'non-distinct in feature composition' (Steriade 1994: 213). She illustrates this merger in (84). The stop has a closure node and a release node, and the release node then merges with the adjacent node that [h] is attached to. The resulting structure is the same as that of a stop, with a closure position and a release position, and a single set of place and laryngeal features.

(84) Release merger in /ph/ clusters (Steriade 1994: 214)



I briefly outline two ways this account is relevant to Sevillian metathesis. First, as with binding theory, the bipositional structure of stops allows glottal features to be realized in various locations, and the release position provides a location for [h] to dock. Second, while her account does not address metathesis itself, it does address the distinction between clusters and segments, both underlyingly and on the surface. In applying this account to Sevillian metathesis, we first start with an unmetathesized [hC] realization, represented as [h], linked to its own A_{Max} node, followed by a [p] with a closure and release node. After metathesis happens, giving [Ch], the release merger illustrated in (84) would occur, so that [h] is linked to the release of the stop.

One potential issue is that [Ch] clusters seem marked in Spanish, and Gylfadottir (2015) has used this argument as a reason to treat them as aspirated stops instead. However, under Steri-

ade's account, the seeming markedness of [Ch] clusters may not be an issue. She argues that 'the relative markedness of different types of onset clusters is determined by the structural and featural similarity between single consonants and consonant clusters....the least marked onsets are identical in structure (i.e., sequence of A positions) and feature composition to single segments' (Steriade 1994: 213). Specifically, for sequences of stops and [h], she proposes that a cluster /ph/ and aspirated stop /p^h/ differ in underlying representation, but are identical on the surface due to the release merger shown in (84). The assumption that clusters and segments can be surface-identical, combined with the assumption that markedness depends on the similarity between clusters and segments, suggests that Sevillian [Ch] sequences may not be so bad. Although these clusters are unusual in Spanish, they are structurally identical to single-segment aspirated stops in aperture theory, and thus may occupy an intermediate status between single-consonant onsets and true onset clusters (e.g. [pr]).

6.2.4 Summary

In sum, Parrell's (2012) proposal that metathesis occurs due to pressures of gestural alignment seems to be the most plausible, in that it faces the fewest obstacles in accounting for the Sevillian patterns. However, to account for more nuances of the pattern, it will need to be augmented to dal with gemination, why metathesis is limited to certain clusters, the lack of tight correlation between parts (e.g. the duration of [h] before and after the consonant),and the actual articulatory patterns (once those are known).

6.3 Phonologization

My experimental results show that Sevillian stop-h sequences are not treated as being aspirated stops at the level of the underlying representation. Like Acenese and Old Khmer, stop-h sequences are clusters. But metathesis is well-established, in the sense that it is produced frequently, even in laboratory speech, and seems to be a production target. In this section, I discuss these sequences

in relation to the literature on phonologization (Section 6.3.1), and compare Sevillian to other languages in which contrastively aspirated stops have developed (Section 6.3.2).

6.3.1 Properties and requirements of phonologization

Phonologization is a process in which gradient phonetic outputs are categorized into discrete phonological representations. Hyman (1976) lays out several stages in phonologization. First, one sound has coarticulatory effects on an adjacent sound. Then, listeners interpret these phonetic effects as intentional, and the effects are exaggerated beyond what is phonetically natural. Hyman (1976) calls this step *phonologization*. Finally, the original segment conditioning the phonetic effect is lost, so words differ only in the new feature. Hyman (1976) calls this final stage *phonemicization*. To illustrate with vowel nasalization, a nasal consonant induces coarticulatory nasalization on a preceding vowel (/VN/ $\rightarrow$ [ṼN]). Then, the extent or intensity of nasalization on the vowel increases beyond the natural coarticulatory effect. Finally, the nasal consonant is lost ([Ṽ]). Nasalization no longer has a clear source, and is reinterpreted as part of the vowel. I follow Barnes' (2002: 28) definition of *phonologization* (which corresponds to Hyman's 1976 *phonemicization*): 'the innovation of changes to phonological representations.'

Hyman (2013) lays out several criteria to diagnose phonologization: (1) a phonetic effect is exaggerated beyond what is natural coarticulation; (2) a phonological rule targets the phonologized feature. A further requirement is (3) that the original source disappears, forcing listeners to reattribute the source of the phonetic effect (Hyman 1976; Ohala 1993). I deduce that a corollary of (3) is that the phonetic effect is no longer optional or variable, but rather consistently realized because it is inherent to the segment.

In addition to my experimental results, evaluating Sevillian stop-h sequences against these criteria makes it clear that they have not phonologized. The only criterion that stop-h sequences fulfill is that they have a natural coarticulatory, phonetic effect that has been exaggerated. In multiple Spanish varieties (including those without metathesis), stop releases in /s/-voiceless stop clusters have been found to be longer than in intervocalic voiceless stops (Torreira 2006). This appears

to be a natural coarticulatory phonetic effect. However, the duration of [h] in Sevillian stop-h sequences is much longer than can be attributed to natural coarticulation. Sevillian stop-h sequences do not fulfill the other criteria of phonolgization: no phonological rule targets aspirated stops, and the source (/s/) is still present, due to sociolinguistic variation and phonological alternations across word and morpheme boundaries.

If Sevillian stop-[h] sequences are not phonologizing, what is their representational status? Barnes (2002) distinguishes phonologization from the 'alteration or addition of phonetic targets to existing phonological representations.' Sevillian stop-h sequences may be at this latter stage. Stop-h realizations are intentional targets (Ruch and Peters 2016), beyond whatever coarticulation arises naturally in /sC/ sequences. Stop-h sequences are also used in perception: Sevillian listeners distinguish /s/-voiceless stop clusters from intervocalic voiceless stops based only on changes in the duration of [h] in [Ch] (Ruch and Harrington 2014). However, [h] still functions 'as [a] cue[] to *other* phonological categories' (Barnes 2002: 28; emphasis mine), not to the stop itself. As long as /s/ is occasionally present, listeners will attribute the [h] in [Ch] to it.

In the next section, I examine cases where aspirated stops *have* developed, as a counterpoint to Sevillian.

6.3.2 Diachronic pathways to contrastively aspirated stops

In considering why stop-h sequences in Sevillian Spanish are not phonologizing into aspirated stops, it is worth looking at how aspirated stops developed in other languages. The set of known pathways to contrastive pre and post-aspiration is fairly restricted. The Sevillian metathesis change that produces stop-h sequences is both similar to and different from other cases where aspirated stops have developed. In this section, I look briefly at pathways to aspirated stops in other languages, and speculate about how Sevillian fits into the attested patterns.

Clayton (2010) argues that preaspiration develops from two major sources: voiceless geminate stops (Nordic languages like Icelandic, Irish and Scottish Gaelic, Sami languages, Forest Nenets, Bora), and, less frequently, /nasal + voiceless stop/ clusters (central Algonquian languages

and Hopi). Blevins and Garrett (1993) and Blevins (2004) also cite voiceless geminates as a source, and further list Carolinian (Turkic), several Polynesian languages, New Caledonian, Modern Greek dialects, and dialects of Southern Italian (Blevins and Garrett 1993).

Blevins and Garrett (1993: 216-220) propose that spontaneous pre- or postaspiration can arise around voiceless geminate stops due to an increase in the magnitude of glottal opening that accompanies gemination, and slight mistiming of that glottal gesture. Pre- and postaspiration can arise if (a) the closure shortens slightly, or (b) the glottal gesture starts slightly too early or too late, through anticipatory or perseveratory coarticulation. They argue that this aspiration, which was originally unintentional, can then be enhanced and reinterpreted as a contrastive feature of the stop.

This connection between gemination and pre- and postaspiration is one of the reasons for my assumption that gemination is a precursor to metathesis in the analysis of Sevillian Spanish in Chapter 3 (Sections 3.3.2 and 3.3.6).[3] As already mentioned, Blevins and Garrett (1993) argue that pre- and postaspiration often arise around geminate stops cross-linguistically, and that aspiration has can be reinterpreted as a feature of the stop. If this is true, then it suggests that a geminate stop plus surrounding aspiration might be representationally ambiguous. Sevillian has [h]-geminate stop sequences, although [h] derives from /s/ instead of arising spontaneously. However, once these sequences exists, they are plausibly subject to the same forces that Blevins and Garrett (1993) propose for spontaneous aspiration. [hCː] sequences could present learners with ambiguous data about the source of [h]. If preaspiration is a natural phonetic consequence of geminate voiceless stops, then a listener who hears [pahtːa] could plausibly reinterpret the aspiration as belonging to the geminated voiceless stop. Once aspiration belongs to the stop, metathesis becomes more plausible since the timing of aspiration is not contrastive cross-linguistically (Steriade 1994; Ladefoged and Maddieson 1996; Kehrein and Golston 2004; Clayton 2010).

Postaspirated stops arise from a broader range of courses, including stress, contrast enhancement, consonant clusters, syllable structure simplification, long VOT due to phonetic factors,

[3]Recall that Sevillian Spanish is not the only dialect to have gemination, but it is the only dialect with metathesis; see discussion in Section 3.3.2.

and language contact.[4] These accounts are fewer and less detailed than those tracing the development of preaspirated stops, but I exemplify several cases below.

- **Basque:** Some varieties of Basque have contrastively aspirated stops, which Hualde (2018) hypothesizes may have come from various sources. He proposes that aspirated stops may have arisen from allophonically aspirated stops: aspirated stops would have occurred as the onset to stressed syllables (as in English). When the stress system changed, shifting stress off these syllables that had allophonically aspirated stops, the link between stress and aspiration was broken. Since aspiration was no longer predictable, it would have phonologized as contrastive.
- **Bantu:** Many Bantu languages have synchronic and diachronic connections between prenasalized voiceless stops and aspiration (*NT $\rightarrow$ NT^{h} $\rightarrow$ T^{h}; NT $\rightarrow$ nɦ $\rightarrow$ ɦ). Givón (1974) argues that the nasal devoiced, and then preaspiration was reinterpreted as postaspiration. Downing and Hamann (2018), however, find that these nasals are often fully voiced, and instead argue that postaspiration arose to enhance the contrast between *ND and *NT sequences (see also Hamann and Downing 2017 and Hinnebusch 1975).
- **Germanic:** Proto Indo-European had voiced aspirated, voiced, and voiceless unaspirated stops. It is unclear how aspiration was introduced to the system, but Iverson and Salmons (2003) suggest that contrast enhancement may have been at play. They suggest that a redundant [glottal width] gesture was added to voiceless stops, which were the unmarked (and laryngeally unspecified) member of the voicing contrast. This resulted in aspiration on voiceless stops.
- **Buyang:** Some dialects of Buyang (Kadai, Southern China) have aspirated stops, although Ancient Buyang did not (Li and Zhou 1998). Aspirated stops come from several sources: ancient voiced stops, fricatives, initial consonant clusters, and loanwords from Han Chinese that contain fricatives or aspirated stops. The majority came from consonant clusters (mostly voiceless stop + liquid), and Li and Zhou (1998: 132) propose that aspiration is 'compen-

[4]This list is non-exhaustive.

sation for the loss of the second consonant' (e.g. /*pl, *pr/ → /p^h/) or sometimes the first (*sk → k^h).

- **Ikalanga**: Ikalanga (Bantu, Zimbabwe and Botswana) has three types of aspirates: regular aspirated stops and an affricate, labialized aspirates, and breathy aspirates, which are phonologically different from regular aspirated stops (Mathangwane 1996). I focus on the first two, which have similar origins. Aspirates derived from three sources (Mathangwane 1996: 204-221). Regular aspirated stops and the aspirated affricate developed from Proto-Bantu close vowels /i̧, u̧/, which caused frication on preceding stops. Labialized aspirates derived from Proto-Bantu *u followed by a non-back vowel in the same syllable, which resulted in /w/. Both of these environments cause slightly longer VOT naturally, due to narrow constrictions, which he proposes could have been reinterpreted as phonemic aspiration.
- **Pali:** Pali underwent various changes in its development from Sanskrit, several of which involved metathesis in clusters with [h] and [s] (Vaux 1998; Suzuki 2002). For example, Sanskrit consonant clusters /hN, hR, SN/ metathesized into /sonorant + h/ in Pali. Clusters of [fricative + stop] changed into voiceless aspirated stops word-initially (Sanskrit: /ˈstana-/; Pali: /t^{h}ana/), and into aspirated geminates word-medially (Sanskrit: /ˈhasta-/; Pali /hatt^{h}a/). Vaux (1998) attributes this change to a general change in simplifying syllable structure: Sanskrit allowed complex onsets and place features in coda position, while Pali did not. To satisfy these new constraints in /sT/ clusters, for example, /s/ was delinked and its laryngeal features docked on the release of the following stop (Vaux 1998: 503-4).
- **Korean:** Aspirated stops in Korean are hypothesized to come from voiceless obstruents (possibly /*k/). Vowel syncope in Proto-Korean resulted in consonant clusters, some of which then became aspirated stops (Ramsey 1991). Hock (1986) also cites metathesis of syllable-final /h/ with a following voiceless stop as a source of aspiration.
- **Proto-Indo-European:** Rasmussen (1987) cites several origins of aspirated stops in Proto-Indo-European: (1) independent /h/; (2) secondary sources, with aspiration arising from

emphatic pronunciation or Siebs' law (devoicing of a root-initial voiced stop when a mobile /s/ attaches to it); (3) there is no clear source, so they are considered to be phonemes in PIE.

- **Latin:** Hock (1986: 206) states that Latin did not have aspirated stops prior to contact with Greek. Early Greek borrowings into Latin also lacked aspiration. As Greek gained prestige, however, Latin speakers started to hypercorrect, introducing aspiration into native Latin words.
- **Indonesian:** Some Indonesian languages have stops with aspiration and ambiguous voicing. These may have arisen through clusters of /voiced stop + h/, allophonic aspiration of originally voiced stops, and language contact (Hock 1986: 626).

These cross-linguistic patterns in how aspirated stops develop have several similarities to Sevillian Spanish metathesis. Several cases of contrastive aspiration seem to have developed from changes in syllable structure, through compensation for loss in one member of a cluster, and/or coalescence of stop-h sequences. Sevillian metathesis has all of these at play: the pressure towards open syllables pushes for metathesis, the coda /s/ in the cluster has been reduced, gemination compensates for reduction, and the resulting stop-h sequences could coalesce. The case most similar to Andalusian appears to be Pali. Like Pali, Sevillian Spanish has coda /s/ reduction to [h], gemination, and metathesis with voiceless stops. Also similar to Pali, Spanish shows a preference for open syllables, and metathesis can be seen as a way to remove the coda.

However, there are crucial differences between Sevillian Spanish and the other reported cases. One difference involves morpheme boundaries. Sevillian metathesis occurs across morpheme boundaries, which may prevent it from phonologizing. It is possible that stop-h sequences phonologize morpheme-internally but not across boundaries. Phonologization morpheme-internally but not across morpheme boundaries seems unlikely, but is not unattested. The Bantu language family provides an illustrative case. Synchronic alternations between nasals/aspiration and contrastive aspiration occur in the same languages, and appear to have arisen from the same processes. For example, Tumbuka (Downing and Hamann 2018) has contrastively aspirated stops /p ^{h}t^{h}k^{h}/, but this contrast only holds robustly in initial position. Tumbuka also has alternations

across morpheme boundaries between nasals (NT) and aspirated stops (NTh). Ikalanga also has both contrastively aspirated stops and synchronic alternations between nasals and aspiration at morpheme boundaries (Mathangwane (1996)). Adjectives that modify class 9/10 nouns take the prefix of the noun, which is nasal. The prefix surfaces as a nasal when the following sound is a voiced stop (/ŋombe + bilí/ → [ŋombe m-bilí], 'two cows'), but as breathy postaspiration when the following sound is a voiceless stop /ptk (/ŋombe + peɲú/ → [ŋombe pʰeɲú], 'alive cattle' (Mathangwane 1996: 231-232). It would be interesting to see if these derived aspirated stops differ phonetically from contrastively aspirated stops. While this pathway to aspiration—contrastive morpheme-internally, but not across boundaries—appears possible, it is uncommon.

Another difference is that in some of these cases, aspiration apparently developed as an allophonic variant in strong prosodic positions, or as way to enhance a voicing contrast. In Basque, aspiration functioned as an allophonic feature in prominent positions, and in Germanic, aspiration may have arisen to enhance a contrast between voicing categories. By contrast, for Sevillian, no case for enhancement can be made. If the goal were to enhance contrastiveness of voiceless stops, aspiration could arise spontaneously in environments without /s/, which it does not. On a more basic level, it does not make sense to say that [h] enhances the voicing contrast. [h] derives from /s/ (not the stops), and manipulations of it could only serve to enhance underlying /s/.

Although the existing literature on diachronic pathways to contrastive aspiration is relatively sparse, Sevillian does not show many similarities to the reported cases reviewed in this section. While this does not mean that phonologization is impossible, it does seem unlikely. I have also argued that there are other reasons stop-h sequences have not phonologized in Sevillian Spanish, namely, a combination of variation and alternations (see Section 3.4.1). Information about variation is not available for the reported cases here, but may prove to be an important source of evidence against phonologization for learners.

6.4 Metathesis in the phonological grammar

The results of the current study suggest that Sevillian metathesis operates on gestures at a relatively late derivational stage. The evidence that it operates on gestures comes from the phonetic study in Chapter 2. Evidence that it occurs late in the derivation comes from the stress judgment task in Chapter 4, where Sevillian listeners treated words with metathesis as if metathesis had not occurred, for the purposes of stress assignment. Do these properties of Sevillian metathesis extend to other cases of synchronic, productive metathesis as well?

In this section, I compare stress-metathesis interactions to stress-epenthesis interactions, which have been argued to occur at different stages in phonology (Elfner 2009). The goal of this comparison is to explore the hypothesis that metathesis might differ from epenthesis in ways that are informative about the structure of the phonological grammar. While epenthesis has been argued to occur at both early and late derivational stages, I suggest that productive metathesis might be limited to late derivational stages. Furthermore, the lateness of metathesis might affect what kinds of representations it operates on: metathesis may manipulate gestures, rather than segmental units. Finally, I speculate that metathesis might be restricted to being derivationally late because of how it is conditioned.[5]

6.4.1 Asymmetries between stress-epenthesis vs. stress-metathesis interactions

Different types of segmental processes seem to interact differently with stress. Stress-epenthesis interactions can be transparent (conditioning factors are visible on the surface: epenthesis before stress) or opaque (conditioning factors are not visible on the surface: stress before epenthesis) (Section 6.4.1.1). Less attention has been paid to stress-metathesis interactions. In Sevillian, the stress-metathesis interaction is opaque: stress occurs before metathesis. For other languages as well, stress appears to occur before metathesis 6.4.1.2).

[5]This section is heavily informed by joint work with Kate Mooney (Gilbert and Mooney To appear).

6.4.1.1 Stress-epenthesis interactions

Stress-epenthesis interactions can be either transparent or opaque (Elfner 2009). In Egyptian Arabic, the interaction is transparent: epenthesis precedes stress. Stress is typically penultimate (85a), and epenthetic vowels in penultimate position are stressed as if they were underlying vowels (85b) (epenthetic vowels are underlined). The epenthetic vowel must be present before stress is assigned, resulting in a transparent interaction.

(85) Egyptian Arabic: transparent stress-epenthesis interaction (Elfner 2009: 20)

a. Underlying vowel: /madrasa/	→	[mad'rasa]	'school'
b. Epenthetic vowel: /bint-na/	→	[bin't<u>i</u>na]	'our daughter'

In Dakota, the interaction between stress and epenthesis is opaque. Regular stress is peninitial (86a), but stress retracts to the initial syllable when the peninitial syllable contains an epenthetic vowel (86b). Stress is initial in (86b) because it was assigned before the epenthetic vowel was added. If epenthesis occurred before stress, stress would have been assigned in the normal pattern (peninitial). The interaction is opaque because the conditioning for initial stress is not visible in the surface form.

(86) Dakota: opaque stress-epenthesis interaction (Elfner 2009: 23-5)

a. Underlying vowel: /wičha-ya-kte/	→	[wi'čhayakte]	'you kill them'
b. Epenthetic vowel: /puz/	→	['puz<u>a</u>]	'to be dry'

Elfner (2009) argues that the fact that languages like Egyptian Arabic and Dakota differ in how stress and epenthesis interact is strong evidence for serial derivations. If languages treat epenthetic vowels differently, then epenthesis must happen at different times. However, parallel analyses are also possible, using HEAD-DEP constraints that prevent non-underlying vowels from being stressed (Alderete 1999; Broselow 2008).

In a different line of research, Hall (2006) distinguishes between two types of inserted vowels: epenthetic vowels vs. intrusive vowels. She argues that the types of inserted vowels behave differently, with intrusive vowels being invisible to other phonological processes, and epenthetic vowels serving to repair structural issues. Several other properties distinguish these

types of epenthetic vowels, leading her to propose that intrusive vowels are 'purely a phenomenon of the gestural layer', while epenthetic vowels are added at the segmental level (Hall 2006: 424). The account also makes predictions about phonetic, timing, coarticulatory properties of different types of inserted vowels, most of which remain to be tested.

The distinction Hall makes between epenthetic and intrusive vowels can be tied to the transparent and opaque stress-epenthesis patterns illusrtated in Egyptian Arabic and Dakota, as both suggest a difference in when the processes occur. More specifically, what she terms epenthetic vowels seem likely to correlate with those that are inserted early, and to participate in transparent stress-epenthesis interactions where epenthesis precedes stress. The inserted vowels that she calls intrusive seem likely to correlate with those that are inserted late, and thus participate in opaque stress-epenthesis interactions. Whether these correlations hold up has not been addressed in the literature.

6.4.1.2 Stress-metathesis interactions

While stress-epenthesis interactions seem to be either transparent opaque, the interaction between stress and synchronic, productive metathesis seems to only be opaque: stress precedes metathesis. In Hall's terms, metathesis would only operate at the level that intrusive vowels operate. In this section, I review several cases of synchronic, productive metathesis and suggest that there may be an asymmetry between epenthesis and metathesis in terms of when they occur. Testing this claim will require extensive typological, phonetic, and phonological work.

There seem to be two types of systems with regular, synchronic metathesis: languages where stress occurs before metathesis but has no impact on it (which I will call Type 1), and languages where stress occurs before metathesis and determines whether metathesis takes place (which I will call Type 2). Type 1 languages include Sevillian Spanish and Lithuanian. In Sevillian Spanish, metathesis occurs regardless of the location of stress (Torreira 2012; Horn 2013). Moreover, the lack of interaction between metathesis and stress assignment in the stress judgment

experiment in Chapter 4 indicates that stress assignment and evaluation occurs at an intermediate level of representation, before metathesis.

In Lithuanian, sequences of coronal fricative + [k] metathesize to [k] + coronal fricative before a following consonant (87) (Seo and Hume 2001).

(87) Lithuanian metathesis (Seo and Hume 2001: 211)

a.	[tvʸeske]	3SG.PAST.IMPERF	'flash briefly'
b.	[tvʸeksti]	INFIN	

Like in Sevillian, Lithuanian metathesis occurs regardless stress location (Seo and Hume 2001). Even though stress does not affect metathesis, stress must precede metathesis because stress is lexical (Halle 1997).

Type 2 languages—where stress precedes metathesis and directly *conditions* metathesis—include Faroese and many Austronesian languages, like Rotuman and Uab Meto. Faroese metathesis is a more constrained version of Lithuanian metathesis. [sk] metathesizes to [ks] in very specific phonological contexts: before another stop, following a vowel or nasal, and only when the preceding vowel is stressed (Seo and Hume 2001). Following a liquid and following an unstressed vowel, the medial stop deletes instead. (88) illustrates the effect of stress. [sk] metathesizes to [ks] when the preceding vowel is stressed (88a). When the preceding vowel is unstressed, the medial stop deletes instead (88b).

(88) Faroese metathesis is conditioned by stress (Seo and Hume 2001: 213)

a.	[ˈraskor]	M.3SG	[ˈrakst]	NEUT.SG	'energetic'	(*[ˈraskt, ˈrast])
b.	[ˈføːrɪsk]	M.3SG	[ˈføːrɪst]	NEUT.SG	'Faroese'	(*[ˈføːrɪskt, ˈføːrɪkst])

Rotuman CV metathesis is also directly conditioned by stress.[6] Blevins and Garrett (1998: 527) analyze Rotuman CV metathesis diachronically: a final unstressed vowel undergoes a stage of extensive coarticulation with a preceding stressed vowel, and finally deletes: $\acute{V}_1CV_2 > \acute{V}_1\check{V}_2C\check{V} > \acute{V}_1\check{V}_2C$. Examples (89a-89b) illustrate metathesis: the final vowel metathe-

[6]Whether Rotuman metathesis is prosodically or morphologically driven is not settled.

sizes leftwards, resulting in a diphthong and no final vowel. This metathesis only occurs when the final vowel is unstressed and is preceded by a stressed vowel.

(89) Rotuman metathesis is conditioned by stress (Blevins and Garrett 1998: 527)

a.	mó<u>se</u>	→	mø<u>́</u>s	'to sleep'
b.	ses<u>é</u>va	→	ses<u>éa</u>v	'erroneous'

Blevins and Garrett (1998) propose that the stressed vowel is long, facilitating the perception of the diphthong resulting from CV metathesis. The properties of a stressed vowel facilitate metathesis, so stress directly allows metathesis.

Another case where metathesis is fed by stress assignment is Uab Meto. In this language, one type of CV metathesis is partially triggered by prosodic pressures: metathesis reduces the lapse between the stressed vowel in the root and the end of the phonological phrase under suffixation (Mooney To appear). When a suffix is added, it would create a right-edge lapse. Under her analysis, the right-edge lapse (*ˈ$\sigma\sigma\sigma$) is solved when the final vowel slot of the root deletes, and the vowel spreads to (and coalesces with) the vowel in the preceding stressed syllable (90). If metathesis is caused by pressure to reduce lapse between a stressed vowel and the end of the phonological phrase, then metathesis must be able to see stress.

(90) Uab Meto metathesis reduces lapse (Mooney To appear: 2)

a.	ˈkok<u>ɪ</u>s	'bread'	ˈko͡<u>ɪ</u>ks-e	'the bread'	(*ˈkokɪs-e)
	ˈ$\sigma\sigma$		ˈ$\sigma\sigma$		(*ˈ$\sigma\sigma\sigma$)

Together, these cases support a serial analysis in which stress precedes metathesis. In other words, the interaction is opaque.

There are few cases of synchronic metathesis, and further research is needed to make typological claims about metathesis more broadly. However, the asymmetry between epenthesis and metathsis, where epenthesis can be transparent or opaque, and metathesis appears to be largely opaque, suggests that the processes may differ in a fundamental way.

6.4.2 Differences in motivations for epenthesis and metathesis

Why does metathesis seem to occur more commonly at late derivational stages, whereas epenthesis can occur at both early and late stages? I speculate that one difference lies in the factors that condition each process, and that metathesis may be similar to late epenthesis. Recally that derivationally early epenthesis has been argued to occur in order to improve syllable structure and metrical parsing (Elfner 2009). Late epenthesis (recall the opaque interaction between stress and epenthesis in Dakota; Elfner 2009), does not affect processes like stress, and I suggested that the opaque, late epenthesis proposed by Elfner (2009) may correspond to Hall's (2006: 391) intrusive vowels (although this correlation remains to be tested). Hall argues that intrusive vowels are purely gestural, and share a cluster of properties that mark them as surface-oriented: they are invisible to other phonology, lack distinct phonetic quality of their own, are affected by coarticulation with surrounding segments, tend to be optional, variable, and conditioned by speech rate, and they do not repair phonotactically illicit structures. Regular metathesis may be like intrusive vowels: it seems to be largely surface-oriented. More specifically, it is often argued to be triggered by surface-oriented constraints on production (coarticulatory pressures, gestural coordination) and perception.

Coarticulatory pressures have been argued to underlie several metathesis patterns. For Sevillian Spanish, recall that one motivation proposed in the literature is a pressure towards in-phase gestural coordination (Section 6.2). Under this explanation, metathesis from [hC] to [Ch] may be due to general properties of human motor coordination: humans perform better when actions start in-phase (at the same time) rather than anti-phase (sequentially). For Rotuman, Blevins and Garrett (1998) argue that metathesis results from perceptual and coarticulatory pressures. Coarticulatory metathesis begins with extensive coarticulation between a final unstressed vowel and a preceding long stressed vowel. Coarticulation allows for a long, easily perceptible diphthong, which compensates for a weak unstressed final vowel. Blevins and Garrett (1998) also argue that coarticulatory metathesis may be constrained in some languages by phonetics and contrast. They state that extensive vowel-to-vowel coarticulation is less likely if a language has a crowded vowel space, vowels with dispersed targets, or a large number of diphthongs, because this

coarticulation would result in merger with an existing category. Furthermore, they argue that coarticulatory metathesis is less likely in languages with long consonants or consonant clusters, whose length inhibits vowel-to-vowel coarticulation. The factors leading to and preventing coarticulatory metathesis are based on coarticulatory or perceptual pressures.

Furthermore, although my ABX task found no evidence that metathesis is perceptually optimizing for [h], most regular, synchronic metathesis patterns have long been argued to improve the perceptibility of one or both segments. The earliest comprehensive accounts of metathesis point to perceptibility as a main motivation. Ultan (1978: 395) states that 'the superficial cause of most metatheses is conversion of one phonologically inadmissible or disfavored sequence into an acceptable one,' and Hock (1985: 532-3) argues that all *regular* metathesis patterns have a 'structural purpose,' which is 'converting phonologically or perceptually 'marked' structures into more acceptable ones.' Later, Hume (1998) argues that 'metathesis occurs in contexts of low salience to enhance perception of certain sounds' (cited in Hume 2001: 7). For Lithuanian, Seo and Hume (2001) and Hume (2001) argue that metathesizing coronal fricatives + velar stops when followed by another consonant improves perceptibility. For example, metathesis in forms like [me<u>ks</u>-ti] 'knit', inf. (cf. [mezg-a] '3-PAST') puts the consonant into post-vocalic position and separates the two consonants, which aids perception. In the non-metathesized form *[meskti], the stops are adjacent, which is not perceptually ideal for either. Similar arguments have been made for pre- and postaspiration: the different forms surface in positions that most maximize the cues (Steriade 1997).

From a non-teleological perspective, perception has also been argued to play a role in metathesis. As discussed in Section 6.1, certain segments with long cues (glottals, liquids, glides) have been observed to undergo metathesis frequently. Arguably, the location of these cues is ambiguous, which sets the stage for listeners to perceive them in non-historical locations. In a slightly different vein, Takahashi (2019) develops a formal account of CC metathesis that relies on cue ambiguity Takahashi (2019) analyzes metathesis as a sequential operation of fusion and splitting, instead of a single operation. The consonants first fuse together, resulting in overlapped

segments that are 'analysed as an unordered complex segment at some point in the derivation' (Takahashi 2019: 718). Then, they split apart and one surfaces on the non-historical side. One reason she gives for preferring a fuse-and-split account is that fusion into an overlapped or complex segment could give rise to the cue ambiguity that makes reinterpretation possible at the splitting stage. Because cue ambiguity is a phonetic, surface property of the sounds involved, I take this to mean that metathesis is not purely phonologically motivated. Although Takahashi (2019) does not frame fusion and splitting articulatorily, her 'fused' stage is reminiscent of an intermediate stage of gestural overlap, with segments unordered in a timing slot, as I have argued for Sevillian.

The perceptual accounts of metathesis—that listeners mishear the location of segments and that metathesis improves perceptibility of one or both segments—remain largely untested. My ABX discrimination task in Chapter 5 tests these questions for [h] metathesis, and finds no supporting evidence among listeners of multiple different language backgrounds. However, this does not mean that other metathesis patterns are not perceptually motivated or facilitated, or that all metathesis patterns are motivated in a single way. It could well be the case that some synchronic, productive metathesis patterns are perceptually motivated, others are motivated by pressures of gestural coordination, and still others have motivations not yet discovered.

In sum, many metathesis patterns have been argued to be conditioned (or facilitated) by acoustic, coarticulatory, and perceptual factors. These pressures apply at the surface, once the speaker is aware of the particular coarticulatory and perceptual pressures at play. If it is true that metathesis is triggered by surface considerations, while vowel insertion can be surface-oriented (intrusive) *or* fix deeper structural problems (epenthesis), then an asymmetry may be expected. In this case, we would expect metathesis to occur after stress assignment, leading to opaque interactions with stress. Vowel insertion, which can be triggered *either* by surface pressures or deeper structural pressures, can occur before or after stress, leading to transparent or opaque interactions with other processes. The comparison between metathesis and epenthesis suggests that processes could be specific to certain levels of phonological derivation, and when they apply determines how they interact with other processes.

CHAPTER 7

Conclusion

The main goal of this dissertation was to investigate the phonetic, representational, phonological, and perceptual properties of a sequence of sounds undergoing change, which has the potential to create a new phonological contrast. Most changes either neutralize contrasts or do not affect the number of contrasts in a language (Gurevich 2003); contrast formation is rare. In this dissertation, I use Sevillian Spanish as a case study: /sC/ clusters are undergoing a metathesis change, surfacing instead as stop-h sequences [Ch]. I present four experiments that look at both the 'what' and the 'why' of the change: What do stop-h sequences look like, phonetically and representationally? Why might this change be taking place? Based on the experimental results, I propose an analysis that relies on the gradual nature of serialism to analyze how the phonological grammar maps /sC/ to stop-h sequences. The mappings from surface to underlying forms, and vice versa, are crucial for explaining participants' responses in the experiments.

7.0 Experimental results

The dissertation presents four experiments, whose results converge: Sevillian stop-h sequences are representationally /sC/ clusters, and metathesis occurs late in the phonology.

The production experiment (Chapter 2) investigates the phonetics of /sC/ sequences. For /s/-voiceless stop and /s/-voiced stop clusters, there is phonetic evidence of gestural overlap: gestures from /s/ overlap those of the consonant in phonetically measurable ways, as predicted by articulatory accounts (Torreira 2006; Parrell 2012). /s/-voiceless stop clusters are the only ones to show full metathesis. The overlap process that applies to /s/-voiced stop clusters may be similar, but since voiced stops spirantize and lack a release in this context, the phonetic outcome looks very different. The results support treating some metathesis patterns as occurring through gestural overlap (e.g. Grammont 1933; Blevins and Garrett 1998; Mooney To appear). /s/-sonorant clusters show no evidence of overlap.

The fill-in-the-blank study (Chapter 3) showed that Sevillian listeners deconstruct [Ch] sequences into underlying /sC/ clusters. [Ch] sequences are derived, and the two components are treated as separable. Listeners of other Spanish dialects (Mexican and Argentinian Spanish) do not undo metathesis of [Ch] sequences to arrive at underlying /sC/. In my analysis, this difference in perception behavior is because the Sevillian grammar has a direct mapping between the underlying /sC/ cluster and the surface metathesized [Ch] form, while the others do not. This direct mapping allows Sevillians to undo metathesis and arrive at the underlying cluster.

Building on the finding that stop-h sequences are derived, the stress judgment task (Chapter 4) tests how these sequences interact with the phonological system at a higher level. Sevillian listeners treat surface-light syllables that are derived by metathesis (/<u>CVS</u>.CV/ → [<u>CV</u>.ChV]) as heavy when judging stress. These results suggest an opaque interaction: stress is assigned and judged at an intermediate level of phonological representation before metathesis has occurred. Metathesis occurs derivationally late.

Finally, the ABX perception task (Chapter 5) investigates potential perceptual causes of laryngeal metathesis. Results show that speakers of different languages and dialects (Arabic, English, French, Mexican Spanish, Argentinian Spanish, Sevillian Spanish) perceive [h] and its linear location more or less accurately depending on their native language phonological inventories and phonotactics. Sevillian listeners do not differ from other Spanish-speaking listeners in this task,

suggesting that [Ch] fails to map to an underlying /sC/ form, unless it is explicitly and plausibly connected to a lexical representation containing /sC/. The results did not support laryngeal metathesis from [hC] → [Ch] as percpetually optimizing at a universal or language-specific level as no listener group perceived stop-h better than h-stop. This result again emphasizes the role of mappings: perception is controlled by language-specific mechanisms, and failure to map leads to failure of perception.

All of the experimental results point to stop-h sequences being represented as clusters of two independent segments, /sC/. Furthermore, the results suggest that metathesis occurs late in the phonology. Metathesis is clearly derived (sometimes with parts coming from different words), it must occur after stress assignment, and it is so surface-level that Sevillians do not recognize [Ch] as different from [C] when [Ch] is presented out of context. I have proposed that variability in /sC/ realizations as well as phonological alternations may provide crucial information to learners that stop-h sequences are derived, not representationally aspirated stops (Chapters 3 and 6).

7.1 Analyzing metathesis

The analyses capture the experimental results by showing that listeners' responses to specific forms in perception can be explained by their grammars' ability to generate these forms in production.

The analysis of three Spanish dialects (corresponding to the fill-in-the-blank task) casts metathesis as a gradual process of reduction: coda /s/ debuccalizes to [h], causes compensatory gemination of the following consonant, and finally metathesizes out of the coda altogether. These steps represent both the diachronic trajectory of metathesis in Sevillian as well as synchronic variation between forms within and across dialects. For Sevillian, the proposed grammar maps /sC/ to surface [Ch]. In perception, then, these listeners are able to 'undo' metathesis by undoing each step of the derivation. In contrast, Argentinian and Mexican Spanish grammars map /sC/ to [sC] and [hC], respectively. Because these listeners' grammars cannot map /sC/ to [Ch] in production, the grammars also cannot map [Ch] back to an underlying form in perception.

The analysis based on the stress experiment shows that stress must precede metathesis in order to explain Sevillians' stress preferences. In the experiment, listeners disprefer antepenultimate stress words with heavy penults (over those with light penults), reflecting patterns in the Spanish lexicon and restrictions on stress and weight. When the penult is originally heavy but light on the surface (as happens with metathesis), listeners still disprefer antepenultimate stress. I propose that this dispreference is because their grammars do not produce antepenultimate stress in words with heavy penults. Instead, stress gets pulled onto the penult in the course of the derivation. Stress must apply before metathesis to explain why surface CVC and CV.CH penults are both treated as heavy: the penult is heavy at the point when stress applies (before metathesis), preventing antepenultimate stress in both. My listeners disprefer antepenultimate stress on forms with derived light penults because their grammars cannot derive these forms.

7.2 Conclusion

In this dissertation, I have investigated the phonetic, phonological, and perceptual properties of an ongoing change in Sevillian Spanish, from [hC] sequences to [Ch] sequences. [Ch] sequences could phonologize into aspirated stops, raising questions about representation and segmenthood. The change provides a unique opportunity to investigate the relationship between phonetic and representational change, as well as probe the way mappings operate between surface and underlying forms. Based on the results of production and perception studies, I have argued that this metathesis has not resulted in a change in representation, and that [Ch] sequences occur through a mechanism of gestural overlap and realignment that happens late in the phonology. I also found no evidence that this change (and potentially other patterns of laryngeal metatheses and pre/postaspiration) is driven by perceptual optimization. Future investigations of what factors are necessary and sufficient for phonologization will further contribute to our understanding of what segments are and how they are represented.

Bibliography

Abrego-Collier, Carissa. 2013. Liquid dissimilation as listener hypocorrection. *Annual Meeting of the Berkeley Linguistics Society* 37:3.

Adisasmito-Smith, Niken. 2004. Phonetic and phonological influences of Javanese on Indonesian. PhD Dissertation, Cornell University, Ithaca, NY.

Alderete, John. 1999. Head dependence in stress-epenthesis interaction. In *Derivational residue in phonological Optimality Theory*, 29–50. Philadelphia, PA: John Benjamins Publishing Company.

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

Anderson, Stephen R. 1974. *The organization of phonology*. New York: Academic Press.

Arnason, Kristján. 1986. The segmental and supersegmental status of preaspiration in modern Icelandic. *Nordic Journal of Linguistics* 9:1–23.

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

Baković, Eric. 2009. Abstractness and motivation in phonological theory. *Studies in Hispanic and Lusophone Linguistics* 2:183–198.

Baković, Eric. 2016. Exceptionality in Spanish stress. *Catalan Journal of Linguistics* 15:9–25.

Bals Baal, Berit Anne, David Odden, and Curt Rice. 2012. An analysis of North Saami gradation. *Phonology* 29:165–212.

Barnes, Jonathan. 2002. Positional neutralization: A phonologization approach to typological patterns. PhD Dissertation, University Of California, Berkeley, Berkeley, CA.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.

Beckman, Jill, Michael Jessen, and Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49:259–284.

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

Beddor, Patrice Speeter. 2009. A coarticulatory path to sound change. *Language* 85:785–821.

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, Indiana University, Bloomington.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104:591–630.

Berg, Thomas. 2001. An experimental study of syllabification in Icelandic. *Nordic Journal of Linguistics* 24:71–106.

Berkson, Kelly Harper. 2019. Acoustic correlates of breathy sonorants in Marathi. *Journal of Phonetics* 73:70–90.

Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-

speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345–360.

Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20:305–330.

Best, Catherine T., and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Language experience in second language teaching and learning: In honor of James Emil Flege*, ed. Ocke-Schwen Bohn and Murray J. Munro, volume 17 of *Language Learning & Language Teaching*, 13–34. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bishop, Jason B. 2007. Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differenes involved in s-aspiration. *Proceedings of ICPhS XVI* 1765–1768.

Bladon, Anthony. 1986. Phonetics for hearers. In *Language for hearers*, ed. Graham McGregor, 1–24. Great Britain: Pergamon Press.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK/New York: Cambridge University Press.

Blevins, Juliette, and Andrew Garrett. 1993. The evolution of Ponapeic nasal substitution. *Oceanic Linguistics* 32:199–236.

Blevins, Juliette, and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74:508–556.

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

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

Broselow, Ellen. 2008. Stress-epenthesis interactions. In *Rules, constraints, and phonological phenomena*, ed. Bert Vaux and Andrew Nevins, 121–148. Oxford/New York: Oxford University Press.

Browman, Catherine, and Louis Goldstein. 1986. Towards an Articulatory Phonology. *Phonology Yearbook* 3:219–252.

Browman, Catherine, and Louis Goldstein. 1988. Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45:140–155.

Browman, Catherine, and Louis Goldstein. 2001. Competing constraints on integestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5:25–34.

Broś, Karolina, Marzena Żygis, Adam Sikorski, and Jan Wołłejko. 2021. Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. *Phonology* 38:1–40.

Butera, Brianna J. 2018. A lenition continuum: Acoustic variability of Spanish stop consonants. PhD Dissertation, University of Wisconsin-Madison, Madison, WI.

Bárkányi, Zsuzsanna. 2002. A fresh look at quantity sensitivity in Spanish. *Linguistics* 40:375–394.

Béland, Renée, and Régine Kolinsky. 2005. One sound heard as two: The perception of affricates in Quebec French by Belgian French speakers. *Journal of Multilingual Communication Disorders* 3:110–117.

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

Campos-Astorkiza, Rebeka. 2012. The phonemes of Spanish. In *The handbook of Hispanic linguistics*, ed. José Ignacio Hualde, Antxon Olarrea, and Erin O'Rourke, Blackwell Handbooks in Linguistics, 89–110. Chichester, West Sussex, UK/Malden, MA: Wiley-Blackwell, 1st edition.

Campos-Astorkiza, Rebeka. 2016. Procesos fonológicos. In *Enciclopedia de lingüística hispánica*, ed. Javier Gutiérrez-Rexach, volume 1, 847–858. London/New York: Routledge Taylor & Francis Group.

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

Carrasco, Patricio, José Ignacio Hualde, and Miquel Simonet. 2012. Dialectal differences in Spanish voiced obstruent allophony: Costa Rican vs. Iberian Spanish. *Phonetica* 69:149–179.

Casserly, Elizabeth D. 2012. Gestures in optimality theory and the laryngeal phonology of Faroese. *Lingua* 122:41–65.

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. *Journal of the International Phonetic Association* 49:1–32.

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

Cho, Taehong, and Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27:207–229.

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

Coetzee, Andries, and Joe Pater. 2011. The place of variation in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, Blackwell

Handbooks in Linguistics. Chichester, West Sussex, UK/Malden, MA: Blackwell Publishing Ltd., 2nd edition.

Colantoni, Laura. 2006. Micro and macro sound variation and change in Argentine Spanish. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 91–102. Somerville, MA: Cascadilla Proceedings Project.

Cole, Jennifer, José Ignacio Hualde, and Khalil Iskarous. 1999. Effects of prosodic and segmental context on /g/-lenition in Spanish. In *Proceedings of the Fourth International Linguistics and Phonetics Conference*, ed. O. Fujimura, B. D. Joseph, and B. Palek, volume 2, 575–589.

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

Cronenberg, Johanna, Michele Gubian, Jonathan Harrington, and Hanna Ruch. 2020. A dynamic model of the change from pre- to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 83:1–22.

Cuetos, Fernando, Maria Glez-Nosti, Analía Barbón, and Marc Brysbaert. 2012. SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica* 33:133–143.

Dautriche, Isabelle. 2017. Words cluster phonetically beyond phonotactic regularities. *Cognition* 163:128–145.

Davidson, Lisa. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38:272–288.

Davidson, Lisa. 2011. Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 37:270–282.

Davidson, Lisa. 2018. Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association* 48:331–356.

Davidson, Lisa, and Jason A. Shaw. 2012. Sources of illusion in consonant cluster perception. *Journal of Phonetics* 40:234–248.

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. *Linguistics* 41:607–652.

Del Saz, María. 2019. From postaspiration to affrication: New phonetic contexts in Western Andalusian Spanish. *Proceedings of ICPhS XIX* 760–764.

Dell, François. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12:31–37.

Downing, Laura J. 2005. On the ambiguous segmental status of nasals in homorganic NC sequences. In *The internal organization of phonological segments*, ed. Marc van Oostendorp and Jeroen van de Weijer, 183–216. Berlin/Boston: De Gruyter Mouton.

Downing, Laura J., and Silke Hamann. 2018. The phonetics of NCh in Tumbuka and its implications for diachronic change. *Papers in Historical Phonology* 3:77–95.

Duchon, Andrew, Manuel Perea, Nuria Sebastián-Gallés, Antonia Martí, and Manuel Carreiras. 2013. EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods* 45:1246–1258.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568–1578.

Dupoux, Emmanuel, Erika Parlato, Sonia Frota, Yuki Hirose, and Sharon Peperkamp. 2011. Where do illusory vowels come from? *Journal of Memory and Language* 64:199–210.

Durvasula, Karthik, Mohammed Qasem Ruthan, Sarah Heidenreich, and Yen-Hwei Lin. 2021. Probing syllable structure through acoustic measurements: Case studies on American English and Jazani Arabic. *Phonology* 38:173–202.

Elfner, Emily. 2009. Syllabification and stress-epenthesis interactions in Harmonic Serialism. *ROA 1047* .

Ennever, Thomas, Felicity Meakins, and Erich R. Round. 2017. A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology* 8:20.

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

Esposito, Christina M., and Sameer ud Dowla Khan. 2012. Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42:123–143.

Farris-Trimble, Ashley W. 2008. Cumulative faithfulness constraints in phonology. PhD Dissertation, Indiana University, Bloomington, IN.

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

Fuchs, Martín. 2018. Antepenultimate stress in Spanish: In defense of syllable weight and grammatically-informed analogy. *Glossa* 3:1–34.

Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20:269–337.

Galarza, Iraida, Gibran Delgado-Días, and Erik Willis. 2014. Nuevamente /s/? Una nueva mirada a la elisión de /s/ implosiva en el español de Puerto rico. Presentation at the 7th Workshop on Spanish Sociolinguistics, Indiana University.

Gallagher, Gillian. 2012. Perceptual similarity in non-local laryngeal restrictions. *Lingua* 122:112–124.

Gallagher, Gillian. 2016. Asymmetries in the representation of categorical phonotactics. *Language* 92:557–590.

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

Garcia, Guilherme. 2017. Weight gradience and stress in Portuguese. *Phonology* 34:41–79.

Garcia, Guilherme. 2019. When lexical statistics and the grammar conflict: Learning and repairing weight effects on stress. *Language* 95:612–641.

Garellek, Marc, and Christina M. Esposito. 2021. Phonetics of White Hmong vowel and tonal contrasts. *Journal of the International Phonetic Association* 1–20.

Garellek, Marc, and James White. 2015. Phonetics of Tongan stress. *Journal of the International Phonetic Association* 45:13–34.

Gerfen, Chip. 2001. A critical view of licensing by cue: Codas and obstruents in Eastern Andalusian Spanish. In *Segmental phonology in Optimality Theory*, 183–205. Cambridge University Press.

Gerfen, Chip. 2002. Andalusian codas. *Probus* 14:247–277.

Gilbert, Madeline. 2022. Conflicting standards and variability: Spirantization in two varieties of Uruguayan Spanish. Manuscript.

Gilbert, Madeline, and Kate Mooney. To appear. Metathesis is late and fake. In *Proceedings of the 2021 Annual Meeting on Phonology*, ed. Peter Jurgec, Liisa Duncan, Emily Elfner, Yoonjung Kang, Alexei Kochetov, Brittney K. O'Neill, Avery Ozburn, Keren Rice, Nathan Sanders, Jessamyn Schertz, Nate Shaftoe, and Lisa Sullivan. Washington, DC: Linguistic Society of America.

Givón, Talmy. 1974. Rule un-ordering: Generalization and de-generalization in Phonology. In *Papers from the Parasession on Natural Phonology*, ed. Anthony Bruck, Robert Allen Fox, and Michael W. La Galy, 103–115. Chicago Linguistic Society.

Goldsmith, John. 1976. Autosegmental phonology. PhD Dissertation, MIT, Cambridge, MA.

Golston, Chris, Wolfgang Kehrein, Jean Léo Léonard, and Samia Naïm. 2013. A prosodic theory of laryngeal timing. In *Base articulatoire arrière. Backing and Backness*, LINCOM Studies in Phonology, 11–44. LINCOM EUROPA.

Gouskova, Maria, and Nancy Hall. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In *Phonological argumentation: Essays on evidence and motivation*, ed. Steve Parker, 203–225. London, UK: Equinox.

Gouskova, Maria, and Juliet Stanton. 2021. Learning complex segments. *Language* 97:151–193.

Grammont, Maurice. 1933. *Traité phonetique*. Paris: Librarie Delagrave.

Gurevich, Naomi. 2003. Functional constraints on phonetically conditioned sound changes. PhD Dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.

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

Gylfadottir, Duna. 2018. The effective borrowing of a phonemic contrast. PhD Dissertation, University of Pennsylvania, Philadelphia, PA.

Hall, Nancy. 2003. Gestures and segments: Vowel intrusion as overlap. Doctoral Dissertation, University of Massachusetts Amherst, Amherst, MA.

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23:387–429.

Hall, Nancy, Bianca Godinez, Megan Walsh, Araceli Carmona, and Sarah Garcia. 2020. Perceptual dissimilation of /l/ and (not) /n/. *Proceedings of the Linguistic Society of America* 5:446.

Hall, Nancy, Bianca Godinez, Megan Walsh, Irene Orellana, and Coleen Villegas. 2019. Evidence for perceptual hypercorrection in American r-dissimilation: A pilot study. *Proceedings of the Linguistic Society of America* 4:52.

Hall, Nancy, Nancy Vasquez, Francisco Aguirre, Muhammad Damanhuri, and Connor Tree. 2017. Long-distance liquid coarticulation in American English. *Supplemental proceedings of the 2016 Annual Meeting on Phonology* .

Halle, Morris. 1997. On stress and accent in Indo-European. *Language* 73:275.

Hallé, Pierre A., and Catherine T. Best. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *The Journal of the Acoustical Society of America* 121:2899–2914.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder, and Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24:592–608.

Hamann, Silke, and Laura J. Downing. 2017. *NT revisited again: An approach to postnasal laryngeal alternations with perceptual CUE constraints. *Journal of Linguistics* 53:85–112.

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

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

Herrero de Haro, Alfredo. 2017. The phonetics and phonology of Eastern Andalusian Spanish: A review of literature from 1881 to 2016. *Íkala, Revista de Lenguaje y Cultura* 22:313–357.

Harrington, Jonathan. 2010. Acoustic phonetics. In *The handbook of phonetic sciences*, ed. William J. Hardcastle, John Laver, and Fiona E. Gibbon, 81–129. Wiley, 1st edition.

Harrington, Jonathan, Felicitas Kleber, and Mary Stevens. 2016. The relationship between the (mis)-parsing of coarticulation in perception and sound change: Evidence from dissimilation and language acquisition. In *Recent advances in nonlinear speech processing*, ed. Anna Esposito, Marcos Faundez-Zanuy, Antonietta M. Esposito, Gennaro Cordasco, Thomas Drugman, Jordi Solé-Casals, and Francesco Carlo Morabito, 15–34. Dordrecht: Springer.

Harris, James. 1983. *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge, MA: MIT Press.

Harris, James. 1991. With respect to accentual constituents in Spanish. In *Current studies in Spanish linguistics*, ed. Héctor Campos and Fernando Martínez-Gil, 447–474. Washington D.C.: Georgetown University Press.

Harris, James. 1992. *Spanish stress: The extrametricality issue*. Indiana University Linguistics Club.

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

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253–306.

Heimisdóttir, Linda. 2015. The phonology of aspiration in Icelandic: A gesture-based approach. PhD Dissertation, Cornell University, Ithaca, NY.

Helgason, Pétur. 2002. Preaspiration in the Nordic Languages. PhD Dissertation, Stockholm University, Stockholm, Sweden.

Heman-Ackah, Yolanda D., Deirdre D. Michael, Margaret M. Baroody, Rosemary Ostrowski, James Hillenbrand, Reinhardt J. Heuer, Michelle Horman, and Robert T. Sataloff. 2003. Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otology, Rhinology & Laryngology* 112:324–333.

Henriksen, Nicholas. 2017. Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception. *Journal of Phonetics* 63:106–126.

Henriksen, Nicholas, and Lorenzo García-Amaya. 2012. Transcription of intonation of Jerezano Andalusian Spanish. *Estudios de fonética experimental* XXI:109–162.

Herbert, Robert K. 1986. *Language universals, markedness theory, and natural phonetic processes*. Berlin, New York: Mouton de Gruyter.

Hermes, Anne, Doris Mücke, and Martine Grice. 2013. Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology* 30:1–25.

Hillenbrand, James, and Robert A. Houde. 1996. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research* 39:311–321.

Hinnebusch, Thomas. 1975. A reconstructed chronology of loss: Swahili class 9/10. In *Proceedings of the Sixth Conference on African Linguistics*, volume 20, 32–41. Ohio State University Working Papers in Linguistics.

Hock, Hans. 1985. Regular Metathesis. *Linguistics* 23:529–546.

Hock, Hans. 1986. *Principles of historical linguistics*. Berlin, Germany: Mouton de Gruyter.

Holmberg, E., and A. Gibson. 1979. On the distribution of [h] in the languages of the world. *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. Marc van Oostendorp, Jeroen van de Weijer, Harry van der Hulst, Jan Koster, and Henk van Riemsdijk, number 77 in Studies in Generative Grammar, 319–354. Berlin, New York: Mouton de Gruyter.

Horn, Meagan E. 2013. An acoustic investigation of postaspirated stops in Seville Spanish. Master's thesis, The Ohio State University, Columbus, Ohio.

Hualde, Jose Ignacio. 2012. Stress and Rhythm. In *The handbook of Hispanic linguistics*, ed. Jose Ignacio Hualde, Anxton Olarrea, and Erin O'Rourke, Blackwell Handbooks in Linguistics. Chichester, West Sussex, UK/Malden, MA: Wiley-Blackwell, 1st edition.

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

Hualde, José Ignacio. 1989b. Procesos consonánticos y estructuras geométricas en español. *Lingüística* 1:7–44.

Hualde, José Ignacio. 2018. Aspiration in Basque. *Papers in Historical Phonology* 3:1–27.

Hualde, José Ignacio, and Sonia Colina. 2014. *Los sonidos del español*. Cambridge: Cambridge University Press.

Hualde, José Ignacio, Ryan Shosted, and Daniel Scarpace. 2011. Acoustics and articulation of Spanish /d/ spirantization. *Proceedings of ICPhS XVII* 906–909.

Hume, Elizabeth. 1998. The role of perceptibility in consonant/consonant metathesis. In *Proceedings of the West Coast Conference on Formal Linguistics 17*, ed. Kimary N. Shahin, Susan Blake, and Eun-Sook Kim, 293–307. Stanford: CSLI.

Hume, Elizabeth. 2001. Metathesis: Formal and functional considerations. In *Surface syllable structure and segment sequencing*, ed. Elizabeth Hume, Norval Smith, and Jeroen van de Weijer, HIL Occasional Papers. Leiden: HIL.

Hume, Elizabeth. 2004. The indeterminacy/attestation model of metathesis. *Language* 80:203–237.

Hussain, Qandeel, and Shigeko Shinohara. 2019. Partial devoicing of voiced geminate stops in Tokyo Japanese. *The Journal of the Acoustical Society of America* 145:149–163.

Hyman, Larry M. 1976. Phonologization. In *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*, ed. Alphonse Julliand, 407–418. Saratoga: Anma Libri, 2nd edition.

Hyman, Larry M. 2013. Enlarging the scope of phonologization. In *Origins of sound change: Approaches to phonologization*, ed. Alan C. L. Yu, 3–28. Oxford: Oxford University Press.

Idsardi, William. 2006. Canadian Raising, opacity, and rephonemicization. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 51:119–126.

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language & Linguistic Theory* 7:217–259.

Itô, Junko, Armin Mester, and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26:571–613.

Iverson, Gregory K., and Joseph C. Salmons. 2003. Laryngeal enhancement in early Germanic. *Phonology* 20:2003.

Jarosz, Gaja. 2016. Learning opaque and transparent interactions in Harmonic Serialism. In *Proceedings of the 2015 Annual Meeting on Phonology*, ed. Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin, and Douglas Pulleyblank.

Johnson, Daniel Ezra, and Jennifer Nycz. 2015. Partial mergers and near-distinctions: Stylistic layering in dialect acquisition. *University of Pennsylvania Working Papers in Linguistics* 21.

Kabak, Bariş, and William J. Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50:23–52.

Kager, René. 2012. Stress in windows: Language typology and factorial typology. *Lingua* 122:1454–1493.

Keating, Patricia, Christina M. Esposito, Marc Garellek, Sameer ud Dowla Khan, and Jianjing Kuang. 2010. Phonation contrasts across languages. In *Working papers in phonetics*, volume 108. UCLA, Los Angeles: UCLA Department of Linguistics.

Keer, Edward William. 1999. Geminates, the OCP and the nature of CON. PhD Dissertation, Rutgers, New Brunswick, NJ.

Kehrein, Wolfgang, and Chris Golston. 2004. A prosodic theory of laryngeal contrasts. *Phonology* 21:325–357.

Kelso, J. A. 1984. Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 246:R1000–R1004.

Kim, Lan, and John Alderete. 2008. Why [h] and aspirated conosnants are different in Korean. *Toronto Working Papers in Linguistics* 28:123–136.

Kim, S.-H. 1989. The behavior of /h/ in Korean. In *Proceedings of the Harvard Workshop on Korean Linguistics III*, 117–126.

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

Kiparsky, Paul. 1973. Phonological representations. In *Three Dimensions of Linguistic Theory*, 3–136. Tokyo: TEC.

Kirby, James. 2018. praatsauce: Praat-based tools for spectral analysis. URL https://github.com/kirbyj/praatsauce.

Kirchner, Robert Martin. 1998. An effort-based approach to consonant lenition. PhD Dissertation, University of California, Los Angeles, Los Angeles.

Kouneli, Maria. 2016. Acoustics of epenthetic vowels in Brazilian Portuguese. *The Journal of the Acoustical Society of America* 140:3109–3109.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82:1–26.

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

Labov, William. 2001. The anatomy of style-shifting. In *Style and sociolinguistic variation*, ed. Penelope Eckert and John Rickford, 85–108. Cambridge: Cambridge University Press.

Labov, William. 2010. *Principles of linguistic change, vol. 3 Cognitive and cultural factors*. Oxford: Wiley-Blackwell.

Ladefoged, Peter, and Keith Johnson. 2011. *A course in phonetics*. Boston, MA: Wadsworth, Cengage Learning, 6th edition.

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

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

Lewis, Anthony. 2001. Weakening of intervocalic /p, t, k/ in two Spanish Dialects: Toward the quantification of lenition processes. PhD Dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign.

Li, Jingfang, and Guoyan Zhou. 1998. Diachronic evolution of initial consonants in Buyang. *Mon-Khmer Studies* 28:115–135.

Lipski, John M. 2011. Socio- phonological variation in Latin American Spanish. In *The handbook of Hispanic sociolinguistics*, ed. Manuel Díaz Campos, 72–97. Oxford, UK: Wiley-Blackwell.

Lisker, Leigh, and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *WORD* 20:384–422.

Llisteri, Joaquim, María Machuca, Carme de la Mota, Montserrat Riera, and Antonio Ríos. 2003. The perception of lexical stress in Spanish. *Proceedings of ICPhS XV* 2023–2026.

Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language & Linguistic Theory* 8:375–425.

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

Lombardi, Linda. 1998. Evidence for MAX-Feature constraints from Japanese. ROA-247, College Park, Maryland.

Lombardi, Linda. 2001. Why Place and Voice are different: Constraint-specific alternations in Optimality Theory. In *Segmental phonology in Optimality Theory*, ed. Linda Lombardi, 13–45. Cambridge University Press, 1st edition.

van der Loo, M. P. J. 2014. The stringdist package for approximate string matching. *The R Journal* 6:111–122. URL https://CRAN.R-project.org/package=stringdist.

Madill, Catherine, Duong Duy Nguyen, Kristie Yick-Ning Cham, Daniel Novakovic, and Patricia McCabe. 2019. The impact of nasalance on cepstral peak prominence and harmonics-to-noise ratio. *The Laryngoscope* 129.

Maguire, Warren, Lynn Clark, and Kevin Watson. 2013. Introduction: What are mergers and can they be reversed? *English Language and Linguistics* 17:229–239.

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

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

Martínez-Celdrán, Eugenio. 1984. Cantidad e intensidad en los sonidos obstruyentes del castellano: hacia una caracterización acústica de los sonidos aproximantes. In *Estudios de fonética experimental*, 71–129. Barcelona: Promociones Publicaciones Universidarias.

Martínez-Celdrán, Eugenio, and Xosé Luís Regueira. 2008. Spirant approximants in Galician. *Journal of the International Phonetic Association* 38:51–68.

Martínez-Paricio, Violeta, and Maria-Rosa Lloret. 2017. /s/-weakening in Andalusian Spanish: From (partial) deletion to (in)complete gemination. Presentation at The XXV Manchester Phonology Meeting.

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

Mathangwane, Joyce Thambole Mogatse. 1996. Phonetics and phonology of Ikalanga: A diachronic and synchronic study. PhD Dissertation, University of California, Berkeley, Berkeley, CA.

McCarthy, John. 2007. *Hidden Generalizations: phonological opacity in optimality theory*. Great Britain: Equinox.

McCarthy, John. 2008a. The gradual path to cluster simplification. *Phonology* 25:271–319.

McCarthy, John. 2008b. The serial interaction of stress and syncope. *Natural Language & Linguistic Theory* 26:499–546.

McCarthy, John, and Kathryn Pruitt. 2013. Sources of phonological structure. In *Linguistic Derivations and Filtering: Minimalism and Optimality Theory*. London: Equinox.

Melero-García, Fernando. 2022. The effect of duration and degree of constriction on the perception of Spanish phonemic voiced and voiceless stops. Presentation at the Linguistic Symposium on Romance Languages 52.

Mielke, Jeff. 2003. The interplay of speech perception and phonology: Experimental evidence from Turkish. *Phonetica* 60:208–229.

Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2003. Looking through opacity. *Theoretical Linguistics* 29:123–139.

Minegishi, M. 2006. Khmer. In *Encyclopedia of Language & Linguistics*, 189–192. Elsevier.

Mooney, Kate. To appear. Phonology cannot transpose: Evidence from Meto. *Phonology* .

Moreno-Fernández, Francisco. 2019. *Variedades de la lengua española*. Routledge introductions to Spanish Language and Linguistics. New York: Routledge, 1st edition.

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. *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. *Spanish in Context* 17:200–220.

Mustafawi, Eiman. 2018. Arabic phonology. In *The Routledge handbook of Arabic linguistics*, ed. Elabbas Benmamoun and Reem Bassiouney, 11–31. London and New York: Routledge Taylor & Francis Group.

Ni Chasaide, Ailbhe. 1985. Preaspiration in phonological stop contrasts: An instrumental study. PhD Dissertation, University College of North Wales, Bangor.

Nycz, Jennifer. 2013. New contrast acquisition: methodological issues and theoretical implications. *English Language and Linguistics* 17:325–357.

Núñez-Cedeño, Rafael, Sonia Colina, and Travis Bradley. 2014. *Fonología generativa contemporánea de la lengua española*. Washington, DC: Georgetown University Press.

Ohala, John J. 1980. The application of phonological universals in speech pathology. In *Speech and Language: Advances in Basic Research and Practice*, 75–97. New York: Academic Press.

Ohala, John J. 1981a. Articulatory constraints on the cognitive representation of speech. In *The Cognitive Representation of Speech*, ed. Terry Myers, John Laver, John Anderson, G. E. Stelmach, and P. A. Vroon, Advances in Psychology. Amsterdam, Netherlands: North-Holland Publishing Company.

Ohala, John J. 1981b. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior, 17th Annual Regional Meeting, Chicago Linguistic Society*, ed. Roberta A. Hendrick, Carrie S. Masek, and Mary Prances Miller, 178–203. Chicago: Chicago Linguistic Society.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The Production of Speech*, ed. Peter F. MacNeilage, 189–216. New York, NY: Springer New York.

Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In *Language change: Contributions to the study of its causes*, ed. L. E. Breivik and E. H. Jahr, number 43 in Trends in Linguistics, Studies and Monographs, 173–198. Berlin: De Gruyter Mouton.

Ohala, John J. 1993. The phonetics of sound change. In *Historical linguistics: Problems and perspectives*, ed. Charles Jones, 237–278. London/New York: Longman.

Ohala, John J., and Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In *Language and its ecology: Essays in Memory*

of Einar Haugen, ed. Stig Eliasson and Ernst H. Jahr, 343–365. Berlin, New York: Mouton de Gruyter.

O'Neill, Paul. 2009. S-aspiration and occlusives in Andalusian Spanish: Phonetics or phonology? *Oxford University Working Papers in Linguistics, Philology & Phonetics* 12:73–85.

O'Neill, Paul. 2010. Variación y cambio en las consonantes oclusivas del español de Andalucía. *Estudios de fonética experimental* XIX:11–41.

Ortega-Llebaria, Marta. 2004. Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In *Laboratory Approaches to Spanish Phonology*, ed. Timothy L. Face, number 7 in Phonology and Phonetics, 237–253. Berlin/New York: Mouton de Gruyter.

Ortega-Llebaria, Marta, and Pilar Prieto. 2010. Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54:73–97.

den Os, Els, and René Kager. 1986. Extrametricality and stress in Spanish and Italian. *Lingua* 69:23–48.

Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics* 40:37–45.

Pater, Joe. 1999. Austronesian nasal substitution and other NC effects. In *The prosody-morphology interface*, ed. René Kager, Harry van der Hulst, and Wim Zonneveld, 310–343. Cambridge: Cambridge University Press.

Piñeros, Carlos-Eduardo. 2016. The phonological weight of Spanish syllables. In *The syllable and stress: Studies in honor of James W. Harris*, ed. Rafael A. Núñez-Cedeño, 271–314. Berlin, Boston: De Gruyter.

Powell, J. V. 1985. An Occurrence of Metathesis in Chimakuan. *Oceanic Linguistics Special Publications* 105–110.

Prince, Alan. 1990. Quantitative consequences of rhythmic organization. In *Parasession on the syllable in phonetics and phonology*, ed. M. Ziolkowski, M. Noske, and K. Deaton, 355–389. Chicago: Chicago Linguistic Society.

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

Pruitt, Kathryn. 2010. Serialism and locality in constraint-based metrical parsing. *Phonology* 27:481–526.

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

Quilis, Antonio. 1981. *Fonética acústica de la lengua español*. Madrid: Gredos.

Ramsey, S. Robert. 1991. Proto-Korean and the origin of Korean accent. In *Studies in the historical phonology of Asian languages*, ed. William G. Boltz and Michael C. Shapiro, volume 77 of *Current Issues in Linguistic Theory*, 215–238. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Rasin, Ezer, and Roni Katzir. 2016. On evaluation metrics in Optimality Theory. *Linguistic Inquiry* 47:235–282.

Rasmussen, Jens Elmegård. 1987. On the status of the aspirated tenues and the Indo-European phonation series. *Acta Linguistica Hafniensia* 20:81–109.

RCoreTeam. 2020. R: A language and environment for statistical computing. URL https://www.R-project.org.

Regan, Brendan. 2017. A study of ceceo variation in Western Andalusia (Huelva). *Studies in Hispanic and Lusophone Linguistics* 10:119–160.

Repiso-Puigdelliura, Gemma, Isabel Benvenuti, and Ji Young Kim. 2021. Heritage speakers' production of the Spanish voiced palatal obstruent /j/: A closer look at orthography and universal phonetic principles. *Heritage Language Journal* 18:1–30.

Ridouane, Rachid, G. N. Clements, and Rajesh Khatiwada. 2011. Language-independent bases of distinctive features. In *Tones and features: Phonetic and phonological perspectives*, ed. John A. Goldsmith, Elizabeth Hume, Leo Wetzels, Harry van der Hulst, Jan Koster, and Henk van Reimsdijk, volume 107 of *Studies in Generative Grammar*, 264–291. Berlin, Boston: De Gruyter Mouton.

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

Roca, Iggy. 1990a. Diachrony and synchrony in word stress. *Journal of Linguistics* 26:133–164.

Roca, Iggy. 1990b. Verb morphology and stress in Spanish. *Probus* 2:321–350.

Roca, Iggy. 1991. Stress and syllables in Spanish. In *Current studies in Spanish linguistics*, ed. Héctor Campos and Fernando Martínez-Gil, 599–635. Washington D.C.: Georgetown University Press.

Rogers, Brandon. 2016. The influence of linguistic and social variables in the spirantization of intervocalic /bdg/ in Concepción, Chile. *Studies in Hispanic and Lusophone Linguistics* 9:207–237.

Romero, Joaquín. 1994. Fricatives and approximants in Andusian Spanish. Presentation at the 24th Linguistic Symposium on Romance Languages.

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

Rosenthall, Sam, and Harry van der Hulst. 1999. Weight-by-Position by Position. *Natural Language & Linguistic Theory* 17:499–540.

Rosner, Burton S., Luis López-Bascuas, José E. García-Albea, and Richard P. Fahey. 2000. Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics* 28:217–224.

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

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. Silvia Calamai, Chiara Celata, and Luca Ciucci, 61–64. Pisa, Italy: Edizioni della Normale.

Ruch, Hanna. 2013. Investigating a gradual metathesis: Phonetic and lexical factors on /s/ aspiration in Andalusian Spanish. *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 41* 19:171–180.

Ruch, Hanna, and Jonathan Harrington. 2014. Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 45:12–25.

Ruch, Hanna, and Sandra Peters. 2016. On the origin of post-aspirated stops: Production and perception of /s/ + voiceless stop sequences in Andalusian Spanish. *Laboratory Phonology* 7:1–36.

Sagey, Elizabeth. 1986. *The representation of features and relations in non-linear phonology*. Published: Doctoral dissertation, MIT. [Published 1990, New York: Garland.].

Salvador, Gregorio. 1968. Neutralización G-/K- en español. volume 4, 1739–1754. Madrid: Revista de filología española.

Sampson, Rodney. 1999. *Nasal vowel evolution in Romance*. Oxford: Oxford University Press.

Sanders, Robert. 2003. Opacity and sound change in the Polish lexicon. PhD Dissertation, University of California, Santa Cruz, Santa Cruz, CA.

Santagada, Miguel, and Jorge Gurlekian. 1989. Spanish voiced stops in VCV contexts: Are they fricative variants or approximants? *Revue de phonétique appliquée* 91:363–375.

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

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. Ana Maria Carvalho and Sara Beaudrie, 189–202. Somerville, MA: Cascadilla Proceedings Project.

Schmidt, Lauren, and Erik Willis. 2011. Systematic investigation of voicing assimilation of Spanish /s/ in Mexico City. In *Selected proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, ed. Scott Alvord. Somerville, MA: Cascadilla Proceedings Project.

Schwarz, Martha, Morgan Sonderegger, and Heather Goad. 2019. Realization and representation of Nepali laryngeal contrasts: Voiced aspirates and laryngeal realism. *Journal of Phonetics* 73:113–127.

Seo, Misun, and Elizabeth Hume. 2001. A comparative account of metathesis in Faroese and Lithuanian. In *Surface syllable structure and segment sequencing*, ed. Elizabeth Hume, N. Smith, and J. van de Weijer, 210–229. Leiden: Holland Institute of Generative Linguistics.

Shaw, Jason, Adamantios I. Gafos, Philip Hoole, and Chakir Zeroual. 2009. Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation. *Phonology* 26:187–215.

Shaw, Jason, Sejin Oh, Karthik Durvasula, and Alexei Kochetov. 2021. Articulatory coordination distinguishes complex segments from segment sequences. *Phonology* 38:437–477.

Shelton, Michael. 2007. An experimental approach to syllable weight and stress in Spanish. PhD Dissertation, Pennsylvania State University, State College, PA.

Shelton, Michael, Chip Gerfen, and Nicolás Gutiérrez Palma. 2012. The interaction of subsyllabic encoding and stress assignment: A new examination of an old problem in Spanish. *Language and Cognitive Processes* 27:1459–1478.

Shelton, Michael, Chip Gerfen, and Nicolás Gutiérrez Palma. 2009. Proscriptions. . . Gaps. . . and Something in Between: An Experimental Examination of Spanish Phonotactics. In *Romance linguistics 2007: Selected papers from the 37th Linguistic Symposium on Romance Languages*, ed. Pascual José Masullo, Erin O'Rourke, and Chia-Hui Huang, 261–275. Amsterdam: John Benjamins Publishing Company.

Shelton, Michael, and Hannah Grant. 2018. Syllable weight in monolingual and heritage Spanish. *Studies in Hispanic and Lusophone Linguistics* 11:395–427.

Shih, Chilin, Bernd Möbius, and Bhuvana Narasimhan. 1999. Contextual effects on consonant voicing profiles: A cross–linguistic study. In *Proceedings of the 14th International Congress of Phonetic Science (San Francisco, CA)*, 989–992.

Silverman, Daniel. 2003. On the rarity of pre-aspirated stops. *Journal of Linguistics* 39:575–598.

Simonet, Miquel, José Ignacio Hualde, and Marianna Nadeu. 2012. Lenition of /d/ in spontaneous Spanish and Catalan. In *Proceedings of the 13th Annual Conference of the International Speech Communication Association (INTERSPEECH 2012)*, 1416–1419.

Soler, Antonia, and Joaquín Romero. 1999. The role of duration in stop lenition in Spanish. *Proceedings of ICPhS XIV* 483–486.

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

Spinu, Laura, Alexei Kochetov, and Jason Lilley. 2018. Acoustic classification of Russian plain and palatalized sibilant fricatives: Spectral vs. cepstral measures. *Speech Communication* 100:41–45.

Stanton, Juliet. 2017. Constraints on the distribution of nasal-stop sequences: An argument for contrast. PhD Dissertation, MIT, Cambridge, MA.

Steriade, Donca. 1993. Closure, release, and nasal contours. In *Nasals, nasalization, and the velum*, ed. Stephen Anderson, Patricia Keating, Marie Huffman, and Rena Krakow, volume 5 of *Phonetics and Phonology*, 401–470. Elsevier, 1st edition.

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

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. UCLA, Los Angeles.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In *Perception in Phonology*, ed. Elizabeth Hume and Keith Johnson, 291–250. San Diego: Academic Press.

Stevens, Mary. 2010. How widespread is preaspiration in Italy? A preliminary acoustic phonetic overview. In *Working Papers of the Department of Linguistics and Phonetics 2010*, ed. S. Schötz and G. Ambrazaitis, 97–102. Lund University.

Stevens, Mary, and John Hajek. 2010. Post-aspiration in standard Italian: some first cross-regional acoustic evidence. In *The 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 2010)*, 1557–1560. Makuhari, Japan.

Suh, Chang-Kook. 2001. Aspects of phonetics and phonology of Icelandic preaspiration. *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*. Number 9 in Studia Humaniora Ouluensia. Oulu, Finland: Oulu University Press.

Suzuki, Yasuko. 2002. Consonant cluster changes in Pali: Toward restricting the phonological patterns (Part 1). *Journal of Inquiry and Research* 75:97–125.

Takahashi, Chikako. 2019. No transposition in Harmonic Serialism. *Phonology* 36:695–726.

Terrell, Tracy D. 1979. Final /s/ in Cuban Spanish. *Hispania* 62:599–612.

Thráinsson, Höskuldur. 1978. On the phonology of Icelandic pre-aspiration. *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. Nuria Sagarra and Almeida Jacqueline Toribio, 113–120. Somerville, MA: Cascadilla Proceedings Project.

Torreira, Francisco. 2007. Pre- and postaspirated stops in Andalusian Spanish. In *Segmental and prosodic issues in Romance phonology*, ed. Pilar Prieto, Joan Mascaró, and Maria-Josep Solé, volume 282 of *Current Issues in Linguistic Theory*, 67–82.

Torreira, Francisco. 2012. Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association* 42:49–63.

Torreira, Francisco, and Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology* 2:331–353.

Torres-Tamarit, Francesc, Clàudia Pons-Moll, and Maria Cabrera-Callís. 2012. Rhotic metathesis in Algherese Catalan: A Harmonic Serialism account. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, 354–364. Somerville, MA: Cascadilla Proceedings Project.

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

Tuller, Betty, and J.A. Scott Kelso. 1990. Phase transitions in speech productions and their perceptual consequences. In *Attention and performance XIII: Motor representation and control*, ed. M. Jeannerod, 429–452. Hillsdalen, NJ: Lawrence Erlbaum, 1st edition.

Tunley, Alison. 1999. Coarticulatory influences of liquids on vowels in English. PhD Dissertation, University of Cambridge, Cambridge, UK.

Ultan, Russell. 1978. A typological view of metathesis. In *Universals of human language*, ed. Joseph H. Greenberg, volume 2, 368–399. Stanford, CA: Stanford University Press.

Vaux, Bert. 1998. The laryngeal specifications of fricatives. *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.

Vida Castro, Matilde. 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. *Lingua Americana* 20:15–36.

Vitevitch, Michael S., and Paul A. Luce. 2016. Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics* 2:75–94.

Vogel, Irene, Angeliki Anthanasopoulou, and Nadya Pincus. 2016. Prominence, contrast and the functional load hypothesis: An acoustic investigation. In *Dimensions of phonological stress*, ed. Jeffrey Heinz, 123–167. Cambridge, UK: Cambridge University Press.

Waltermire, Mark. 2004. The effect of syllable weight on stress in Spanish. In *Laboratory Approaches to Spanish Phonology*, ed. Timothy L. Face, 171–191. Berlin: Mouton de Gruyter.

Webb, Charlotte Ann Yates. 1974. Metathesis. PhD Dissertation, The University of Texas at Austin, Austin, Texas.

Werker, Janet F., and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49–63.

Whalen, D.H., Catherine T. Best, and Julia R. Irwin. 1997. Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics* 25:501–528.

Widdison, Kirk. 1995. Two models of Spanish s-aspiration. *Language Sciences* 17:329–434.

Wolf, Matthew. 2007. For an autosegmental theory of quirky mutation. In *University of Massachusetts occasional papers in linguistics 32: Papers in Optimality Theory III*, 315–404. Amherst, MA: GLSA.

Yoon, Tae-Jin. 2012. Laryngeal metathesis as perceptual optimization. *Studies in Phonetics, Phonology and Morphology* 18:413–435.

Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12:85–129.

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

Zoll, Cheryl. 1996. Parsing below the segment in a constraint-based framework. PhD Dissertation, University of California, Berkeley, Berkeley, CA.

Zsiga, Elizabeth. 1993. Features, gestures, and the temporal aspects of phonological organization. PhD Dissertation, Yale University, New Haven, CT.