Preface to "Sound parallels?"

Presented in this manuscript is a lightly edited version of my term paper for Mairi-Louise McLaughlin's graduate seminar course "Linguistic History of the Romance Languages" given in the Spring 2021 term at UC Berkeley. This was my first serious endeavor into linguistics, in an attempt to address an odd phonological pattern I sensed some years back between Ibero-Romance and Chinese language families. There are certain claims about Chinese history and Chinese languages that could have benefited from more citational evidence, but for the most part, I am quite satisfied with how this paper turned out. May I say, what an utmost amusing experience it was to find my hypothesis completely disproved!

Adrianne Zhong

August 2021

Sound parallels? Contrasting consonant change within Ibero-Romance and Chinese language varieties

Adrianne Zhong

1 Introduction

An important question in phonology, one that has inspired and continues to inspire much debate, is what Hinskens poses as the Why of sound change: "Does [a particular instance of sound change] result from internal dynamics or from extensive contact with a related dialect or another language?" (2020: 7). In this term paper, I address this question for consonant changes that on the surface are paralleled between Ibero-Romance and Chinese varieties.

I can identify two global events in 2019 that together mark the inception of this project: (1) In June 2019, protests erupted in Hong Kong against the 2019 Hong Kong extradition bill, and (2) on March 21, 2019, Wikipedia's Today's featured article was on Jorge Ben's 1970 album, *Fôrça Bruta*, spurring within me a lasting interest in Brazilian music, in particular Bossa Nova and Tropicália movements.[1] Following the HK protest movement and obsessively consuming 70s Brazilian music exposed me to Cantonese and Portuguese, at which point I was struck by certain paralleled phonological differences when they are compared to Mandarin[2] and Spanish:

1. Certain *n*'s in Spanish and Mandarin (henceforth Sp and Man) correspond to *m*'s in Portuguese and Cantonese (Pt and Can): *una flor* ['u**n**a 'floɾ] vs. *uma flor* ['u**m**a 'floR] 'a flower', cf. 男 *nán* [*na**n***ˊ] vs. *naam4* [*na**m***˩˩˧] 'male'.[3]

[1] Along with *Fôrça Bruta,* I would greatly recommend the albums *Gal Costa* by Gal Costa (1969) and *Brasil* by João Gilberto (1981) as starting points. For a modern continuer with extraordinary songwriting, check out *<atrás/além>* by O Terno (2019). Be warned, this music is dangerously enthralling.

[2] To clarify, I use the word "Mandarin" to refer to Standard Beijing Mandarin, and "Mandarin dialects" to refer to the Mandarin dialect supergroup.

[3] I include Pinyin (Man) and Jyutping (Can) romanizations and IPA tone markers here and here alone in my introduction, just for comprehensiveness. These aspects are not relevant for the rest of the paper.

2. Certain *h*'s in Sp/Man correspond to *f* 's in Pt/Can: *hacer* [a'θeɾ] vs. *fazer* [**f**a'zeR] 'to do', cf. 火 *huǒ* [***x***uɔ˨˩˦] vs. *fo2* [***f***ɔ˥] 'fire'.
3. Instances of affricate palatalization in Sp and Man that are lacking in Pt and Can: *luchar* [lu'**ʧ**aɾ] vs. *lutar* [lu'taR] 'to fight', cf. 身 *shēn* [***ʂ***ən˥] vs. *san1* [***s***ɐn˥] 'body'.

In short, I sensed a systematic parallel relationship in sound changes between the variety pairs, representable as Sp : Pt :: Man : Can,[4] and in my excitement, hypothesized that all these parallel consonant changes arose from parallel mechanisms of contact and internal sound change.[5]

This paper is a report of an extensive survey of linguistics literature I conducted to test this hypothesis. Naturally, the extent of my analysis will be limited by the availability and reliability of existing studies.[6] As well, given the broad scope of my inquiry, I had to compromise between depth and breadth of coverage: my analysis of each of the sound changes is far from a comprehensive, and I concentrate only on the standard forms of the four languages, giving examples from regional varieties only as they pertain to the discussed phonological change. With that being said, within its limited scope, my investigation was in certain ways fruitful. It will soon become clear that I discovered many surprises along the way.

In section 2, I give brief sketches of each of the four languages, illustrating both diachronic development with a focus on language contact, and synchronic phonology with a focus on consonants. In section 3, I describe the methodology guiding my literature survey, and explore sequentially correspondences between *n* vs. *m* and *h* vs. *f* in Sp/Man vs. Pt/Can, and affricate palatalization in Sp/Man not observed in Pt/Can. In section 4, I discuss my findings and address Hinskens's question of Why, and briefly sketch possible directions for future research.

---

[4] It is also interesting to note that while Spanish and Mandarin each have only 5 vowel phonemes, Portuguese has 14 (9 oral and 5 nasal), and Cantonese 11 (4 short and 7 long).
[5] Of course, at the base of my hypothesis was the assumption that all these sound changes were regular.
[6] Unfortunately, a substantial amount of Chinese linguistic research is published in Chinese, which I cannot yet easily read.

2.1 A Brief Linguistic Overview of Ibero-Romance

Spanish and Portuguese are the two most widely-spoken Romance languages today, with an estimated 471.4 and 232.4 million native speakers as of 2021 respectively (Eberhard et al. 2021). As Ibero-Romance languages they developed from Latin, and share many traits in lexis and grammar with each other.[7] Most of the phonemic inventory of consonants is shared between the two varieties (Appendix A), the primary difference being that only Spanish has palato-alveolar affricates, and only Portuguese the postalveolar fricatives.

The following diachronic overview of Spanish and Portuguese is based primarily off of Alkire & Rosen's (2010). Modern Spanish and Portuguese linguistic history arguably begins in 711 with the Moorish conquest of the Iberian Peninsula. It took several centuries for the various Iberian kingdoms to regain the lost territory, but not before Moorish occupation left an impact on language, mostly in lexicon.[8] In 1093 the kingdoms of Galicia and Portugal split from León and Castile, starting the linguistic split between Galician-Portuguese and Spanish.

Pertaining to Spanish, Basque is a genetically unrelated non-Romance language that originates from mountainous North Spain; its existence and contact with Castilian Spanish form a large part of language contact discourse with regards to the well-studied *f* > *h* sound change in Spanish, discussed in section 3.2. Just as well, during the Italian renaissance and its aftermath, there was a reinvigorated interest in the usage of Latin, leading to many *cultismos*, i.e. lexical borrowings from Latin to Spanish, which need to be taken into consideration for *f* > *h*.[9]

---

[7] Other Ibero-Romance varieties include Astur-Leonese, Aragonese, Catalan, and Occitan, though there are many subtleties in their classification, in particular for those falling between Ibero- and Gallo-Romance. See Bossong (2016).

[8] There are hundreds of words that have readily identifiable Arabic origins. For instance: *alcachofa, alcachofra* (Sp, Pt) 'artichoke' from Arabic *al-ḫarshoof*, same meaning; *enchufar* (Sp) "to plug in/connect/perform nepotism" from Andalusi Arabic *Juf,* 'stomach'; *aceite, azeite* (Sp, Pt) 'olive oil' from Arabic *az-zayt*, same meaning.

[9] Without a doubt, the colonial expansion of the Spanish and Portuguese empires in the 15th to 18th centuries has also led to significant language contact, in their colonial occupancy of Africa, the Americas, and Southeast Asia, and their enslavement of indigenous peoples. Though interesting (and quite important), addressing this is beyond the scope of this paper.

2.2 A Brief Linguistic Overview of Chinese

There are today 921.2 million native speakers of Chinese languages, far exceeding all other languages in this statistic (Eberhard et al. 2021). It is largely agreed that there are seven mutually unintelligible Chinese dialect groups: the Mandarin supergroup[10] (including Standard Mandarin Chinese), Wu, Xiang, Gan, Hakka, Yue (including Cantonese), and Min,[11] developing from ~10th century Middle Chinese[12] in a similar way that Romance languages developed from Latin (Thurgood 2003). Traits shared across the Chinese languages include being tonal and syllable-timed, and having very little inflectional morphology (Szeto et al. 2018). Chinese logographic orthography reflects the shared morpho-syllabic structure across Chinese languages: each syllable has necessarily a tone contour[13] and a monophthong or diphthong vowel nucleus; and optionally initial (consonant and/or glide) and/or coda (either consonant or glide) segments (Chen 1976). Of the analyzed sound changes, *h* vs. *f* and affricate palatalization pertain to the syllable initial, while *n* vs. *m* to the coda. There is a large overlap between the phonemic inventory of Cantonese and Mandarin consonants (see Appendix A), the most notable difference being that Cantonese's phonemic inventory lacks the retroflexes found in Mandarin.

An important factor of Chinese linguistic history is that there has been a relatively continuous thread of hereditary monarchical regimes known as dynasties, from prehistoric times up until 1912. Individual dynasties lasted anywhere between a couple of decades to multiple centuries (276 years, tied by Ming and Qing dynasties, 1368-1644 and 1636–1912). Long

[10] The Mandarin dialect group is classified as a supergroup given its size; though varieties within the supergroup share many features, there is also substantial linguistic variation. See Szeto et al. (2018).
[11] There is some controversy (often politically motivated) with the referring to Chinese varieties as dialects, or as distinct languages. I follow Szeto et al. (2018)'s lead and refer to these groupings as *dialect groups*.
[12] With the exception of the Min dialect group, which is theorized to have split from Old Chinese. There are certain phonological traits that are believed to be conserved from OC, much like Sardinian for spoken Latin.
[13] Unlike in many European languages where inflection plays a grammatical role, distinct tone contours distinguish different morphemes in Chinese. A common example is that the single syllable morpheme [*ma*] in Mandarin can mean 'mother', 'numbing-spicy', 'horse', or 'to scold', based on which of the four tone contours it is pronounced with.

periods of political stability within a dynasty allowed for territorial expansion, whereas the violent political upheaval between dynasties led to mass refugee emigration. The political capital of these dynasties oscillated between various cities mostly in the plains of North China, where Mandarin varieties were indigenously spoken, and used as the official government language.[14]

Most dynasties were led by families of Han Chinese ethnicity, with two important exceptions: the Yuan dynasty (1271-1368) led by Mongols, and the Qing dynasty (1636-1912) by Manchus. Both dynasties would conduct government business in Mandarin instead of imposing their native, Altaic languages on the large existing population of ethnic Han Chinese (Hashimoto 1986). These century-long periods of native speakers of non-Sinitic varieties governing in Sinitic languages provides a very unique situation of language contact, which is widely believed to have caused, as we will shortly see, various unique sound changes in Northern Chinese varieties, in particular Mandarin.[15]

On the other hand, the Yue varieties, of which nowadays Hong Kong Cantonese stands as the prestige dialect, have their inception around 300 BC, when Han Chinese migrated to the Southern China Coast and encountered indigenous speakers of non-Chinese Tai varieties. In particular, there was a large migratory wave during the violent political upheaval during the Tang dynasty (705-907) (Yue-Hashimoto 1991). Due to the difficult, mountainous terrain in South China, there were not any additional waves of mass migration between the Southern China Coast and the rest of China until the 20th century. Intermarriage and cultural exchange between Han

[14] The word Mandarin has its origins in the Portuguese missionaries' noticing the particular variety of Northern Chinese that was spoken by governmental officials, when they arrived in 1582 during the Ming dynasty. They would call the variety *Mandarim*, having etymological roots in the Portuguese word *mandar* 'to order/command', which itself is derived from MANDĀRE in Latin (same meaning). There is very little evidence of language contact with Portuguese causing sound change in Chinese languages (Zhong 2021).

[15] The language reformations beginning in 1912 that established Beijing Mandarin as the model for Standard Chinese is also of much interest, in particular to more modern sound changes, but is outside of the scope of my investigation.

and Tai ethnic groups, as well as linguistic isolation serve as a backdrop for the sound changes in Cantonese.[16]

3 Investigation

First, I analyzed lexically-related glosses of the Swadesh 207 word list to get a rough idea of the frequency of each phonological difference between Sp vs. Pt and Man vs. Can.[17] Summary tables are given in Appendix B, sorted by frequency of each class of phonological difference. In the Swadesh 207 word list glosses, I was surprised to find that there were zero instances of *-n* vs. *-m* in Sp vs. Pt, cf. 9 instances in Man vs. Can; while for *h* vs. *f*, there were 3 instances of /∅/ vs. /f/ in Sp vs. Pt, cf. 4 instances of /xu/ vs. /f/ in Man vs. Can; and for affricate palatalization, there were 4 cases of /tʃ/ vs. /jt/ in Sp vs. Pt., cf. 45 cases of /tʂ/ vs. /ts/ (and corresponding fricatives) in Man vs. Can.[18] With these rough frequencies in mind, I began my literature search.

3.1 *n* vs. *m*

One of my first findings was that *m*'s at the end of syllables in Portuguese orthography do not correspond to the phoneme /m/, but rather signals nasalization, e.g. *um* [ũ] 'one', *bom* ['bõ] 'good', *embora* [ẽ'bɔɾa] 'although'.[19] The only example of an *n* > *m* change in Portuguese is with the feminine indefinite article *uma*, derived from the Latin word UNAM,[20] which Parkinson explains "developed in the seventeenth century through glide epenthesis *ũa* > *uwa* > *uma*

[16] The mid-20th century rise of Hong Kong is also of much interest, but is beyond the source of this paper.
[17] There is much controversy over the claimed anthropological universality of the lexicon of Swadesh 207. This criticism need not apply, however, as I use this word list to get a rough gauge of kinds of phonological variation.
[18] It is important also to note that between each pair, there are sound changes that do not occur, in the other, e.g. although vowel nasalization was observed 29 times in Sp vs. Pt, it was not observed in Man vs. Can; and although lenition of voiceless stops /-k, -p, -t/ was observed 18 times in Man vs. Can, it was not observed in Pt vs. Sp.
[19] Of course, never having formally studied Portuguese, but rather obtaining nearly all of my knowledge of the language from Brazilian song lyrics, I did not know this.
[20] Compare with glosses in other Romance languages that preserve the intervocalic [n]: *una* (Spanish), *une* (French), *una* (Italian).

possibly influenced by the orthographic form of the masculine, *(h)um*." (1988: 147). This single sporadic example alone does not indicate any systematic sound change.

On the other hand, I learned that it was not Cantonese that underwent *n* > *m*, but rather Mandarin that underwent *m* > *n*. While Cantonese preserves the six codas of Middle Chinese (/-m, -n, -ɲ, -p, -t, -k/, with /-ɲ/ > /-ŋ/), Mandarin, only has two: /-n, -ŋ/[21]. In particular, the MC codas /-m/ and /-n/ regularly merge into /-n/ in the diachronic development of Standard Mandarin (e.g. 心, 新 [*siəm, *siən] > [ɕin], Chen 1976: 211). Unlike the aforementioned Spanish-Portuguese case, there is very robust evidence for this sound change. Zee (1985) illustrates *m* > *n* alongside other processes of coda nasal lenition for various Chinese varieties within and without the Mandarin dialect supergroup, to draw out subtleties of Chen (1972)'s proposal of a nasal lenition cline, roughly:

(1) /-m, -n/ > /-n/; (2) /-n, -ŋ/ > /-ŋ/; (3) /-ŋ/ > /-N/; (4) /-N/ > /Ṽ/; (5) /Ṽ/ > /V/

with Yue and Keijia dialects on the conservative extreme (before 1), followed by Northern Mandarin dialects (1), Wu dialects (2-3), and finally Southwest Mandarin dialects (4-5).

Interestingly, a literature search with the queries "nasal merger" and "bilabial alveolar merger" gives results only for studies of the /m, n/ coda merger in Chinese varieties, suggesting to me that this sound change is extremely rare outside of the Chinese language family. With regards to language contact, LaPolla notes that "Mantaro Hashimoto … suggests that the preservation of final -n and -ŋ in Mandarin while all the stop endings and -m were lost might be due to the fact that these two finals are found in Manchu" (2010: 6862); Hashimoto writes, "when the Yuan empire collapsed and the Qing empire was replaced by the Chinese Nationalists,

[21] Very strikingly, Mandarin has lost all unvoiced stop codas /-p, -t, -k/ from MC (e.g. 拉 [*la**p**] > [la] 'to pull', 不 [*fuə**t**] > [bu] 'one', 北 [*pək] > [pei] 'North'), explaining the corresponding. Like Cantonese, Mandarin has also undergone a /-ɲ/ > /-ŋ/ shift from MC.

neither the Mongolian nor the Manchu population were exterminated or driven back to the northern steppes. They [became] the very northern 'Chinese' people we are talking about" (1986: 95). It seems quite plausible that it was due to the very unique language contact situation with Altaic varieties that the initial /-m, -n/ > /-n/ coda merger occurred at all. It could then be that the rest of the cline developed through more typical processes of nasal lention, as nasal lenitions /Vn, Vŋ/ > /Ṽ/ are extremely common sound changes found throughout the world (Ohala 1974).[22]

3.2 *h* vs. *f*

The /f/ > /h/ sound change in Spanish is very well attested, and has received much scholarly attention (Lloyd 1987; Brown & Raymond 2012; Sala 2013). In modern day Spanish, there are many words that have undergone /fV/ > /hV/ (e.g. *hacer* 'to do', *horno* 'oven'), but just as well many that have not (*fuente* 'fountain', *fantasma* 'ghost'). *Cultismo* borrowings from Latin (see section 2.1) account for the reintroduction of some glosses beginning with *f*; yet, even with this, there are numerous irregularities, e.g. *humo* 'smoke' vs. *fumar* 'to smoke'.[23] Lloyd (1987) discusses the phonemic isolation and phonetic instability of /f/, and gives a phonetic explanation of the sound change, positing the unvoiced bilabial fricative /ɸ/ as intermediate, i.e. /f/ > /ɸ/ > /h/ (before leniting to /Ø/, the modern null pronunciation of *h*). It is certainly very interesting that [ɸ] is an allophone of /f/ in certain modern-day Andean, Palanquero, Caribbean, and Puerto Rican varieties of Spanish (Iacobo 2014).

[22] It's even observed in Portuguese, with 29 instances of [Vn] > [Ṽ] and 3 instances of [Vŋ] > [Ṽ] observed in the Swadesh 207 word list gloss comparison between Sp vs. Pt.
[23] Brown & Raymond (2012) show in an analysis of the 1499 work *La Celestina* that Frequency in a (phonologically) Favorable Context is a fair predictor of whether a given word would undergo *f* > *h*.

A well-known (and perhaps infamous) explanation for the /f/ > /h/ sound change is the Basque Substratum Hypothesis: that there was a superstratum-substratum relationship between Latin and Basque leading to the development of Castilian Spanish, and because Basque did not have in its phonemic inventory the labiodental fricative /f/, its speakers would use /h/ instead. Sala (2013) provides a strong argument against this hypothesis, citing the occurrence of /f/ > /h/ in Sardinian and Calabria in Italy, how borrowed words from Latin into Basque underwent /f/ > /b/ as opposed to /f/ > /h/, and that there is substantial lack of other Basque linguistic features found in Spanish. As a matter of fact, the /f/ > /h/ occurs very often cross-linguistically: Foulkes (1997) catalogs nine cases from genetically unrelated varieties. Ultimately, well-developed phonetic theories and cross-linguistic ubiquity have led linguists to believe that contact with Basque serves at best a complementary role to the internal mechanisms behind the sound change.

On the other hand, between Mandarin and Cantonese, it is Cantonese that has developed /h/ > /f/ from Middle Chinese (Chen & Newman 1984), in the opposite direction of Spanish! Garrett & Johnson give examples of occurrences of /h/ > /f/ of in Middle English > Modern English (e.g. [kou**x**e] > *cou**gh***, [lau**x**e] > *lau**gh***, [rou**x**e] > *rou**gh***, 2013: 25)) and in Old English > Buchan Scots (e.g. [**xw**a:] > [**f**a:] 'who',[24] [**xw**i:t] > [**f**ejt] 'white', 2013: 29). A. Garrett explains that [xw] > [f] is an instance of fusion, a cross-linguistically common articulatory sound change where two separate segments fuse into one; in this case, the labiovelar fricative [f] shares articulatory features of both the velar fricative [x] and labiovelar approximant [w] (personal communication, May 3, 2021). Upon a reexamination of the Swadesh 207 word list, all four examples of *h* vs. *f* between Man and Can are either [xw-] vs. [f-] or [xu-] vs. [f-]. I was not able to find any studies proposing language contact as the source of this sound change, suggesting that it is explainable solely through internal mechanisms.

[24] This is the exact same change in Man vs. Can for 花 [***x****wa*] vs. [***f****a*] 'flower'!

3.3 Affricate Palatalization

Alkire & Rosen classify palatalization as a kind of compromise articulation: "when a gesture aiming at one point of articulation is adjacent to a gesture aiming at a different point of articulation, the two tend to become simultaneous," then giving examples in English: *did you* [dɪ**dʒ**ə], *can't you* [kæ̃**tʃ**ə] (2010: 56). In Spanish, the affricate palatalization sound change [kt], [k], [ks], [lt] > [tʃ] occurs regularly, mediated through the creation of secondary diphthongs [kt] > [jt], e.g. STRĬCTU 'tight' > [estrekto] > [estrejto] > [estretʃo] *estrecho* (Alkire & Rosen 2010: 82-83). It is in the penultimate stage we find the corresponding Portuguese gloss, [estrejtu] *estreito*. Through this regular sound change, all 5 cases of [-tʃ-] vs. [-jt-] in the Swadesh 207 word list are explained. There are excellent phonetic studies of this palalalization; for instance, Baker (2004)'s PhD thesis is entirely on palatal segments in Spanish, and in it he discusses theories of articulation and phonetic detail, the diachronic development of palatals from an Optimality-Theoretic framework, as well an acoustic analysis study he conducts. On the other hand, I was not able to find any studies proposing language contact as a source of affricate palatalization in Spanish. As it turns out, palatalization is one of the most common sound changes in the world (Bateman 2007).[25] In her thesis, Bateman (2007) conducts a cross-linguistic survey of 117 languages from 86 genera, finding that just under half of them (58) show palatalization. It is likely due to the ubiquity of this sound change, amongst other types of palatalizations, that linguists largely believe that the phenomenon is explainable solely through internal sound change mechanisms.

Reconstructions of Early Middle Chinese include unvoiced and voiced retroflex and palatal sibilants [ʂ, ʐ, ɕ, ʑ], and their corresponding affricate counterparts [tʂ, tʂʰ, dʐ, tɕ, tɕʰ, dʑ]

[25] For frication, there are many examples within Romance, e.g. Italian *cento* [tʃɛnto] 'hundred'; Acadian French *tiens* [tʃɛ̃] 'you/I keep'; Galician *cheio* [ˈtʃeo] 'full'. Even in Brazilian Portuguese, there is now a development of affricate palatalization after /t, d/ + front vowel, e.g. *gente* [ˈʒen.tʃi] ‘people’ (Ferreira & Holdt 2014: 133).

(Baxter 1992). Mandarin preserves [ʂ, ʐ, ɕ, ʈʂ, ʈʂʰ, tɕ, tɕʰ] after a series of voiced-unvoiced and fricative-affricate mergers (Chen 1976) while Cantonese has lost all traces of these palatalized initials. There is substantial evidence that this change occurred relatively recently: Bauer (2005) points out that two Cantonese-English dictionaries produced by English missionaries in 1828 and 1856 show distinction between palato-alveolar and alveolar [ʃ] and [s], glosses in these dictionaries with [ʃ] match with Mandarin glosses with [ʂ]. Zee (1996), after an extensive literature review of studies from 1856 to 1990, concludes that this [ʃ, s] > [s] merger was likely complete by the late 1930s.

Interestingly, many southern dialects of Mandarin have also lost retroflex initials (Szeto et al. 2018: 246). As a matter of fact, there is some evidence that this loss of retroflexion is an ongoing sound change in the Changsha dialect of Xiang Chinese.[26] LaPolla cites Chen (1996)'s study on bi-direction diffusion between Chinese and Tai in the Yunnan province of China, "that in some cases there has been a simplification of the sound system, e.g. the loss of the distinctions … between /ts/ and /tʂ/ in the Chinese spoken by ethnic Chinese … In other cases there has been an increase in phonemes due to the influence of loan words in the language, e.g. the development of /kh, tsh, tɕ, h/ in the Tai of Luxi county" (2010: 6863).[27]

The [ʃ, s] > [s] merger seems to be cross-linguistically rare; like in the case of the [-n, -m] > [-n] merger in Mandarin, I struggled to find literature on this sound change outside of Chinese languages. It is curious that despite centuries of contact between Tai and Yue languages, this sound change occurred only in the late 1800s/early 1900s, though there are some preliminary signs that the more recent language contact between Yue and Hakka Chinese varieties in Hong Kong, the latter not having palatalized initials, could be the source of this merger; before mass

[26] Unfortunately, there does not seem to be much linguistic focus on this ongoing sound change, at least not in English. See Michaelyus (2013)'s forum post.
[27] Unfortunately, I cannot follow this citation, as I do not read Chinese.

Yue migration to Hong Kong from Canton in the 19th and 20th centuries, the island was occupied primarily by speakers of Hakka varieties (Fat 2005).

4 Discussion and Closing Remarks

It is said that one of the most delightful experiences in research is to arrive at results that entirely contradict your hypothesis and overturn all expectations. As has been made apparent in the above analysis, the phonological parallels that I perceived between the Ibero-Romance and Chinese varieties were surface-level at best. The innovations of all three consonant changes seemingly paralleled between Sp/Pt and Man/Can occurred in opposite directions: /m/ > /n/ sporadically (with just a single instance!) in Pt, while /m/ > /n/ regularly in Man; both /f/ > /h/ and /t/ > /tʃ/ occurred in Spanish, while /h/ > /f/ and /ʃ/ > /s/ occurred in Cantonese.

Addressing now Hinskens's question of Why, concerning the extent to which a given sound change is caused by language contact vs. internal mechanisms, I suggest that the rarer a sound change is observed cross-linguistically, the more weight should be given to hypotheses of language contact; and conversely, for cross-linguistically common sound changes, more weight should be given towards mechanisms of internal sound change although the effects of language contact should not be ruled out in particular with regards to the actuation problem, i.e. what spurs the sound change occurrence at a particular time in the variety's linguistic history. This was made apparent in my difficulty in finding literature arguing for language contact for the sound change /h/ > /f/ in Cantonese, and affricate palatalization in Spanish.

It is interesting to consider that the two sound changes for which a strong case could be made about language contact, are both mergers that seem to be extraordinarily rare: [m, n] > [n] in Mandarin codas, and [ʃ, s] > [s] in Cantonese initials. Hashimoto (1976) expounds on this

hypothesis of the prominence of language contact as the source of much linguistic variation within China, with Altaic and Tai for Northern and Southern Chinese varieties, respectively, and goes on to make an even stronger claim: that all Chinese varieties exist on a dialect continuum with Tai anchoring the gradient in the South, Altaic in the North; and as one travels northward from South China, linguistic features of the encountered varieties shift from being Tai-like to being Altaic-like.[28] He provides substantial evidence to his claim by showing that the phonological, morphological, and syntactic structures of multiple Chinese varieties correlate strongly (although not without exception) with their north-south spatial location, and posits a wave theory-esque explanation for the gradient of linguistic features, i.e. there is substantial contact not only with Tai and Altaic at the ends, but also within Chinese varieties that cause to the diffusion of linguistic features, amounting to the observed gradient. His claim would provide a very strong framework to explain the palatalization (and loss thereof) sound changes discussed in section 3.3, though I would have to look into his claims more before deciding for myself.

A dimension that I was not able to include in this current study is that of sociolinguistic variation, which is believed to play just as important of a role in sound change as does phonetics (personal communication with A. Garrett, May 3, 2021). Although I am absolutely certain that there is a lot to be gained from considering the sociolinguistic dimension for the sound changes I analyzed, I would like to redirect this attention towards future studies in the synchronic domain. In so many ways, we live in unprecedented times: The Covid-19 pandemic has led to a great slow-down in migration; as the pandemic slowly ends, there are unique opportunities to study the effects of language contact on sound change from restarting migratory patterns. On the other

[28] This is very reminiscent of the heat equation in physics, in particular the solution with fixed temperature boundary conditions, e.g. for a metal rod with its ends connected to thermal baths of temperature T_1 and T_2, no matter what initial temperature distribution across rod (i.e. the rod may initially be very hot at certain points, very cold in others) or irregularities in kind of metal and metal alloys the rod is made of, with enough time it will settle to a unique steady state with a continuous temperature gradient between T_1 and T_2.

hand, there are nascent studies of the innovation and usage of gender-inclusion language,[29] which may provide much elucidation on the very intriguing interactions between political ideology and language usage/adoption. Finally, the socioeconomic rise of Mainland China, political unrest in Hong Kong, and strained international relations between East and West provide for an excellent testing ground for the sociolinguistics effects that political alignment and language attitude have on sound change within and without Chinese languages.

Moving forward, contrastive studies are important to distinguish the roles that these diverse factors of sound change play. Because language is such an extraordinarily multifaceted phenomenon, involving simultaneous interactions between acoustics, human cognition, and social dynamics, a theory of language change, even just at the level of phonology, necessarily needs to incorporate factors in multiple levels. Hinskens (2020) proposes a compelling holistic model for sound change, and it is my belief that contrastive studies between genetically unrelated languages may play a crucial role in elucidating fundamental principles.

[29] For instance Papadopoulos et al. (2021) compares and contrasts the gendered nature and gender-inclusive innovations in Spanish and Mandarin Chinese. Licata (2019) examines how listeners' attitudes towards the usage of the local Genoese vs. standard Italian varieties in Liguria, Italy is conditioned on the speaker's gender.

Acknowledgements

I would like to thank Andrew Garrett for insightful conversations. A special thanks goes out to Rachel Weiher for providing helpful feedback on an earlier draft of this paper. And last, but certainly not least, I would like to give a very (*vraiment!*) big thank you to Mairi-Louise McLaughlin, for organizing and teaching an extraordinary iteration of the c202 Linguistic History of the Romance Language graduate seminar. I cannot possibly fathom conducting this linguistics investigation to the extent that I have without her guidance and support.

Bibliography

Alkire, T., & Rosen, C. (2010). *Romance languages: A historical introduction*. Cambridge University Press.

Appendix:Portuguese Swadesh list. (2021, April 15). In *Wiktionary*. https://en.wiktionary.org/wiki/Appendix:Portuguese_Swadesh_list

Appendix:Spanish Swadesh list. (2021, April 15). In *Wiktionary*. https://en.wiktionary.org/wiki/Appendix:Spanish_Swadesh_list

Baker, G. K. (2004). *Palatal phenomena in Spanish phonology*. University of Florida.

Bateman, N. (2007). *A crosslinguistic investigation of palatalization* (Doctoral dissertation, Univ. of California San Diego).

Bauer, R. S. (2005). Two 19th century missionaries' contributions to historical Cantonese phonology.

Baxter, William H. (1992), *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter, ISBN 978-3-11-012324-1.

Bossong, G (2016). In *The Oxford guide to the Romance languages* (Vol. 1). Oxford University Press, 63-72.

Brown, E. L., & Raymond, W. D. (2012). How discourse context shapes the lexicon: Explaining the distribution of Spanish f-/h-words. *Diachronica*, *29*(2), 139-161.

Cantonese Swadesh list. (2021, April 15). In *Panglossa*. https://panglossa.fandom.com/wiki/Cantonese_Swadesh_list

Chen Baoya. (1996). *Yuyan jiechu yu yuyan lianmeng (Language contact and language coalescence)*. Beijing: Yuwen Chubanshe.

Chen, M. Y. (1972). *Nasals and nasalization in Chinese: explorations in phonological universals*. University of California, Berkeley.

Chen, M. Y. (1976). From middle Chinese to modern Peking. *Journal of Chinese Linguistics*, 113-277.

Chen, M. Y., & Newman, J. (1984). From Middle Chinese to Modern Cantonese (Part 1). *Journal of Chinese Linguistics*, *12*(1), 148-198.

EasyPronunciation [Computer software]. (2021). Retrieved from https://easypronunciation.com/en/cantonese-jyutping-phonetic-transcription-converter

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2021. *Ethnologue: Languages of the World.* Twenty-fourth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Fat, L. C. (2005). A dialect murders another dialect: the case of Hakka in Hong Kong.

Ferreira, L., & Holt, D. E. (2014). On the partially divergent phonology of Spanish, Portuguese and points in between. *Portuguese-Spanish interfaces: Diachrony, synchrony, and contact,* 123-150.

Foulkes, P. (1997). Historical laboratory phonology — Investigating /p/ > /f /> /h/ changes. *Language and Speech*, 40(3), 249-276.

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. *Origins of sound change: Approaches to phonologization*, 51, 97.

Hashimoto, M. J. (1976). 'Language diffusion on the Asian continent', *Computational Analyses of Asian and African Languages* 3.49-63.

Hashimoto, M. J. (1986). The altaicization of northern Chinese. *Cornell Linguistic Contribution*, *5*, 76-97.

Hinskens, F. (2020). The expanding universe of the study of sound change. *The Handbook of Historical Linguistics*, 2, 5-46.

Hu, L. (2017, February 25). 二. *Mandarin pronunciation vs. other dialects and Japanese*. StackExchange. https://chinese.stackexchange.com/a/22901

Iacobo. (2014, June 16). *Any Spanish speech variety where F is pronounced as ɸ?* StackExchange. https://linguistics.stackexchange.com/a/26769

LaPolla, R. J. (2010). Language contact and language change in the history of the sinitic languages. Procedia-Social and Behavioral Sciences, 2(5), 6858-6868.

Licata, G. (2019). Language Attitudes in Liguria: Effects of Gender on the Perception of Genoese. 49th Linguistic Symposium on Romance Languages, University of Georgia. May 1–4, 2019.

Lloyd, P. M. (1987). *From Latin to Spanish: Historical phonology and morphology of the Spanish language* (Vol. 173). American Philosophical Society.

Mandarin Swadesh list. (2021, April 15). In *Panglossa*. https://panglossa.fandom.com/wiki/Mandarin_Swadesh_list

Michaelyus. (2013, February 20). *Taking the Changsha dialect as an example of Xiang Chinese / Hunanese, the loss of the retroflex consonants is in process* [Online forum post]. Chinese-forums. https://www.chinese-forums.com/forums/topic/40068-a-merging-of-the-alveolar-and-retroflex-initials-in-the-sinitic-language-family/?tab=comments#comment-301913

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In: C. A. Ferguson, L. M. Hyman, & J. J. Ohala (eds.), *Nasálfest: Papers from a symposium on nasals and nasalization*. Stanford: 559 Language Universals Project. 289-316.

Papadopoulos, B., Bedin, C., Duran, J., Duarte, J., Fliege, C., Tang, C., & Yi, I. (2021). Gender in Language Project. https://www.genderinlanguage.com/gender-in-language-project

Parkinson, S. (1988). Portuguese. *The Romance Languages*, 131-169.

Sala, M. (2013). Contact and borrowing. *The Cambridge history of the romance languages,* 2, 187-236.

Szeto, P. Y., Ansaldo, U., & Matthews, S. (2018). Typological variation across Mandarin dialects: An areal perspective with a quantitative approach. *Linguistic Typology*, *22*(2), 233-275.

Thurgood, G. (2003). A subgrouping of the Sino-Tibetan languages: The interaction between language contact, change, and inheritance. *The Sino-Tibetan Languages*, 3, 21.

Weinberger, Steven. (2015). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu

Yue-Hashimoto, A. (1991). The Yue dialect. *Languages and Dialects of China* (Journal of Chinese Linguistics Monograph Series Number 3, William S. Wang ed.), pp. 294-324.

Zee, E. (1985). Sound change in syllable final nasal consonants in Chinese/汉语鼻音韵尾的音变. *Journal of Chinese linguistics*, 291-330.

Zee, E. (1996). Phonological Changes in Hong Kong Cantonese. *Current Issues in Language & Society*, *3*(2), 192-198.

Zhong, A. (2021). Macanese Patúa. Presentation for c202 Linguistic History of the Romance Languages graduate seminar.

Appendix A: Phonemic Inventory of Consonants, adapted from Weinberg (2015).

Table A.1 Spanish

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular
Plosive	p			t				k g	
Nasal	m			n			ɲ		
Trill				r					
Tap or Flap				ɾ					
Fricative	β	f	θ ð	s				x ɣ	
Affricate					tʃ				
Lateral fricative									
Approximant							j		
Lateral approximant				l			ʎ		

Table A.2 Portuguese

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular
Plosive	p b		t d					k g	
Nasal	m		n				ɲ		
Trill									ʀ
Tap or Flap									
Fricative		f v		s z	ʃ ʒ				
Affricate									
Lateral fricative									
Approximant			ɹ				j		
Lateral approximant			l				ʎ		

Table A.3 Mandarin

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular
Plosive	p			t				k	
Nasal	m			n				ŋ	
Trill									
Tap or Flap									
Fricative		f		s		ʂ ʐ	ɕ		χ
Affricate				ts		tʂ	tɕ		
Lateral fricative									
Approximant							j		
Lateral approximant				l					

Table A.4 Cantonese

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p			t				k			
Nasal	m			n							
Trill											
Tap or Flap											
Fricative		f		s							h
Affricate				ts							
Lateral fricative											
Approximant							j				
Lateral approximant				l							

Appendix B: Phonological differences in Swadesh 207 word list cognates, sorted by frequency

Table B.1 Non-allophonic sound changes between Spanish and Portuguese (European), adapted from Wiktionary ("Appendix:Spanish Swadesh", "Appendix:Portuguese Swadesh").[1]

Categorical Difference		Specific Detail	Example from Swadesh 207	
Category	**#**	**IPA (# instances)**	**Swadesh #. English Gloss**	**Sp vs. Pt Glosses Orthography and IPA**
Nasalization (Pt)	29	[Vn, Vŋ, Vm] vs. [Ṽ] (29, 3, 1)	14. when	cuándo [ˈkw**a**ndo] vs. quando [ˈkw**ɐ̃**du]
Fricative Velarization (Sp)	13	[x] vs. [ʎ] (8)	36. woman	mujer [muˈ**x**ɛɾ] vs. mulher [muˈ**ʎ**ɛɾ]
		[x] vs. [ʒ] (3)	126. to turn	girar [**x**iˈɾaɾ] vs. girar [**ʒ**iˈɾaɾ]
		[x] vs. [s] (1)	46. bird	pájaro [ˈpa**x**aɾo] vs. pássaro [ˈpa**s**ɐɾu]
		[x] vs. [ʁ] (1)	135. to push	empujar [empuˈ**x**aɾ] vs. empurrar [ẽpuˈ**ʁ**aɾ]
Lateralization (Sp)	8	[ʎ] vs. [l] (4)	149. star	estrella [esˈtɾe**ʎ**a] vs. [iʃˈtɾe**l**ɐ]
		[ɲ] vs. [n] (2)[2]	179. year	año [ˈa**ɲ**o] vs. ano [ˈɐ**n**u]
		[ʎ] vs. [ʃ] (2)	151. rain	lluvia [ˈ**ʎ**uβja] vs. chuva [ˈ**ʃ**uvɐ]
Fricativization (Sp)	13	[ð] vs. [d] (7)	156. stone	piedra [ˈpje**ð**ɾa] vs. pedra [ˈpɛ**d**ɾɐ]
		[ɣ] vs. [g] (6)	167. fire	fuego [ˈfwe**ɣ**o] vs. fogo [ˈfo**g**u]
Intervocalic Alveolar Lenition (Pt)	6	[l (3), n (3)] vs. [∅]	148. moon	luna [ˈlu**n**a] vs. lua [ˈluɐ]
Fricative Palatalization (Sp)	5	[tʃ] vs. [jt] (5)	177. night	noche [ˈno**tʃ**e] vs. noite [ˈnɔ**jt**i]
Initial Fricative Lenition (Sp)	4	[∅] vs. [f] (3)[3]	166. smoke	humo [ˈumo] vs. fumo [ˈ**f**umu]
		[∅] vs. [ʒ] (1)	145. freeze	helar [eˈlaɾ] vs. gelar [**ʒ**iˈlaɾ]

Notes:

[1] There were 33 cases of allophones of /s/: [ʃ, z, s] vs. [s, θ]; 18 of /b/: [v, b] vs. [b, β], and 11 of /r/: [r, ʁ] vs. [l, ɾ, ɾ].

[2] Interestingly, there was one instance of [n] vs. [ɲ]: 121. to walk, caminar [kamiˈnaɾ] vs. caminhar [kɐmiˈɲaɾ].

[3] There were two instances of [fV] vs. [fV], and six instances of [fC] vs. [fC].

Table B.2 Sound changes between Mandarin and Cantonese, adapted from Panglossa Wiki ("Mandarin Swadesh", "Cantonese Swadesh"). IPA from EasyPronunciation (2021).

Categorical Difference		Specific Phonological Detail	Example from Swadesh 207	
Category	**#**	**IPA (# instances)**	**Swadesh #. English Gloss**	**Can vs. Man Glosses Orthography, IPA (sans tone)**
Retroflex Palatalization (Man)	87	[ʂ, ʈʂ, ʈʂʰ] vs. [s, ts, tsʰ] vs. (45)	20. few	少, [**ʂ**ɑʊ] vs. [**s**iu]
		[s, ts, tsʰ, k, kʰ, h] vs. [ɕ, tɕ, tɕʰ] (35)	90. heart	心, [**ɕ**ɪn] vs. [**s**ɐm].
		[j, s, ts, tsʰ] vs. [ʐ] (7)	63. meat	肉, [**ʐ**ʊu] vs. [**j**ʊk]
Init Fricative Velarization (Man)	21	[x-] vs. [h-] (11)	185. good	好, [**x**ɑʊ] vs. [**h**ou]
		[xu-] vs. [w-] (7)	108. To leave	活, [**x**uɔ] vs. [**w**ut]
		[xu-] vs. [f-] (4)[1]	59. flower	花, [**x**ua] vs. [**f**a]
Final Stop Lenition (Man)	18	[∅] vs. [-k, -p, -t] (6 each)	22. one	一, [i] vs. [ja**t**]
Final /m/ (Can) vs. /n/ (Man)	9	[-n] vs. [-m] (9)[2]	52. 森林 (forest)	森林, [sə**n** lɪ**n**] vs. [sɐ**m** lɐ**m**]
Init Nasal Lenition (Man)	8	[u-, w-] vs. [ŋ-, m-] (8)	1. 我 (I)	我, [**u**ɔ] vs. [**ŋ**ɔ]
/aɻ/ (Man) vs /ji/ (Can)	3	[aɻ] vs. [ji] (3)[3]	23. 二 (two)	二, [**aɻ**] vs. [**ji**]

Notes:

[1] There were four instances of conserved [f] vs. [f].

[2] There were 44 instances of conserved [-n] vs. [-n].

[3] Surprisingly, all three examples of /aɻ/ vs. /ji/ have completely different tones (˥˩ vs. ˨˩, ˧˥ *vs.* ˨˩, ˩˩ *vs.* ˧˥) and semantic contexts (23. 二 'two', 39. 'child', and 73. 'ear'). See Hu (2017)'s StackExchange answer.