Open Library of Humanities

Modelling opacity and variation in Gran Canarian Spanish apocope

Karolina Broś, University of Warsaw, PL, k.bros@uw.edu.pl

Aleksei Ioulevitch Nazarov, Utrecht University, NL, a.i.nazarov@uu.nl

In this paper, we present novel data from Spanish spoken on Gran Canaria which show an interaction of two lenition processes: final consonant deletion and vowel apocope. We show that in certain positions in an utterance, these processes optionally combine in a fed counterfeeding interaction. Furthermore, the variation present in the dialect due to process optionality uncovers a *latent opacity* pattern, i.e. an additional opaque interaction that is not motivated by any individual input-output mapping, but only by the quantitative aspect of variation. This takes the form of mutual counterfeeding, a rarely reported phenomenon, thus creating a novel test case for theories of opacity. The second part of the paper provides a formal analysis of the opacity-ridden data, taking into account process optionality and variation. The theoretical analysis and learning simulations using the Expectation-Driven Learner demonstrate that a probabilistic variant of Serial Markedness Reduction can capture both fed counterfeeding and *latent opacity* without recourse to additional mechanisms beyond the original framework, as opposed to analyses in other serial frameworks such as OT-CC. Our analysis shows that opacity with optional processes is a complex problem that has to be specifically addressed with probabilistic frameworks.

1. Introduction

In this paper, we present a previously unreported case of incipient[1] language change that is currently taking place in the Spanish of the Canary Islands. As will be shown in the subsequent sections, two lenition processes identified in our data – consonant deletion and vowel apocope – lead to surface forms that show opaque interactions. At the same time, at least one of the two processes is optional. Thus, the same underlying forms lead to variable outputs, some of which in themselves are difficult to generate in a constraint-based framework.

Furthermore, the kind of opaque interactions we observe is quite complex. First, the data show an instance of fed counterfeeding (Kavitskaya & Staroverov 2010), i.e. a situation in which one process first feeds another process and is then counterfed by it (in our case, final consonant deletion feeds vowel apocope, but is then counterfed by it: /VCVC#/ → VCV# → [VC#] *→ V#). Second, an additional opaque interaction is revealed only when looking quantitatively at variation. More specifically, the probability of vowel apocope is substantially different depending on whether the vowel is word-final in the underlying form (/VCV#/ → [VC#]), or whether consonant deletion has applied to make the vowel word-final (/VCVC#/ → VCV# → [VC#]). As we will explain in §3.2, accounting for this quantitative variation requires an additional opaque pattern in which consonant deletion is counterfed by apocope (/VCV#/ → [VC#] *→ V#) while apocope is counterfed by consonant deletion (/VCVC#/ → [VCV#] *→ VC#): a mutual counterfeeding rather than a fed counterfeeding interaction (Wolf 2011). Interestingly, we can generate each of the individual surface forms without assuming this additional opaque pattern, but we would not be able to capture their relative frequencies. We refer to this phenomenon, in which an opaque interaction is motivated purely by quantitative patterns in the data, as *latent opacity*.

As we will show in this paper, the mutual counterfeeding effect forms a particularly interesting challenge for current formal frameworks, which prompts us to pursue an analysis that takes variation into account. The main questions we would like to answer with our data and analysis are i) whether (and how) surface variation driven by process optionality can be captured using generative frameworks, ii) what the implications of variation are for the opaque processes analysed, and iii) whether it is possible for fed counterfeeding and mutual counterfeeding to be analysed using the same mechanisms as regular counterfeeding.

The paper is structured as follows. In §2, we present the dialect and the data, including a quantitative analysis of productions made by 18 native speakers. In §3, we discuss the opacity effects in the data and provide a formal analysis using Serial Markedness Reduction (SMR, Jarosz

---

[1] In this paper, we refer to an incipient change in the sense of a process that seems to be 'new' in the dialect and is both optional and highly restricted in terms of the environments in which it applies and the population in which it is observed. These restrictions are described in detail in §2.

2014). In §4 we use a learning implementation of the SMR framework to see whether an optimal probabilistic grammar can be found that accounts for the available quantitative data. In §5, we discuss some implications of our analysis for modelling variation with opacity, and consider alternative analyses. §6 concludes the paper.

2. Data

In this paper, we are interested in the interaction of two processes taking place in the Spanish of the Canary Islands. More specifically, we focus on one area: the northern part of Gran Canaria. The data presented come from 18 speakers of the dialect collected in 2016 on Gran Canaria, in the course of semi-structured interviews, using a Zoom H4N digital recorder and a Shure SM10a headworn microphone. Prior to the analysis, the data were transcribed using automatic alignment (EasyAlign, Goldman 2011) and then realigned manually in Praat (Boersma & Weenink 2019) by three annotators.[2]

The processes of interest can be classified as instances of lenition (i.e. sound weakening), one of which is more prevalent than the other, both in terms of phonological context and in terms of its sociolinguistic profile. The first process, consonant weakening, is widespread in the whole speech community (see §2.1), while the other – vowel apocope – is only incipient, and occurs in specific positions only and usually in the speech of younger males (§2.2). Thus, to provide a quantitative analysis of vowel apocope and its interaction with consonant deletion (§2.3), we looked at the speech of young and middle-aged male speakers. The 10 main recordings analysed in this section correspond to young males aged 18–25. A further 8 pieces of data were taken from males aged 37–59 and served as comparison. In the subsequent sections, we explain the two reported processes and provide examples from the corpus. This is followed by a detailed explanation of the environments in which they occur and by the quantitative analysis which informs us on their exact rates of occurrence. The surface distributions that result from the quantitative analysis will then serve as the basis for the formal analysis.

1.1. Consonant weakening in Gran Canarian Spanish

According to the literature, Spanish as spoken on Gran Canaria is well-known for multiple weakening processes and frequent consonant elisions (Alvar 1972; Oftedal 1985; Almeida & Díaz Alayón 1988). While syllable- and word-final consonant deletions are well-extended in rural areas, in urban communities and certain geographical areas of the island weakening without deletion, especially *s* aspiration, is the dominant output. Our data are in line with these general

---

[2] The dataset forms part of a bigger corpus gathered in 2016, encompassing a total of 111,317 phones produced by 44 native speakers of the dialect. The corpus is described in detail in Broś (2022) and samples are available online at www.karolinabros.eu. For the purposes of this paper, only the speech of young and middle-aged males was analysed.

observations as they show syllable-final and word-final consonant weakening in spontaneous speech. One of the outcomes of this weakening is full elision, examples of which are presented in words taken directly from the collected corpus in (1).

(1)   Word-final consonant deletion in the Gáldar dialect[3]
        cosas   /kosas/   [ˈko.sa]   'things'
        hacer   /aseɾ/   [a.ˈse]   'to do'
        papel   /papel/   [pa.ˈpe]   'paper'
        canción   /kansjon/   [ka(n).ˈsjo]   'song'

Our data show that consonant deletion is prevalent in most speakers. More specifically, the representatives of the (sub)dialect tend to delete most word-final consonants unless they can resyllabify them into the onset of the following syllable (e.g. *por aquí* /poɾ#aki/ [po.ɾa.ˈki] '(around) here'), although in many cases these segments are deleted anyway and the following onsets remain unrepaired (e.g. *montamos un panel* /montamos#un#panel/ [mon.ˈta.**mo.u**m.pa.ˈne(l)] 'we assembled a panel'). As for the scope of application, consonant deletion is well advanced and has narrowed down to the word domain. It applies (variably) whenever there is a word-final consonant, independently of bigger constituents such as phrases or sentences. However, it is much more frequent phrase-finally. To provide some numbers, while the rate of consonant deletion in word-final position in general is slightly higher than 50% (55% in our young speakers), it rises to over 90% phrase-finally, which is the position we will focus on in this paper. Furthermore, it must be noted that the process applies regardless of age or gender, the only difference lying in the relative rates of application vis à vis other forms of weakening, e.g. debuccalisation of /s/ to [h] or [ɦ].[4]

1.2. Vowel apocope in Gran Canarian Spanish

The second process of concern in this paper is vowel apocope, i.e. the deletion of word-final unstressed vowels. Some examples from our corpus are provided in (2).

(2)   Apocope (deletion of final unstressed vowels) in the Gáldar dialect
        cosa   /kosa/   [ˈkos]   'thing'
        Tenerife   /teneɾife/   [te.ne.ˈɾif] 'Tenerife'
        eso   /eso/   [ˈes]   'this'
        perfecto   /peɾfekto/   [peɾ.ˈfekt]   'perfect'

---

[3] Here and elsewhere in the paper, underlying representations are given in slashes. Note that all final consonants can undergo elision, but other forms of weakening such as devoicing and fricativisation, gliding or velarisation can also occur. We will not go into any further detail here, as these processes are not the subject of the formal analysis.

[4] We do not pursue this question further as it is outside the scope of the paper. For a sociophonetic analysis of consonant weakening in the dialect, see Broś (2022).

It must be stressed that vowel apocope has **not** been reported in the literature on the dialect to date. Thus, this paper presents novel data. To provide full information on the contexts of occurrence of apocope and generalisations, we had to take a closer look at the corpus, both quantitatively and qualitatively. We list the results of this inquiry below.

First, as shown by the examples in (2), word-final vowels undergo deletion, which results in word-final codas, regardless of the number of consonants (cf. *perfecto*). Crucially, stressed vowels are not affected by the process. It does not apply in words such as *papá* 'daddy'. Additionally, it is worth noting that vowel apocope applies only to **final** unstressed vowels. For instance, in the word *ofertas* 'offers' the initial unstressed vowel is retained as it occupies a strong (initial) position. Similarly, unstressed vowels in words with antepenultimate stress, such as *pájaro* / paxaɾo/ 'bird' are retained beyond the final syllable ([ˈpa.fiaɾ] and not *[ˈpafiɾ] or *[ˈpafi.ɾo]). Furthermore, we determined that there is usually no apocope in monosyllables, even if they are function words, and that apocope seems to be less often applied in verb forms (e.g. *se negaba* / se#negaba/ [se.ne.ˈɣa.(β)a] '(s)he was denying', *te enteras* /te#enteras/ [ten.ˈte.ɾa] 'you find out', *estuve* /estube/ [eh.ˈtu.βe] 'I was'). It is also often blocked or hidden by other processes, e.g. intervocalic stop deletion and the resultant vowel merger/simplification: *nada* /nada/ [ˈna] 'nothing', *lesionado* /lesionado/ [le.sjo.ˈna] 'injured', *relajado* /relaxado/ [re.la.ˈfia] 'relaxed'.[5]

The above description provides a general picture of vowel apocope in the dialect. The process has further restrictions, however. Importantly, it is not word- but phrase-final, and dependent on information load and intonation.[6] For instance, when information is incomplete and an explanation or a second part of the message follows, there is no apocope. The same applies to hesitations and incomplete sentences. Such phrases are characterised by rising intonation and often final vowel or syllable lengthening (possibly an intonational boundary process). When information is completed, the phrase or sentence is finished and the intonation is level or falling, vowel apocope occurs. Some examples of phrases containing the context necessary for apocope to occur, as well as examples of phrases excluded from the count are provided in Appendix 1.

Two more observations should be mentioned. First, vowel apocope may be incomplete – whereas numerous cases of full elision can be found in the data, in some cases the vowel is fully devoiced and some remnant of it is still present in the signal (see **Figure 1**).

---

[5] Also, note that intervocalic stop lenition is, possibly, one of the reasons why verbs behave differently than other words, cf. *se negaba* [se.ne.ˈɣa.(β)a] 'he/she was denying' and other words in which the intervocalic /b/ is either realised as a very weak approximant or, most often, deleted and the flanking vowels are merged as one long stressed vowel.

[6] Given the specific nature of apocope, this had to be determined manually, by listening to the recordings and inspecting the spectrograms.
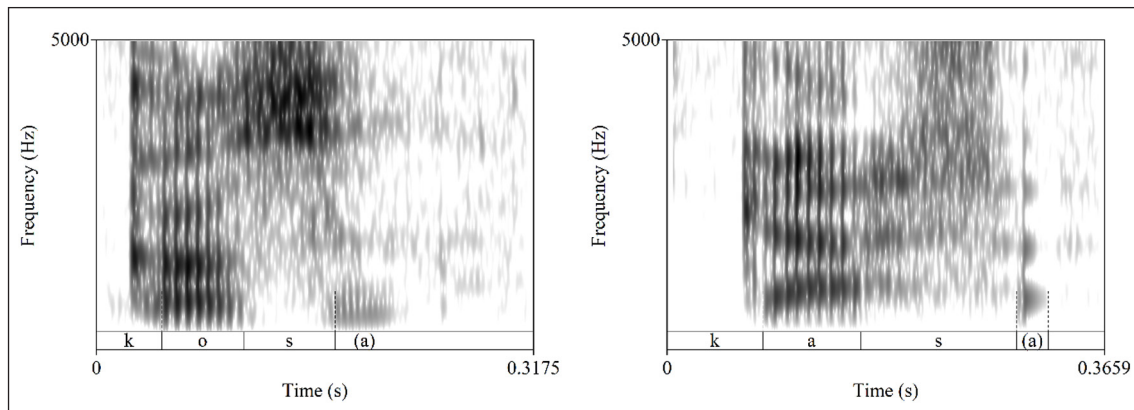
Figure 1: Spectrograms of the words *cosas* 'things' (left) and *casa* 'house' (right) showing incomplete vowel deletion: a voicing trail without a formant structure and complete devoicing (with remnants of formants present), respectively.

Second, as for the final consonants left after apocope applies, there seems to be some sort of emphatic strengthening. For instance, the word *curioso* /kuɾioso/ 'curious' is reduced to [gu.ˈrjos] with a lengthened [s], while in the words *ofertas* /ofeɾtas/ 'offers' > [o.ˈfeɾt] or *gente* /xente/ 'people' > [ˈhent] there is a strong plosion with aspiration on the [t] despite the fact that stops are usually produced with a weak plosion or no plosion at all in this dialect (Broś & Lipowska 2019) and Spanish in general has no stop aspiration (see the spectrograms in **Figure 2**).
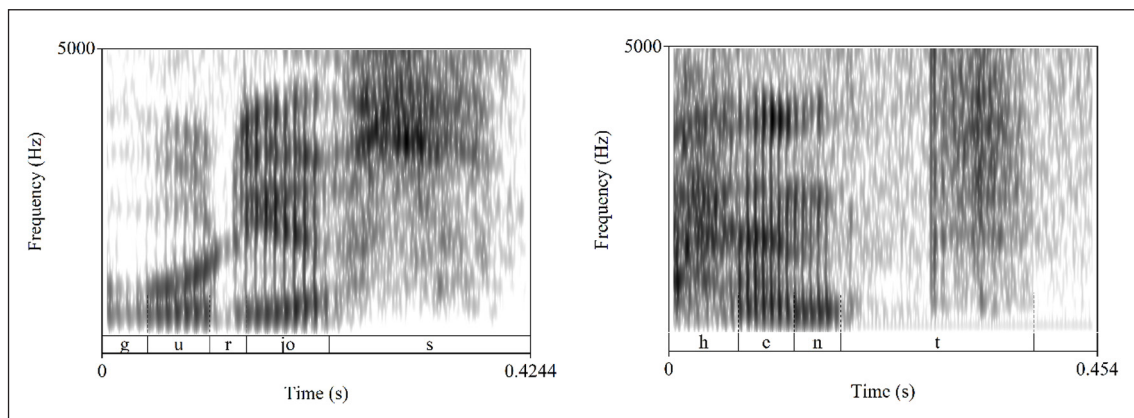
Figure 2: Spectrograms of the words *curioso* 'interesting' (left) and *gente* 'people' (right) showing full vowel apocope with the strengthening of the resultant final consonants.

Given the above we can conclude that the process of vowel apocope is only incipient as it applies optionally and only in the outer domains of prosodic structure. In our data, vowels tend to be dropped at the end of an intonational phrase. In some cases, vowel deletion is incomplete, which means that some parts of the signal are still present and visible on a spectrogram, as

shown in **Figure 1**. In most cases, however, we have full deletion (as in **Figure 2**), i.e. unstressed post-tonic vowels are removed, and the loss of the whole final rhyme tends to be accompanied by some degree of strengthening of the resultant final segment.[7] All in all, the observed changes seem to be driven by ongoing generalised lenition typical of the Gran Canarian dialect.[8]

Another important aspect of vowel apocope in the dialect is that it is socially restricted. As already mentioned at the beginning of §2, it seems to be produced almost exclusively by young male speakers. It can be found to some extent in the speech of middle-aged inhabitants of the island but it does not occur in older speakers, and it can occasionally take the form of vowel devoicing and shortening in some females. The latter, however, is not as yet systematic.[9]

In order to provide the most reliable quantitative data possible, we should pursue the age-related differences in the application of vowel apocope further. Besides, the age factor provides additional evidence that apocope is an incipient process. More specifically, when comparing young speakers (10 speakers aged 18–25) with the older generation (8 speakers aged 37–59), one can observe a substantial difference in the frequency of apocope but not C deletion. Out of a total of 199 contexts across the middle-aged speakers, 81 show either full or incomplete vowel apocope, while 58 show consonant deletion. In the case of the younger speakers, we counted 192 contexts, with 142 cases of full and incomplete apocope and 56 cases of C deletion. The overall percentage of lenitions in the investigated contexts is 58% in middle-aged speakers, which is substantially less than in the younger age group in which 86% of all final sounds are weakened. At the same time, however, C deletion in C-final words happens 95% of the time in middle-aged speakers, and 92% of the time in the younger population, which means that this process does not

---

[7] Further study is needed to see whether this is a compensatory effect, emphasis, domain-final lengthening (Byrd 2000) or gestural masking (Browman & Goldstein 1990), i.e. the presence of the vowel gesture overlapping with a different gesture, resulting in there being no audible sound.

[8] First, the studied dialect is characterised by a series of lenition processes at different advancement stages depending on phonological and social factors (see e.g. Broś et al. 2021). These include weakening of consonants in intervocalic and syllable-final positions, vowel merger, gliding and many others. Phrase-final consonant deletion is a case in point. Apocope is, in our opinion, another change driven by the tendency to drop weak segments and retain strong prosodic positions such as stressed vowels. Second, several lenition processes interact with other processes, which makes them phonological rather than phonetic. Final /s/, for instance, is resyllabified across word boundaries and voiced before V-initial words. Intervocalic lenition is blocked by the deletion of a preceding consonant. Apocope interacts with word-final stop devoicing, e.g. *haciendo* /asjendo/ [a.ˈsjent] 'doing' or *trabajando* /tɾabaxando/ [tɾa.βa.ˈfiant] 'working' and with intervocalic stop lenition and resultant vowel merger (see fn 6). Were the two analysed processes low-level phonetic phenomena, such effects would not ensue and the numbers we show in Table 1 would not point to either C or V deletion as majority options.

[9] Females seem to have different strategies for emphasis or prosodic marking. Although vowel apocope does seem to happen sporadically in some women, it cannot be reliably counted based on our database. A reviewer suggested that this discrepancy between males and females may mean that the change is led by males, which is potentially relevant to social theories of sound change. While this is an important sociolinguistic observation, contrary to the fact that change is usually led by middle-class women and touching upon the famous Labovian gender paradox (Labov 2001), we cannot undertake a discussion on this issue for reasons of space.

differ depending on the age. The difference between the groups lies in the application of vowel apocope. Here, we divided the words into V-final and C-final since both can have apocope but only the latter ones can undergo C deletion and apocope depends on whether C deletion applied. As a result, only 30% in V-final words and a mere 13% in C-final words undergo apocope in the middle-aged group, compared to 61% and 36%, respectively in the younger age-group.[10] These differences have been tested statistically and are illustrated in **Figure 3**. Two sample *t*-tests run in R (R Core Team 2017) showed a significant difference between the two age groups both in overall apocope application ($t(16) = -4.297$, $p < 0.001$) and in V-final and C-final words separately ($t(16) = -3.738$, $p = 0.002$ and $t(16) = -4.057$, $p = 0.001$, respectively). No statistical difference was found for C deletion ($t(16) = 0.238$, $p = 0.81$).
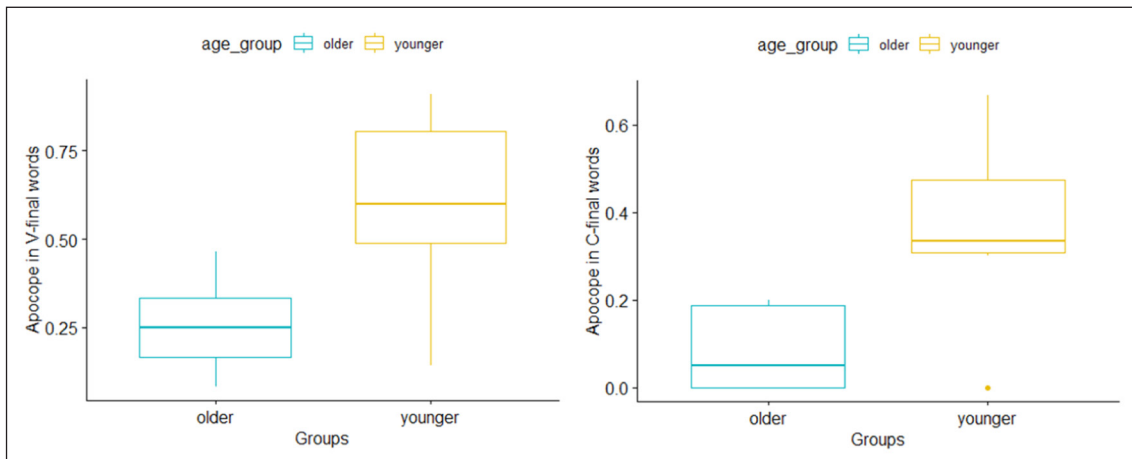

Figure 3: Differences in the proportion of vowel apocope in V-final and C-final words between younger and older speakers.

All in all, given our empirical results, since the occurrence of apocope is much less frequent in older speakers compared to the young ones, we can conclude that vowel apocope is an incipient change, especially given that no instances whatsoever have been detected in the speech of a yet older generation (males or females over 60 years old) in the initial corpus of Gran Canarian speech. The quantitative comparison between the middle-aged and the young speakers suggests that vowel apocope is an ongoing change that seems to be spreading among the young speakers of the dialect. It is therefore this community that we will focus on in the rest of the paper. Most importantly, we will provide data on the rates of consonant and vowel deletion that will be used in the subsequent formal analysis based on the representatives of the young generation.

---

[10] The details of the calculations made for the younger group are provided in Table 1. The interaction of the two processes and the discrepancies between V-final and C-final words are explained in more detail in §2.3. Also, note that the percentages given here refer to full apocope only in order to ensure comparability with the younger group. If we take incomplete apocope into account, middle-aged speakers apply the process 47% of the time in V-final words and 26% of the time in C-final words (cf. Table 1, which shows an overall of 85% and 51%, respectively).

1.3. Consonant deletion and vowel apocope in interaction

Perhaps the most interesting aspect of the processes described in this paper is that they interact in phrase-final position, i.e. where apocope optionally applies. This is illustrated in (3).

(3)   Overlap of consonant deletion and vowel apocope

hijos	/ixos/	[ˈih]	'children'
cosas	/kosas/	[ˈkos]	'things'
los valientes	/los#balientes/	[lo.ba.ˈljent]	'the brave'
metros	/metɾos/	[ˈmetɾ]	'metres'
ofertas	/ofeɾtas/	[o.ˈfeɾt]	'offers'

The data in (3) show that with the two processes operating side by side there are some interesting results. Most often, plural nouns and adjectives will lose the final segment and may additionally lose the resultant word-final vowel under certain circumstances. Naturally, however, not all phrase-final words have a word-final consonant and/or unstressed vowel. If the word is consonant-final but with a stressed syllable, the consonant usually deletes but the vowel is retained. If the word is vowel-final, the stressed vowel is retained as well.

Since neither of the described processes applies 100% of the time, their interaction may lead to a wide range of outcomes. If we look only at the words in which these processes have a chance to apply, i.e. vowel- and consonant-final words with final unstressed vowels, the possibilities are as follows. In vowel-final words there may be no change, incomplete vowel apocope or full vowel apocope. In consonant-final words, there may be no change, consonant deletion only or consonant deletion accompanied by incomplete or full vowel apocope. To illustrate these outcomes with an example, the word *plaza* /plasa/ 'square', an example of a V-final context, and its plural form *plazas* /plasas/ 'squares', an example of a C-final context, are presented in (4) together with a list of possible outcomes.

(4)   Sample outcomes of V-final and C-final forms

a. V-final UR   /plaza/
    no change   [ˈplasa]
    incomplete apocope   [ˈplasḁ]
    full apocope   [ˈplas]

b. C-final UR   /plazas/
    No C deletion   [ˈplasah][11]
    C deletion   [ˈplasa]
    C deletion + incomplete apocope [ˈplasḁ]
    C deletion + full apocope   [ˈplas]

---

[11] Note that when consonant deletion does not apply, the final consonant is usually weakened; /s/ weakens to [h].

Thus, we have several output options among the V-final and C-final words, some of which overlap. For instance, the output [ˈplas] can be a result of full apocope in the word *plaza,* and of C deletion + full apocope in the word *plazas,* etc. Additionally, it is worth underlining that C deletion applies only once, i.e. only to the underlyingly final segment. Whenever vowel apocope 'uncovers' a final consonant, that consonant is not weakened any further. In the word *plaza,* an output form *[ˈpla] is impossible. This is an opaque interaction of the fed counterfeeding type. As the combination of feeding and counterfeeding has proved problematic in formal analyses to date (Kavitskaya & Staroverov 2010), it is important to investigate this complex pattern in formal terms.

As we have seen, the interaction of C deletion and apocope leads to a variety of surface forms, but it is also crucial to look into how often a given form occurs in the dialect. To obtain quantitative data, we turned to the surface sounds used in the speech of the young males (see §2.2 above), as they are the ones who use both processes systematically. The results are presented in **Table 1**, which shows the number of contexts fulfilling the criteria described in §2.2 per speaker, together with the number of vowel apocope instances in each context, including incomplete apocope, and the number of C deletions (when applicable). Thus, phrases containing the context for apocope (both vowel-final and consonant-final words) were selected based on the criteria of information load, intonation and stress, after which we counted instances of deletion that actually occurred. The table shows quantitative results, including percentages.[12]

As can be observed in **Table 1**, the data included a total of 192 contexts of apocope, 131 in vowel-final words (68%) and 61 in consonant-final words. All speakers have both vowel and consonant deletions in most cases. However, the probability of occurrence of vowel apocope changes depending on the word type. In vowel-final words, speakers delete final unstressed vowels 61% of the time. An additional 24% of the words has incomplete apocope, which means that an overwhelming majority of unstressed vowels is weakened in absolute phrase-final position. In consonant-final words, however, only 36% of the tokens exhibit vowel apocope (plus 15% incomplete deletions), and the conditional probability of vowel apocope given that consonant deletion has applied is $22/56 = 39\%$. Under either calculation, full apocope applies in a minority of cases, and even the rate including incomplete apocope is lower than for vowel-final words (probability of full or incomplete apocope given consonant deletion: $31/56 = 55\%$). At the same time, the consonant deletion rate is very high – as many as 92% of final consonants are deleted in these words. This makes consonant deletion without apocope the preferred lenition strategy in consonant-final words.

---

[12] These percentages are calculated out of the total number of V- or C-final tokens, so that the percentages of all variants (no C deletion, C deletion but no apocope, C deletion and apocope) add up to 100%. The probability of occurrence of a process will be discussed in the next paragraph.

Subject (age)	V-final contexts	C-final contexts	V-final apocope	%	C-final apocope	%	C deletion	%
Ccr (25)	12	10	9 (3)	75% (25%)	3	30%	10	100%
Aai (23)	33	9	16 (12)	48% (36%)	5 (3)	55% (33%)	9	100%
Jjo (18)	10	5	5 (1)	50% (10%)	2	40%	5	100%
Ch (24)	7	6	1 (4)	14% (57%)	0 (4)	0% (67%)	6	100%
Ma (24)	11	6	5 (3)	45% (27%)	2 (1)	33% (17%)	5	83%
Mi (23)	10	3	7 (2)	70% (20%)	2	67%	3	100%
Jje (24)	12	10	6 (3)	50% (25%)	3	30%	7	70%
Aal (24)	11	3	10	91%	1 (1)	33% (33%)	3	100%
Jss (25)	8	3	7	88%	1	33%	2	67%
Aar (16)	17	6	14 (3)	82% (18%)	3	50%	6	100%
Totals	**131**	**61**	**80 (31)**	**61% (24%)**	**22 (9)**	**36% (15%)**	**56**	**92%**

Table 1: Vowel apocope and consonant deletion in relevant contexts as produced by young males. Incomplete apocopes are provided in parentheses.

All in all, it can be concluded that although apocope is still not fully established in the language and seems to be restricted mostly to (young) males, it is nevertheless very frequent in phrase-final contexts corresponding to falling intonation (completing information), especially in vowel-final words. At the same time, the very high final consonant deletion rate is important as it suggests a phonological effect, since the mean consonant deletion rate for the same speakers, calculated based on all consonant-final words, regardless of the position in the sentence or phrase, is only 55%.[13] Compared to this much lower rate, consonant deletion is (near-)categorical rather than merely optional in the phrase-final context.[14] Finally, perhaps the most important observation about the data is the difference in the proportion of tokens of apocope between V-final and C-final words. As already mentioned, the numbers are 61% and 36% (39% if conditioned on the occurrence of consonant deletion), respectively. This marked difference in apocope rates between underlyingly V- and C-final words, even when no final consonant is observed, should be accounted for in any formal model of this phenomenon. It will be referred to as a *latent opacity* effect (i.e. dispreference for apocope when followed by a deleted consonant) that is not motivated by any specific surface form, but instead by the frequency distribution between surface forms. We discuss it in detail in §3.2.

3. Formal analysis

In §2, we saw quantitative data from Canary Islands Spanish showing an interesting interaction of two processes of lenition affecting word-final syllables. The general observation is that word-final consonants are systematically deleted regardless of the nature of the preceding segment, whereas final vowels are elided only when unstressed (apocope) and in an appropriate pragmatic/intonational context. As a result, in many cases, whole phrase-final rhymes are deleted. However, deletion never takes place *ad infinitum*: whenever apocope leads to the creation of a final coda, this coda is not weakened any further.

Although phonetic effects and the partial variability in vowel deletion vs. devoicing confirm the incompleteness of the sound changes in question, there is no doubt that whenever these processes apply in their full form, phonological effects can be directly observed (see arguments presented in §2). In addition, there is a *latent opacity* effect that emerges from quantitative data. Given these observations, an analysis of the data in a generative phonology framework may pose a challenge. However, constraint-based models operating under the assumption of violability

---

[13] This also demonstrates that there is no UR restructuring and word-final /s/ is still in the underlying representation.

[14] It must be noted that words with the same characteristics in terms of information load and intonation but with stressed final vowels and words in which other lenition processes apply were excluded from the counts. Thus, there are many more potential contexts for final consonant deletion, albeit without any influence on the apocope results. The percentages of final consonant deletion regardless of prosody and pragmatics are provided for comparison.

should be nevertheless able to account for opaque surface structures, as we show in the following subsections.[15]

3.1 Fed counterfeeding opacity

An important observation concerning the data from this dialect is that the interaction between apocope and consonant deletion, and the resultant ban on multiple deletions are due to opacity. If we look at surface forms like [ˈpas] (**Table 2**), we can immediately notice the underapplication of an otherwise prevalent process of consonantal lenition. We are dealing with a counterfeeding rule order (Kiparsky 1971). If we were to establish rules for the discussed processes, we would note that apocope counterfeeds consonant deletion as the output of apocope meets the context of application of consonant deletion, but deletion does not apply. This counterfeeding relationship between the two rules is not so straightforward, however, given that in a subset of cases, i.e. in consonant-final words, consonant deletion provides the context for, and thus feeds, apocope. As a result, we have an instance of fed counterfeeding (Kavitskaya & Staroverov 2010; see also Baković 2011).

		Transparent	**Opaque 1**	**Opaque 2**
UR:		*hacer* **'to do'** /aser/	*paso* **'step'** /paso/	*pasos* **'steps'** /pasos/
C deletion		[a.ˈse]	–	[ˈpa.so]
Apocope		–	[ˈpas]	[ˈpas]
SR:		[a.ˈse]	[ˈpas]	[ˈpas]

Table 2: Rule derivation of vowel + consonant deletion cases from Gran Canarian Spanish.

In **Table 2**, the transparent case shows the application of consonant deletion at the end of a word. Apocope does not apply because the word-final vowel is stressed. There are two types of opaque outputs that can be produced, assuming that the two rules apply whenever their structural descriptions are met. In the case of a vowel-final form (e.g. *paso* 'step'), we can see the counterfeeding relationship between apocope and consonant deletion. The latter cannot apply since it is ordered before apocope. In the case of a consonant-final form (e.g. *pasos* 'steps'), we see that consonant deletion first feeds apocope and is then counterfed by it. This can be referred to as fed counterfeeding on environment, since the processes potentially feed each other but each process applies at most once. This is especially problematic for parallel OT, in which no rule ordering can be imposed. Also, we can imagine that the situation will get even more complicated

---

[15] In our analyses, we focus on phrase-final positions and hence phrase-final C deletion given that apocope takes place only phrase-finally. We also base our analyses and the investigated surface distributions on full apocope cases. We assume that incomplete vowel deletions are not subject to phonological analysis.

with words such as *pájaros* 'birds', in which transparent feeding would lead to the deletion of most parts of the word (e.g. /paxaɾos/ → [ˈpa] mapping reflecting the following changes: /paxaɾos/ → [ˈpaɦaɾo] → [ˈpaɦaɾ] → [ˈpaɦa] → [ˈpaɦi] → [ˈpa]).

As pointed out by an anonymous reviewer, our case is similar to a fed counterfeeding interaction in Lardil (Kavitskaya & Staroverov 2010; Baković 2011; see also references therein). In Lardil words longer than 2 syllables, final vowels and non-coronal consonants are deleted. Final vowel deletion feeds final consonant deletion, but not *vice versa*: *mungkumungku → mungkumu*, **mungkum* 'wooden axe' (Kavitskaya & Staroverov 2010: 2). Both the Gran Canarian Spanish and the Lardil interaction involve word-final consonant and vowel deletion. However, in Lardil, the processes apply in the opposite order (vowel deletion before consonant deletion). In addition, Gran Canarian Spanish conditions vowel deletion through stress (stressed vowels remain), whereas Lardil does so through word length. Finally, no variation is reported for the Lardil case, which makes that case simpler in some aspects. As will be shown in the next subsection, variation is crucial in understanding surface options and all process interactions involved in our data, and later modelling them in a successful manner.

3.2 Modelling variation: latent opacity

As indicated in §2, our data crucially involve variation so all surface variants must be generated (/pasos/ → [ˈpasos ~ ˈpaso ~ ˈpas], /ˈpaso/ → [ˈpaso ~ ˈpas]). However, underlyingly V-final words tend to surface **without** their unstressed final vowel (61% /paso/ → [ˈpas] vs. 39% /paso/ → [ˈpa.so]), while underlyingly C-final words, when their final C is deleted, tend to surface **with** an unstressed final vowel (36% /pasos/ → [ˈpas] vs. 56% /pasos/ → [ˈpa.so]; 39% vs. 61% when only tokens with consonant deletion are considered). We will now argue why this is a *latent opacity* effect.

Since V- and C-final words have to be modelled with the same grammar, the difference in vowel deletion rates between them must be captured. In a variable ranking grammar (such as the one we will be working with, see §4.1–2), this means we must have rankings that generate various combinations of surface patterns, as indicated in **Table 3**.[16] Logically speaking, apart from the ranking under which both processes apply in all word types, which we can refer to as A – full lenition, there can be a ranking under which neither process applies (B – no lenition); a ranking under which final vowels are retained, but final consonants are deleted (C – consonant deletion only), and a ranking under which final vowels are deleted, but final consonants are retained (D – apocope only). However, as we argue below, to obtain a higher rate of apocope in

---

[16] A similar argument can be made for variable weighting grammars. For Maximum Entropy grammars, the argument is a bit more complex, since these generate probabilities without perturbing constraint weights. Unfortunately, we cannot present this argument within the scope of this paper.

V-final words compared to C-final words, we also need an additional ranking (E – mixed pattern) where final consonants and final underlying vowels are deleted, but underlying vowels which become final after consonant deletion are retained. The surface pattern derived by each ranking will henceforth be called a (surface) variant.

Rankings	Surface variants	Descriptive name	/pasos/	/paso/
A	Variant A	full lenition	ˈpas	ˈpas
B	Variant B	no lenition	ˈpasos	ˈpaso
C	Variant C	C deletion only	ˈpaso	ˈpaso
D	Variant D	apocope only	ˈpasos	ˈpas
E	Variant E	mixed pattern	ˈpaso	ˈpas

Table 3: Surface variants to be modelled.

The need for variant (and ranking) E can be shown as follows. If there are rankings that generate only variants A, B, C, and D, and speakers pick a different ranking at each instance of grammar use (cf. Boersma 1998), the following conundrum ensues. We know that faithful realization of /pasos/ occurs 8% of the time (in the remaining 92% of the cases, C deletion applies), meaning that rankings B and D together should be picked in no more than 8% of language use. Furthermore, /pasos/ → [ˈpaso] occurs 56% of the time, meaning that ranking C is picked 56% of the time, and /pasos/ → [ˈpas] occurs 36% of the time, meaning that ranking A is picked 36% of the time. This is shown in **Table 4**.

Rankings	Surface variants	Descriptive name	/pasos/	/paso/	Frequency of picking ranking
A	Variant A	full lenition	ˈpas	ˈpas	36%
B	Variant B	no lenition	ˈpasos	ˈpaso	8%
D	Variant D	apocope only	ˈpasos	ˈpas	
C	Variant C	C deletion only	ˈpaso	ˈpaso	56%

Table 4: Attempt at modelling data without a ranking for Variant E.

A model without a mechanism to generate Variant E cannot match the correct rates for vowel deletion for C-final and V-final words. This is because the model generates /pasos/ → [ˈpaso] at the same rate as /paso/ → [ˈpaso], which grossly overestimates how often the latter occurs: 56% of the time instead of the attested 39%. At the same time, if /paso/ → [ˈpas] is only generated by rankings A and D, its rate of occurrence will be grossly underestimated. Ranking D cannot

be chosen more than 8% of the time (because of /pasos/ → [ˈpasos]), while ranking A must be chosen 36% of the time (/pasos/ → [ˈpas]), meaning that /paso/ → [ˈpas] could occur at most 8% + 36% = 44% of the time instead of the attested 61%. To correctly predict the rate of /paso/ → [ˈpas], we must assume that there is also a different ranking that generates /paso/ → [ˈpas] but not /pasos/ → [ˈpasos] or [ˈpas]: the mixed E pattern.

Note that this mixed pattern is, in itself, opaque. C-final words like /pasos/ undergo consonant deletion, but fail to undergo vowel apocope afterwards: /pasos/ → [ˈpaso] *→ ˈpas. V-final words like /ˈpaso/ undergo vowel apocope, but fail to undergo consonant deletion afterwards: /paso/ → [ˈpas] *→ ˈpa. This constitutes a chain shift (VC# → C#, C# → ∅#; Baković 2011), but it is also a case of mutual counterfeeding (Wolf 2011): consonant deletion counterfeeds vowel deletion and *vice versa*. Since this opaque mapping is not necessary to derive any of the individual surface forms observed in the language, but only to model the quantitative pattern of variation, we refer to this as a case of *latent opacity*.

3.3 Evaluation of the Gran Canarian data under Serial Markedness Reduction

In the remainder of this section, we present an analysis of the Gran Canarian data in the framework of Serial Markedness Reduction (SMR, Jarosz 2014), which offers constraints on ordering in derivations that help model opacity of various types. Thus, it is appropriate for generating vowel apocope and consonant deletion in interaction. While other frameworks exist that could in principle model the opaque data,[17] SMR has one major advantage – there is an existing probabilistic learner for it (Jarosz et al. 2018), which allows us to test numerically how well a probabilistic ranking version of SMR can capture the variation in our data (see §4).

3.3.1 Deriving full lenition – Variant A

SMR is a version of Harmonic Serialism (McCarthy 2008) that enables extrinsic process ordering by tracking for each candidate which markedness constraints were newly satisfied by that candidate and each of its derivational predecessors. This is represented in *Mseq*, an integral part of the candidate. The order in which markedness constraints are satisfied is controlled by so-called serial markedness (SM) constraints which mandate a certain order in which a pair of markedness constraints has to be satisfied in a derivation. Thus we have the iterative evaluation mechanism

---

[17] Other frameworks include Hauser & Hughto's (2020) Contextual Faithfulness approach, which is briefly discussed in §5, and a probabilistic rule-based framework (e.g. Tajchman et al. 1995) in which the derivation would be possible but only under very specific assumptions (see Appendix 2). We thank an anonymous reviewer for the suggestion to discuss this. Finally, the framework that has been used to model a similar case, albeit without including variation, is OT-CC (McCarthy 2007). We will show in §5.1, however, that it is suboptimal as problems arise with generating the mixed pattern (Variant E).

of Harmonic Serialism with the addition of constraints that guide the overall derivation, allowing for some extrinsic ordering of processes.[18]

In our case, we need to derive outputs with consonant deletion and apocope. The constraints driving these processes are *Final-C and *UnstrV, respectively, as defined below.

(5)   Gran Canarian Spanish case – basic constraint definitions
       *Unstr[essed]V   Assign a violation mark for every unstressed vowel.[19]
       *Final-C   Assign a violation mark for every consonant standing in word-final position.

The two deletion processes can be modelled by ranking these markedness constraints above Max(seg).[20] Additionally, since we are using a general markedness constraint to ensure apocope, we have to make sure that unstressed vowels other than final are left unscathed. This is effected via undominated, positional faithfulness constraints Max(V)/Initial (Beckman 1998) and Contiguity (McCarthy & Prince 1994), defined as in (6).[21]

(6)   Constraint definitions: Max(V)/Initial and Contig[uity]
       Max(V)/Initial   Assign one violation mark for every word-initial input vowel that has no output correspondent.
       Contig[uity]   Assign a violation mark for every pair of non-adjacent stem input segments whose output correspondents are adjacent.

An illustration of the interaction of these constraints is provided in (7), where for /paso/ 'step', unstressed vowel deletion (7b) wins for its lack of a *UnstrV violation, while for /akolito/ [aˈkolit] 'acolyte',[22] deleting the penultimate (7e) or initial (7f) unstressed vowels is not harmonically improving due to high-ranking Contiguity and Max(V)/Initial, respectively.

---

[18] The way SMR functions resembles the *LUMseq* and Prec constraints used to impose precedence relations among faithfulness constraint violations in OT-CC (McCarthy 2007).

[19] In principle, we might postulate a positional constraint here, stating that there should be no final unstressed vowels. However, this would lead to a ranking paradox similar to the one described by Kavitskaya & Staroverov (2010); see also footnote 36 in §5.1.1.

[20] Note that we use one non-positional faithfulness constraint, Max(seg), instead of Max(V) and Max(C) separately, as in the OT-CC analysis in §5.1.1. However, an analysis using Max(V) and Max(C) instead of Max(seg) would yield the same results.

[21] According to McCarthy (2003), Contiguity should be treated as a contextually restricted faithfulness constraint and can be divided into I-Contig (a special version of Max banning internal deletions) and O-Contig (a special version of Dep banning internal epenthesis).

[22] This example is used to illustrate the behaviour of both initial and non-final unstressed vowels. The corpus contains similar words, e.g. *adelante* 'ahead' and *entonces* 'so' with penult stress.

(7) Evaluation of the words *paso* 'step' and *acólito* 'acolyte' using positional faithfulness constraints

/paso/	MAX(V)/INITIAL	CONTIGUITY	*UNSTRV	MAX(seg)
a. ˈpaso < >			*!	
b. ☞ˈpas <*UNSTRV@4>[23]				*
/akolito/				
c. aˈkolito < >			***!	
d. ☞ aˈkolit <*UNSTRV@7>			**	*
e. **aˈkolto <*UNSTRV(V)@5>		*!	**	*
f. **ˈkolito <*UNSTRV@1>	*!		**	*

Now, to model final consonant deletion in words like *pasos* 'steps', *FINAL-C must be ranked above MAX(C), but below *UNSTRV. In our data, all final consonant deletion appears in contexts where it interacts with vowel apocope, so the effect of *FINAL-C will be shown within this interaction. Furthermore, since we are dealing with a fed counterfeeding interaction between consonant deletion and vowel apocope, we need to ensure that the former applies first, followed by the latter, and that consonant deletion does not happen (again) after apocope. For this, we use an SM constraint, SM(*FINAL-C,*UNSTRV), as defined below.

(8) Definition of the key Serial Markedness constraint
     SM(*FINAL-C,*UNSTRV)   Assign a violation mark for every satisfaction of *FINAL-C
                              that follows a satisfaction of *UNSTRV in a candidate's *Mseq*

Ranking this constraint above *UNSTRV will make the derivation converge on the candidate with non-iterative consonant deletion, which is the desired result. The derivation is presented in (9) using two examples that differ minimally and hence best illustrate the differences and similarities between V-final and C-final stems.

---

[23] Following Jarosz (2014), we indicate the *Mseq* (order of markedness constraint satisfactions) for every candidate in angled brackets, indicating for each unfaithful mapping the markedness constraint it violates and at which segment in the input (locus) this happens. Since *loci* will not be crucial in our case, we will not indicate them for any following SMR tableaux; see also footnote 24.

(9)　Evaluation of the words *paso* and *pasos* using SMR (Variant A – full lenition)[24]

Step 1

	CONTIG	SM(*FINAL-C,*UNSTRV)	*UNSTRV	*FINAL-C	MAX(seg)
Input: pasos < >					
a. ˈpa.sos < >			*	*!	
b. ˈpass <UNSTRV >	*!			*	*
c. ☞ˈpa.so <*FINAL-C >			*		*
Input: paso < >					
a. ˈpa.so < >			*!		
b. ☞ˈpas <UNSTRV>				*	*

Step 2

	CONTIG	SM(*FINAL-C,*UNSTRV)	*UNSTRV	*FINAL-C	MAX(seg)
Input: ˈpa.so <*Final-C>					
a. ˈpa.so <* FINAL-C>			*!		
b. ☞ˈpas <* FINAL-C,UNSTRV>				*	*
Input: ˈpas <UNSTRV>					
a. ☞ˈpas <UNSTRV>				*	
b. ˈpa <UNSTRV,* FINAL-C>		*!			*

---

[24] In our SMR derivations, we count constraint satisfactions (markedness reductions) only, following the proposal by Jarosz (2014: §3.2). For simplicity, we do not keep track of constraint satisfaction *loci* (Jarosz 2014:§5.2), as all consecutive markedness satisfactions in our analysis interact with one another (i.e., final consonant deletion makes final vowel deletion possible).

Step 3

	Cᴏɴᴛɪɢ	SM(*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ)	*Uɴsᴛʀⱽ	*Fɪɴᴀʟ-C	Mᴀx(seg)
Input: ˈpas <*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ>					
a. ☞ˈpas <*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ >				*!	
b. ˈpa <*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ,*Fɪɴᴀʟ-C>		*!			*

The tableaux in (9) show that the serial markedness constraint prevents consonant deletion that follows vowel apocope. In this way, we can account for the fed counterfeeding pattern found in Gran Canarian Spanish.

3.3.2 Deriving alternative surface patterns – Variants B-E

In §3.3.1 we showed a successful evaluation of the data presented in (9) under the SMR framework. However, the variation seen in the data in §2 must also be accounted for (see also §3.2). As we will show in this subsection, the SMR framework is able to account for all the five surface variants presented in **Table 3**.

As we have seen in (9), to generate VC deletion in consonant-final and vowel apocope in vowel-final words (Variant A), we need to make sure that the serial markedness constraint is ranked above both markedness constraints that compose it; this constraint is undominated, like Cᴏɴᴛɪɢ and Mᴀx(V)/Iɴɪᴛɪᴀʟ. The markedness constraints mandating deletion, in turn, must be ranked above the Mᴀx(seg) constraint. Thus, the correct ranking is: Mᴀx(V)/Iɴɪᴛɪᴀʟ, Cᴏɴᴛɪɢ, SM(*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ) >> *Uɴsᴛʀⱽ >> *Fɪɴᴀʟ-C >> Mᴀx(seg).

To derive faithful forms, i.e. with no final C deletion nor final unstressed V deletion (as in /paso/ → [paso] and /pasos/ → [pasos], Variant B), we need both markedness constraints to be ranked below Mᴀx(seg). From **Table 1** we know that the probability of words such as *pasos* to surface as [pasos] is 8%.[25] The probability of getting faithful [paso] from /paso/, on the other hand, is 39% (but this mapping is also generated by rankings C and E, so this higher probability can be accounted for). Thus, we need to rank Mᴀx(seg) above *Uɴsᴛʀⱽ, which yields Mᴀx(V)/Iɴɪᴛɪᴀʟ, Cᴏɴᴛɪɢ >> Mᴀx(seg) >> *Uɴsᴛʀⱽ, *Fɪɴᴀʟ-C; SM(*Fɪɴᴀʟ-C,*Uɴsᴛʀⱽ) can be ranked anywhere, since satisfaction of either markedness constraint is unmotivated given high-ranked Mᴀx(seg). This is illustrated in (10).[26]

---

[25] Actually, this probability refers to a debuccalized variant [pa.soh], but we ignore this detail because this would over-complicate the analysis by adding additional constraints responsible for debuccalisation.

[26] The SM constraint is not necessary to derive this variant and its ranking may be different than in Variant A. Here and in the rest of the section, only active constraints will be included in the ranking.

(10)   Step 1 derivation of the words *pasos* and *paso* with faithful candidates as winners
(Variant B – no lenition)

	Contig	Max(seg)	*Unstrv	*Final-C
Input: pasos < >				
a. ☞ˈpa.sos < >			*	*
b. ˈpass <Unstrv >	*!	*		*
c. ˈpa.so <*Final-C>		*!	*	
Input: paso < >				
a. ☞ˈpa.so < >			*	
b. ˈpas <Unstrv>		*!		*

The third option (Variant C) is one where only consonant deletion applies. As mentioned in §2, underlyingly consonant-final forms may surface without their final consonants, but keep their last vowels, while vowel-final forms may keep their final vowels. This is effected by demoting *Unstrv below *Final-C. The ranking Max(V)/Initial, Contig >> *Final-C >> *Unstrv >> Max(seg) ensures such a state of affairs (see derivation in 11); SM(*Final-C,*Unstrv) can be ranked anywhere, since apocope never applies.

(11)   Derivation of the words *pasos* and *paso* with consonant deletion only (Variant C – consonant lenition only)

Step 1

	Contig	*Final-C	*Unstrv	Max(seg)
Input: pasos < >				
a. ˈpa.sos < >		*!	*	
b. ˈpass <Unstrv >	*!	*		*
c. ☞ˈpa.so <*Final-C>			*	*
Input: paso < >				
a. ☞ˈpa.so < >			*	
b. ˈpas <Unstrv>		*!		*

Step 2

		Contig	*Final-C	*Unstrv	Max(seg)
Input: ˈpa.so <* Final-C>					
a. ☞ˈpa.so <* Final-C>				*	
b. ˈpas <* Final-C,Unstrv >			*!		*

In step 1 in (11), high-ranked *Final-C rules out final consonants, which triggers final consonant deletion in *pasos* but blocks final vowel deletion in *paso* (high-ranked Contig prevents *Unstrv from being satisfied in *pasos*). In step 2, it is shown that *pasos* only undergoes final consonant deletion but not vowel apocope because it is more important to avoid final consonants than it is to have no unstressed vowels.

There is also the possibility that only apocope applies (Variant D), yielding /pasos/ → [ˈpasos], but /paso/ → [ˈpas]. Such a situation is ensured by ranking Max(seg) between *Unstrv and *Final-C, as demonstrated below, yielding the full ranking Max(V)/Initial, Contig >> *Unstrv >> Max(seg) >> *Final-C. SM(*Final-C,*Unstrv) can be ranked anywhere, since the attested candidates never involve consonant deletion.

(12)   Derivation of the words *pasos* and *paso* with apocope only (Variant D – apocope only)

	Contig	*Unstrv	Max(seg)	*Final-C
Input: pasos < >				
a. ☞ˈpa.sos < >		*		*
b. ˈpass <Unstrv >	*!		*	*
c. ˈpa.so <*Final-C>		*	*!	
Input: paso < >				
a. ˈpa.so < >		*		
b. ☞ˈpas <Unstrv>			*	*

Finally, one more scenario has to be taken into account: Variant E (§3.2). To model the *latent opacity* effect, that is, the relative underapplication of vowel deletion in C-final words compared to V-final words, we need a ranking in which vowel apocope applies in V-final words only,

i.e. /paso/ → [ˈpas] but /pasos/ → [ˈpa.so]. This ranking requires another SM constraint, SM(*UNSTRV,*FINAL-C), violated once for every instance of *UNSTRV satisfaction after an instance of *FINAL-C satisfaction. Together with high-ranked SM(*FINAL-C,*UNSTRV), this constraint, if ranked above *UNSTRV, blocks consonant deletion and vowel apocope from occurring in the same derivation, thus obtaining Variant E, as illustrated in (13). If SM(*UNSTRV,*FINAL-C) is ranked below *UNSTRV and all other constraints are ranked the same, Variant A is obtained. For rankings B-D, SM(*UNSTRV,*FINAL-C) can be ranked anywhere, as the corresponding surface variants never involve apocope after final consonant deletion.

(13)   Second step of derivation of the words *pasos* and *paso* with high-ranked
        SM(*UNSTRV,*FINAL-C) (Variant E – mixed pattern)

        Step 2

	CONTIG	SM(*FINAL-C,*UNSTRV)	SM(*UNSTRV,*FINAL-C)	*UNSTRV	*FINAL-C	MAX (seg)
Input: ˈpa.so <*FINAL-C>						
☞ ˈpa.so <*FINAL-C>				*		
ˈpas <*FINAL-C,*UNSTRV>			*!		*	*
Input: ˈpas <*UNSTRV>						
☞ ˈpas <*UNSTRV>					*	
ˈpa <*UNSTRV,*FINAL-C>		*!		*		*

The first step for the ranking and inputs in (13) is identical to the first step in (9), where the ranking *UNSTRV >> *FINAL-C >> MAX(seg) leads to the deletion of the final segment (CONTIG blocks deletion of a medial unstressed vowel). In the second step, shown in (13), the role of both SM constraints becomes crucial. As in (13), SM(*FINAL-C,*UNSTRV) blocks the deletion of the final consonant in [ˈpas], but now SM(*UNSTRV,*FINAL-C) also blocks the deletion of the final vowel in [ˈpaso]. A high ranking for both SM constraints thus leads to the deletion of just final consonants or just final vowels.

This shows that the proposed SMR analysis is able to derive all the attested phonological surface variants, i.e. both vowel and consonant deletion in all types of words, the absence of consonant and vowel deletion, the absence of vowel deletion but presence of consonant deletion, the absence of consonant deletion but presence of vowel deletion, and both consonant and vowel deletion but with vowel deletion applying only in V-final words. The relevant rankings are summarised in Appendix 3.

Now that we know the proposed analysis can derive all surface options provided that minimal reranking is allowed, we will test with a learning algorithm whether a probabilistic version of our analysis can derive the attested frequencies of the options. This is the goal of §4.

4. Learning

As can be seen in §3.3, the deletion patterns of Gran Canarian Spanish can be accounted for in Serial Markedness Reduction (Jarosz 2014). However, it is important to ensure that all attested variants can be generated by the same probabilistic grammar, and that the rates of consonant and vowel deletion can be matched by this grammar (cf. the importance of ranking E). In addition, the discoverability of this analysis from ambient language data is important, since younger male speakers have indeed internalized this pattern.

Here, we use Jarosz's (2015) probabilistic ranking grammars to represent optionality and variation (§4.1) and the concomitant Expectation-Driven Learning framework (§4.2) to learn the optimal probabilistic constraint rankings from the Gran Canarian Spanish data. This method works for both parallel OT and for Harmonic Serialism and has an existing implementation for SMR (Jarosz et al. 2018), making it an ideal candidate for a probabilistic representation of the analysis sketched in §3.3.

We will show the results of several learning simulations to tease apart the effects of rankings A (requires one of the Serial Markedness (SM) constraints in the analysis) and E (requires both SM constraints; see §3.3 and Appendix 3). The results demonstrate that both rankings are necessary to fully account for the pattern, confirming the need for *latent opacity*.

4.1 Probabilistic grammar framework

Jarosz's (2015) framework operates on strictly ranked constraints, as opposed to weighted-constraint alternatives such as Harmonic Grammar (Legendre et al. 1990) and related approaches like Maximum Entropy Grammar (Goldwater & Johnson 2003). Like Stochastic OT (Boersma 1998), Jarosz's framework defines probabilities over rankings. Differently from Stochastic OT, Jarosz (2015) represents these probabilities directly: for every pair of constraints, the grammar represents what the probability is that one of these constraints is ranked over the other. Accordingly, Jarosz names these grammars Pairwise Ranking Grammars. Assigning probabilities to rankings allows the expression of variation: multiple rankings are possible given the grammar,

with potentially different outcomes for the same input (an example will be given below). The reason for representing these probabilities directly rather than through weights as in Stochastic OT comes from learning efficiency (Jarosz 2015): they allow for Expectation-Driven Learning.

An example of a Pairwise Ranking Grammar is given in **Table 5**, where ranking probabilities over three constraints – *UNSTRV, *FINAL-C, MAX(seg) – are represented. This grammar represents a fixed ranking *FINAL-C >> MAX(seg), as can be seen in the top right cell of the tableau with 100% probability for *FINAL-C >> MAX(seg) and in the bottom left cell with 0% probability for MAX(seg) >> *FINAL-C. *UNSTRV has a variable ranking with a tendency to rank in between the former two constraints. This can be seen in the middle row and the centre column of the table. The top centre cell indicates 70% probability for *FINAL-C >> *UNSTRV (the mid left cell correspondingly indicates 30% probability for *UNSTRV >> *FINAL-C), so *UNSTRV is most likely below *FINAL-C. The mid right cell indicates an 80% probability for *UNSTRV >> MAX(seg), while the bottom centre cell correspondingly indicates a 20% probability for MAX(seg) >> *UNSTRV. This tells us that there is a tendency for *UNSTRV to rank above MAX(seg).

	... >> ***FINAL-C**	... >> ***UNSTRV**	... >> **MAX(seg)**
*Final-C >> ...		70%	100%
*UnstrV >> ...	30%		80%
Max(seg) >> ...	0%	20%	

Table 5: Example of a Pairwise Ranking Grammar.

Like in Stochastic OT, every time the grammar is used, a specific ranking of these constraints is sampled from the grammar (see Jarosz 2015 for the sampling procedure).[27] For the grammar in **Table 5**, the most likely ranking is *FINAL-C >> *UNSTRV >> MAX(seg), which yields final C and V deletion: /pasos/ → [ˈpas] (Variant A, presuming high ranking of SM(*FINAL-C,*UNSTRV)). However, there is a chance of sampling *FINAL-C >> MAX(seg) >> *UNSTRV (since there is a 20% probability that MAX(seg) >> *UNSTRV), which would lead to the deletion of final consonants only: /pasos/ → [ˈpaso] (Variant C). Crucially, there is no chance in this grammar that MAX(seg) >> *FINAL-C >> *UNSTRV, which would lead to /pasos/ → [ˈpasos] (Variant B), since the probability of MAX(seg) >> *FINAL-C is 0. The relative likelihood of each of these rankings means that /pasos/ → [ˈpas] will occur most often, /pasos/ → [ˈpaso]

---

[27] To avoid conflicts between sampled pairwise rankings (e.g., A >> B, B >> C, C >> A), Jarosz specifies that all cells of the matrix be put in a single random order (new order picked for every new sample); going through the matrix cells in this order, the algorithm samples a 0 or a 1 where the number in the cells determines the probability of sampling 1; after a cell (= pairwise ranking) has been set to 0 or 1, the algorithm sets the probability of any ranking that's implied by transitivity to 1 and the probability of any incompatible ranking to 0 before moving to the next cell. This guarantees the sampled ranking will always be consistent within itself.

less often, and /pasos/ → [ˈpasos] will never occur. The precise probability of a mapping given the grammar can be estimated by taking many samples from the grammar (in our case, 1000) and counting how often a ranking is chosen under which this mapping wins; for instance, if 832 of 1000 sampled rankings yield /pasos/ → [ˈpaso], the probability of /pasos/ → [ˈpaso] will be estimated as 83.2% (the same procedure is used in Stochastic OT, Boersma 1998, and Noisy Harmonic Grammar, Coetzee and Pater 2011).

Since this framework is based on ranked constraints, and the properties of GEN and EVAL remain unaltered, it can be straightforwardly applied to HS: every time the grammar is used, a ranking is picked, and this ranking fully determines the HS derivation. For instance, if from the grammar in **Table 4** the learner samples *FINAL-C >> *UNSTRV >> MAX(seg), the HS derivation will be /pasos/ → /ˈpaso/ → /ˈpas/ → [ˈpas] (corresponding to Variant A); if a different ranking is sampled from the Pairwise Ranking Grammar, this ranking may determine another HS derivation. Such a setup is different from the MaxEnt implementation of probabilistic HS (Staubs & Pater 2016) in which each step of the HS derivation is made probabilistic (changing EVAL). The probabilistic nature of the latter setup alters some of the properties of HS and requires some ad-hoc adjustments (see Staubs & Pater 2016).

4.2 The Expectation-Driven Learning framework

To simulate the learning of the Canary Islands Spanish patterns, we use the batch version of Jarosz's (2015) Expectation-Driven Learning (EDL) mechanism, which learns Pairwise Ranking Grammars from data using the general principles of Expectation Maximization (EM; Dempster et al. 1977), a method of machine learning that is guaranteed to maximize a model's fit to the training data even when the learning problem has great complexity. When using a serial framework like HS, this is especially relevant, since the outcome of the model is mediated by potentially many derivational steps, each of which reflects on the overall ranking that must be learned. Our choice for EDL is also motivated by the fact that the only existing implementation of learning SMR grammars is in EDL (Jarosz 2016; Jarosz et al. 2018).

The learner starts with an initial grammar hypothesis (we used a uniform distribution over all pairwise rankings) and then iterates a cycle consisting of the E(xpectation)-step (compute the expected ranking probabilities given the current grammar hypothesis and the data set) and the M(aximization)-step (replace the current grammar hypothesis by the expected ranking probabilities just found at the E-step) until convergence (the M-step does not significantly change the grammar hypothesis) or until timeout. The details of how this learner updates the grammar hypothesis are given in Appendix 4.

The application of this method to HS is straightforward: it only requires an implementation of standard HS and a way of checking whether the output of the HS derivation given a particular

ranking and input matches the intended mapping. We use a slightly updated version of Jarosz et al.'s (2018) code, which integrates Expectation-Driven Learning and SMR (as well as other variants of HS); our updates to their code allow for the definition of faithfulness constraints that use context (MAX(V)/INITIAL and CONTIGUITY) and for a more general application of Serial Markedness constraints.

4.3 Simulation setup

4.3.1 Datasets

For the simulations, we use a dataset that includes the words *paso(s)* 'step(s)' /paso(s)/, as well as words with multiple unstressed vowels (*pájaro(s)* 'bird(s)'/paxaɾo(s)/) and words with initial unstressed vowels and consonant clusters in which rhyme apocope leads to a final complex coda: *oferta(s)* 'offer(s)' /ofeɾta(s)/, *metro(s)* 'metre(s)' /metɾo(s)/. Since different frequencies of the processes analysed here are associated with the singular vs. plural forms, we include both options in the simulations. For all aforementioned words (inputs), output candidates representing each attested pronunciation are offered to the learner at frequencies obtained from the data described in §2; the resulting mappings are shown in **Table 6**.

Input	Output	Frequency
/paso, paxaɾo, metɾo, ofeɾta/	[ˈpa.so,ˈpa.xa.ɾo,ˈme.tɾo, oˈfeɾ.ta]	39
	[ˈpas,ˈpa.xaɾ,ˈmetɾ, oˈfeɾt]	61
/pasos, paxaɾos, metɾos, ofeɾtas/	[ˈpa.sos,ˈpa.xa.ɾos,ˈme.tɾos, oˈfeɾ.tas]	8
	[ˈpa.so,ˈpa.xa.ɾo,ˈme.tɾo, oˈfeɾ.ta]	56
	[ˈpas,ˈpaxaɾ,ˈmetɾ, oˈfeɾt]	36

Table 6: Frequencies of mappings offered to the learner.

4.3.2 Gen and con

Since our data only includes deletion of consonants and unstressed vowels, and there is no way of satisfying the crucial markedness constraints *FINAL-C and *UNSTRV through epenthesis, we restrict *GEN* in our simulations to deleting any consonant or unstressed vowel – no insertion or change operations are considered.[28] Stress assignment is not modelled: stress is marked as an inherent property of a vowel, which is a necessary simplification.

---

[28] *FINAL-C or *UNSTRV could be satisfied by consonants' turning into vowels or *vice versa*, which should then be blocked by high-ranked IDENT(vocalic). For stressed vowel deletion, inherently disallowed, we would need to include the high-ranked constraint IDENT(V)/STRESS.

Based on this setup, three different constraint sets are used to investigate the importance of rankings A and E (Appendix 3). The basic constraint set consists of *Final-C, *Unstrv, Max(seg), and Contig, as well as Max(V)/Initial, which does not allow rankings A or E. Then the effect of adding Serial Markedness constraints to this basic constraint set is studied: SM(*Final-C,*Unstrv) and SM(*Unstrv,*Final-C) are considered. The former SM constraint, as discussed in §3.3, is used in rankings A and E to block the deletion of a final consonant when this final consonant arises from the deletion of the following vowel: /pasos/ → /ˈpa.so/ → /ˈpas/ (*→ˈpa) → [ˈpas]. The latter SM constraint is used in ranking E to block vowel deletion after consonant deletion has applied, helping apocope apply only in V-final words.

In our simulations, the two SM constraints are added one by one: first SM(*Final-C,*Unstrv), then SM(*Unstrv,*Final-C). This is because ranking A only crucially involves SM(*Final-C,*Unstrv), while ranking E crucially involves both SM constraints. This yields three models in total whose setups and abbreviated names are summarized in **Table 7**.

noSM	1SM	2SM
No SM	SM(*Final-C,*Unstrv)	SM(*Final-C,*Unstrv), SM(*Unstrv,*Final-C)

Table 7: Overview of models to be trained.

All these models are trained on the dataset in **Table 6**.

4.3.3 Parameter settings and evaluation

With regards to the parameters of learning, the standard settings given in Jarosz et al.'s (2018) implementation are kept (batch learning, sample size for learning and evaluating is 1000, depth of search is 8), except for the number of iterations, which is set to 15 (instead of the default 10) to ensure that our simulations always converge despite the complexity of the dataset. Each of the three models is learned 20 times with a fully unbiased initialization (50% probability for all pairwise rankings). For each of these simulations, two metrics are computed: the mean absolute error (MAE; the average difference between how often a mapping occurs in the dataset and how often the model predicts it will occur),[29] and the data log-likelihood (the log of the probability that the current model will generate exactly the training data). Log-likelihood is a standard measure of model success (closer to 0 means better model fit), whereas MAE is a way to gauge

---

[29] In calculating the MAE, all predicted candidates that are not in the dataset are grouped together as 'other' and their predicted frequencies summed. This, if anything, overestimates the MAE, since there are fewer candidates to divide the total absolute error between.

how far off the model is from the target percentages on average, making it an alternative where log-likelihood is not interpretable.

4.4 Results

Table 8 shows the numerical results of the simulations. These results demonstrate that having SM constraints improves the fit of the model to the frequency distribution: no SM constraints yields a data log-likelihood of $-\infty$ (because some attested forms have a predicted probability of 0) and a high MAE (about 19), while models with SM constraints do have nonzero probability for all attested forms leading to finite negative log-likelihood and a markedly lower MAE (with non-overlapping confidence intervals (CIs)). Furthermore, the 2SM model does better than the 1SM model: it has a higher data log-likelihood and a lower MAE (both with non-overlapping confidence intervals). In fact, the 2SM model is only an average of 3 percentage points off on the relative frequency of each form.

	noSM	**1SM**	**2SM**
MAE	19.1 (19.0–19.7)	8.2 (8.1–8.6)	3.4 (3.3–3.7)
Log-likelihood	$-\infty$ (0 probabilities for attested forms)	-6.567 (−6.574; −6.555)	-6.505 (−6.510; −6.479)

Table 8: Numerical results of simulations for the three models, averaged across 20 runs (95% confidence intervals given in parentheses).

Qualitatively, as predicted, the model with no SM cannot handle the opaque interaction in /paso(s)/ and /paxaɾo(s)/: it is unable to produce final VC deletion, deleting all segments up until the stressed vowel instead: /paxaɾos/→[ˈpa], since SM(*FINAL-C,*UNSTRV) is not available (cf. ranking A, Appendix 3). /metɾo(s)/ and /ofeɾta(s)/ are mapped to attested candidates, but the frequency distribution is not captured adequately (Appendix 5, Table 11).

The 1SM model, as predicted, can generate all mappings in the data. However, since it does not allow ranking E, in which final unstressed vowels delete in underlyingly V-final words but not underlyingly C-final words, it cannot match the relative frequencies of vowel deletion in V-final and C-final words well: it predicts that it will happen equally often in both (Appendix 5, Table 13).

This is solved in the 2SM model. This model is able to generate the difference between underlyingly V-final and underlyingly C-final words ranking by not fixing the ranking between SM(*UNSTRV,*FINAL-C) and *UNSTRV, but representing a tendency for SM(*UNSTRV,*FINAL-C) to rank above MAX(seg) (see **Table 9**), so that the crucial subranking SM(*UNSTRV,*FINAL-C) >> *UNSTRV >> MAX(seg) from ranking E (Appendix 3), which blocks vowel deletion

specifically in underlyingly C-final words, appears often enough to ensure that C-final and V-final words receive the correct percentages of vowel apocope. In fact, this model very closely tracks the frequency distribution in the data file (Appendix 5, Table 15).
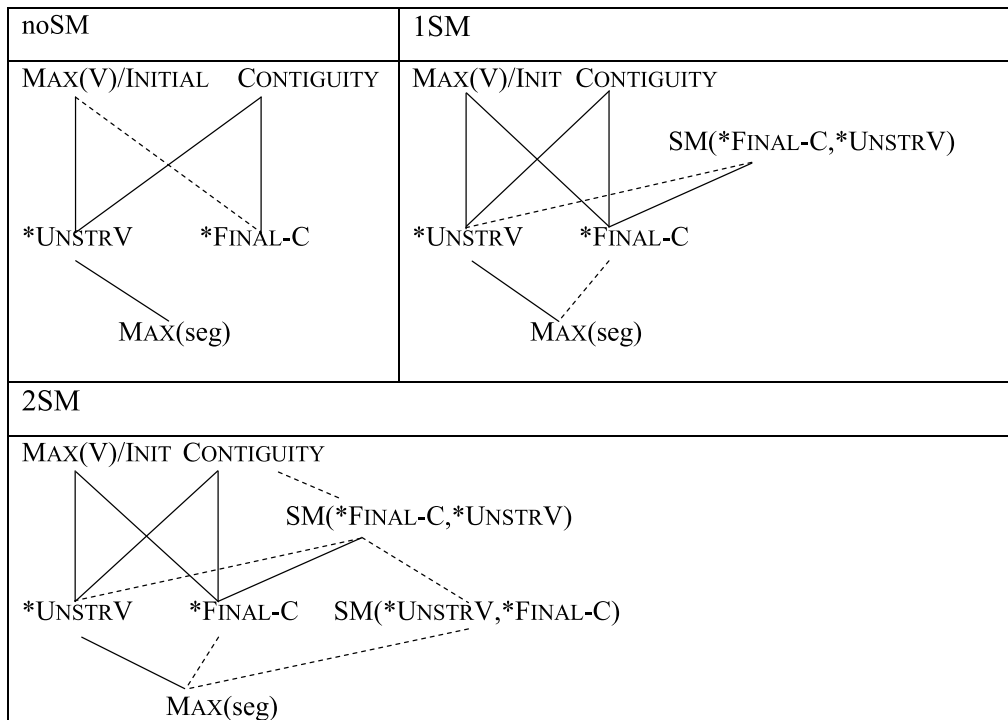
noSM	1SM
MAX(V)/INITIAL CONTIGUITY *UNSTRV *FINAL-C MAX(seg)	MAX(V)/INIT CONTIGUITY SM(*FINAL-C,*UNSTRV) *UNSTRV *FINAL-C MAX(seg)
2SM	
MAX(V)/INIT CONTIGUITY SM(*FINAL-C,*UNSTRV) *UNSTRV *FINAL-C SM(*UNSTRV,*FINAL-C) MAX(seg)	

Table 9: Hasse diagrams for probabilistic rankings resulting from the noSM, 1SM, and 2SM models.

Table 9 shows the resulting ranking probabilities for the noSM, 1SM and 2SM models. The Hasse diagrams are to be read as follows. A solid line indicates that the relevant ranking has a probability of at least 90% for all 20 runs for that model. A dashed line indicates that the relevant ranking has a probability of at least 70% for all 20 runs for that model, but it has a probability below 90% for at least 1 run.

In Table 9, it can be observed that, as more SM constraints are added, the ranking probabilities among the indicated constraint pairs increase or remain the same (no line < dashed line < solid line), meaning that the ranking of these constraint pairs gradually becomes more predictable.[30] For instance, *FINAL-C >> MAX(seg) has a probability of 65% for noSM (no line) and a probability of 89-90% for the 1SM and 2SM models (dashed line). The latter corresponds more closely to the desired 8% occurrence of final consonant retention, as it predicts MAX(seg) >> *FINAL-C about

[30] For all rankings shown in the diagrams, their minimum probabilities among 20 runs monotonically increase from noSM to 1SM to 2SM, except for the rankings among MAX(V)/INITIAL, *UNSTRV, and MAX(seg), whose minimum probabilities very slightly decrease, fluctuating by 1–3%.

10% of the time. This increase in constraint ranking predictability corresponds to an increase in accuracy in the models' predictions. In noSM, ranking A is unavailable because of the crucial role of SM(*Final-C,*Unstrv), leading the learner to overestimate the chance of full faithfulness to prevent mappings like /pasos/ → [ˈpa] from occurring too often (see Appendix 5, Table 11). In noSM and 1SM, Variant E is unavailable (since it crucially involves both SM constraints), leading the learner to settle on overestimating the rate of vowel deletion in vowel-final words and underestimating it in consonant-final words (see Appendix 5, Table 13). It is only the 2SM model that steers clear of this over- and underestimation and closely matches the attested distribution.

5. Discussion

In §2, we presented an intriguing case of process interaction in Gran Canarian Spanish that requires special attention in phonological terms. First, it shows a fed counterfeeding pattern that combines feeding of vowel apocope by consonant deletion with underapplication of the latter. Needless to say, this type of interaction has only been reported or analysed in the literature a few times, including Kavitskaya & Staroverov (2010) and Baković (2011). Second, our data show that variation leads to an emergence of an additional pattern in surface realisations: the rate at which vowel apocope applies is different depending on whether it is applying to an underlyingly final vowel (in V-final words) or one that is created by consonant deletion (in C-final words). As noted in §3, this *latent opacity* effect adds complexity to the formal analysis. Additionally, the interaction type it reveals, i.e. mutual counterfeeding, is exceptionally rare and barely attested across the world's languages (see Wolf 2011). To the best of our knowledge, no case of combined fed counterfeeding and mutual counterfeeding in one language has been reported to date, which makes our case all the more relevant for phonological theory.[31] Crucially, in §4, we provide a successful account of all the data in the framework of SMR (Jarosz 2014), using Expectation-Driven Learning (Jarosz 2015) to find grammars that provide a good fit to the data. Our probabilistic analysis demonstrates the fundamental role of SM constraints in generating the attested pattern of variation. There are a few remaining issues that we would like to discuss before our concluding remarks.

5.1 Opacity, variation and alternative analyses

As mentioned in the previous sections, a crucial point of our analysis is that it accounts for variable surface distributions. Our SMR analysis, in which ranking probabilities are optimised by machine, offers a comprehensive treatment of opacity-ridden variation. In §4 we showed that the variable surface distributions can be successfully mapped with the use of two serial markedness constraints mandating precedence relations between the two analysed processes. To the best of our knowledge, this is the first attempt to address opacity with variation using

---

[31] Interestingly, Kavitskaya & Staroverov (2010) mention three types of problematic cases that cannot be solved without modifying existing OT frameworks dedicated to solving opacity. We show that more than one such case can occur in the same language variety.

a learning algorithm and the first application of the SMR framework (and the EDL learner) to opacity patterns with variation.[32]

Importantly, the simulations in §4 show that the Canary Islands Spanish pattern of variation can be learned as long as the learner has access to the necessary Serial Markedness constraints. Alternatives to such constraints exist, including PREC constraints in Optimality Theory with Candidate Chains (OT-CC, McCarthy 2007), and contextual faithfulness constraints in parallel OT or HS (Hauser & Hughto 2020). Both approaches have been claimed to need additional mechanisms to deal with fed counterfeeding (Kavitskaya & Staroverov 2010; Hauser & Hughto 2020). Below (§5.1.1), we show an alternative analysis in OT-CC in which we model fed counterfeeding without additional mechanisms. We then show that mutual counterfeeding poses a greater challenge, however. This is followed by an alternative analysis using contextual faithfulness (§5.1.2), which shows, *contra* Hauser & Hughto (2020), that fed counterfeeding is possible in this framework, and that there is potential to analyse the current data in parallel OT.

5.1.1 Analysis of the data in OT-CC

OT-CC (McCarthy 2007) uses a derivational grammar framework and PREC(EDENCE) constraints to account for opacity. The candidates in an OT-CC tableau are entire derivations (candidate chains). Only candidate chains whose harmony with respect to the constraint ranking improves at each derivational step may be considered in a tableau, and the surface candidate that is pronounced corresponds to the most harmonic (winning) candidate chain in the tableau. PREC constraints apply to these chains, and if they are sufficiently high-ranked, they can block derivations with certain orders of process application.

In our data, final consonant deletion never applies to the result of vowel deletion. This can be captured by the constraint PREC(MAX(C), MAX(V)), as defined in (14).[33]

(14)   PREC(MAX(C),MAX(V)): Assign one violation mark for:
         (i) every pair of steps in a candidate chain in which MAX(C) is violated after MAX(V) (i.e., a derivation in which consonant deletion feeds apocope), and
         (ii) every step in a candidate chain in which MAX(V) is violated without a preceding MAX(C) violation (i.e., a derivation in which apocope happens without being fed by consonant deletion). (cf. McCarthy 2007: 98).

The PREC constraint in (14) must be ranked above *FINAL-C so that a violation of *FINAL-C is preferred to deleting a final consonant after vowel apocope, but below MAX(V) due to the ranking metaconstraint introduced by McCarthy (2007): PREC constraints must be outranked by the second (later) faithfulness constraint in their definition. Without this metaconstraint, PREC

---

constraints would lead to undesirable typological consequences (see McCarthy 2007: 101–102; Wolf 2011): a process can be blocked if it does not counterbleed another specific process.[34] Notably, Serial Markedness Reduction is similar to OT-CC in terms of the opacity-inducing mechanism but does not need the ranking metaconstraint (Jarosz 2014: 7–8), because Serial Markedness constraints are satisfied even when only one of the relevant markedness constraints is satisfied in the derivation, and thus could not motivate the typological problem described by McCarthy.

The OT-CC tableau for Variant A (/pasos/ → ['pas] and /paso/ → ['pas]) is presented in (15). In the tableau, candidates that do not represent a harmonically improving chain and are disqualified and indicated with two asterisks; they are still shown for clarity of comparison.

(15)   Derivation of /pasos/ → ['pas] and /paso/ → ['pas] in OT-CC

/pasos/	Contig	*Unstrv	Max(V)	Prec(Max(C), Max(V))	*Final-C	Max(C)
a. ˈpasos < >		*!			*	
b. ˈpaso <Max(C)@5>[35]		*!				*
c. ☞ ˈpas <Max(C)@5, Max(V)@4>			*		*	*
d. **ˈpass <Max(V)@4>	*!		*	*	*	
e. **ˈpa <Max(C)@5, Max(V)@4, Max(C)@3>			*	*!		**
/paso/						
f. ˈpaso < >		*!				
g. ☞ ˈpas <Max(V)@4>			*	*	*	
h. **ˈpa <Max(V)@4, Max(C)@3>			*	**!		*

---

[34] McCarthy's example: [i] only deletes if it has been able to trigger palatalization; /…ki…/ → […kʲ…] (deletion because counterbleeding occurs) but /…ri…/ → […ri…], where [r] does not palatalize (no deletion because no counterbleeding occurs).

[35] OT-CC's equivalent of Jarosz's *Mseq* is the *LUMseq*, also indicated in angled brackets, which is the record of a candidate's subsequent faithfulness violations with their respective *loci* (segments numbered from beginning of the word).

Tableau (15) shows that deleting a final consonant is harmonically improving (15a-b), and so is deleting the final vowel exposed by consonant deletion (15b-c). Deleting a final consonant exposed by vowel apocope (15e,h), would be harmonically improving if PREC(MAX(C),MAX(V)) were not ranked above *FINAL-C, so this ranking is crucial in modelling the fed counterfeeding. Note that, while (an additional) violation of PREC(MAX(C),MAX(V)) is sufficient to rule out candidates (15e,h), a violation of PREC(MAX(C),MAX(V)) is tolerated in the winning candidate (15g), because this violation allows eliminating a violation of higher-ranked *UNSTRV.

Now the remaining issue to solve is variation. We have seen that Variant A (full lenition) is easily derived in OT-CC.[36] Variant C (consonant deletion only) can be modelled by swapping the ranking of MAX(V) and *UNSTRV, which will make apocope no longer harmonically improving. To derive apocope but not consonant deletion (Variant D), in turn, we have to swap MAX(C) and *FINAL-C. Variant B (no deletion) can be modelled by additionally swapping the ranks of MAX(C) and *FINAL-C. However, it is not possible to model Variant E, in which only underlyingly final vowels delete. The straightforward OT-CC tool for this would be another PREC constraint: PREC(MAX(V),MAX(C)), with a definition identical to (14), except that MAX(V) and MAX(C) are swapped. This would penalise any derivation in which apocope takes place after consonant deletion. However, this constraint cannot be ranked high enough: to block apocope, it should be ranked above *UNSTRV, but the ranking metaconstraint forces it to be lower than MAX(C). Since we have already established in §3 that *UNSTRV >> *FINAL-C >> MAX(C), this means that PREC(MAX(V),MAX(C)) cannot be above *UNSTRV. In this low ranking, the PREC constraint is not able to reach its intended effect. Thus, although OT-CC is able to derive fed counterfeeding of the type presented here, it is unable to account for the *latent opacity* in our data.

It is worth mentioning that according to Wolf (2011), mutual counterfeeding can be accommodated in OT-CC if a different version of PREC constraints is assumed: each constraint of the format PREC(A,B) would be split into *B-THEN-A, which penalizes violation of A after violation of B, and A←B, which penalizes violation of B without preceding violation of A. Wolf argues that in this case the ranking metaconstraint should only apply to constraints of the A←B type. In our case, ranking *MAX(C)-THEN-MAX(V) above *UNSTRV in addition to having *MAX(V)-THEN-MAX(C) ranked above *FINAL-C would correctly derive the mutual counterfeeding interaction, while MAX(C)←MAX(V) and MAX(V)←MAX(C) do not play a crucial role in the analysis and can be ranked lower. Thus, with an important change to the tenets of OT-CC (cf. McCarthy 2007),

---

[36] This stands in contrast with previous accounts. For instance, Kavitskaya & Staroverov (2010) point to a ranking paradox in an OT-CC analysis of fed counterfeeding in Tundra Nenets, which leads them to propose markedness constraints whose violations depend on the current as well as the previous derivational steps, which they call Previous Step constraints. This, in turn, requires that PREC constraints be modified to contain an antifaithfulness requirement (E-PREC constraints). In the case of Gran Canarian we avoid the ranking paradox by using a context-free *UNSTRV rather than a contextual markedness constraint (*UNSTRV#), as shown in (15). Thus, fed counterfeeding can in principle be accounted for without any modification to the original OT-CC.

whose typological consequences have not been explored further, the analysis of our data would be technically possible as an alternative to SMR. Unfortunately, there is no available learner with which we could test whether the surface variation can be correctly generated.

5.1.2 Analysis of the data under contextual faithfulness constraints

Another alternative to using SMR was advanced by Hauser & Hughto (2020). In principle, it could be used in parallel OT, albeit only for counterfeeding interactions. However, we should bear in mind that Hauser & Hughto show that contextual faithfulness only works as a general solution for opacity when it is used in HS rather than parallel OT. The HS version of Hauser & Hughto's proposal cannot be easily implemented in our current learner due to the need for faithfulness constraints referring directly to the UR (Faith-UO; Hauser & Hughto 2020:§3.2). Nevertheless, in order to explore an alternative solution in parallel OT, we decided to consider a model with contextual faithfulness constraints. Interestingly, Hauser & Hughto state that their proposal is not fit for solving fed counterfeeding. In this context, we would like to show that contextual faithfulness does work for at least some fed counterfeeding cases, such as ours.

We explore Hauser & Hughto's model with the same constraints as in the *noSM* model in addition to two contextual faithfulness constraints, defined in (16).

(16)   Definitions of contextual faithfulness constraints, following Hauser & Hughto (2020)
         MAX/_V: Assign one violation mark for every input segment that is followed by a vowel in the input and has no output correspondent.
         MAX/_C: Assign one violation mark for every input segment that is followed by a consonant in the input and that has no output correspondent.

MAX/_V is violated when an input prevocalic segment is deleted, as is the case in the mapping /pasos/→[ˈpa]. This means that it can limit final consonant deletion to only consonants that never preceded a vowel in the input. Thus, the constraint can take over the function of SM(*FINAL-C,*UNSTRV) in the SMR analysis (§3.3.1).

MAX/_C is violated by final vowel deletion in C-final inputs (/pasos/→[ˈpas]) but not in V-final inputs (/paso/→[pas]), which means that it can ensure there are grammars where vowel deletion happens only in V-final inputs. This means it can take over the function of SM(*UNSTRV,*FINAL-C) in the SMR analysis (§3.3.2).

With the above constraints in place, surface forms can be generated successfully. The results of the simulations are presented in Appendix 6, showing that the model's accuracy is somewhere between our 1SM and 2SM models. Thus, even if this might not be the optimal version of a contextual faithfulness analysis, we can conclude that parallel OT and contextual faithfulness can be ingredients of an alternative analysis of our data.

5.2 Gran Canarian Spanish and the nature of opacity

Apart from showing unusual complexity in terms of opaque process interactions, our study also raises another important question: morphophonological restrictions. In OT, opacity is most often tied directly to cyclicity (Kiparsky 1971; 2000; Bermúdez-Otero 1999). Kiparsky (2015: 21) states explicitly that opacity is "a side effect of domain stratification" and that there are at most two levels of opacity corresponding to the changes in rankings between the three strata in Stratal OT opacity (Bermúdez-Otero forthcoming). Instead, we contend that the two processes involved necessarily act on the same stratum and morphological structure is not responsible for the opacity effect. Without any doubt, apocope should be assigned to the phrase-level stratum given its restricted application. However, positing C deletion at the word level is problematic because this process is in competition with other repair strategies: the final consonant can be devoiced or aspirated (if it is an *s*). Resyllabification is another complicating factor: word-final consonants tend to form an onset of the following word whenever the latter begins with a vowel. In the case of the *s*, a weakened variant is preserved in the newly formed onset (weak glottal [h], and in the Spanish from the Canary Islands, its voiced version, [ɦ], e.g. *los ejemplos* 'the examples' /los#exemplos/ [lo.ɦe.ˈɦem.plo]). Since deletion is avoided in resyllabification contexts, positing that deletion occurs at the word level, where no information concerning the following word is available to prevent unnecessary elision, is problematic. Thus, as far as Stratal OT is concerned, we are forced to assume that both deletion processes presented in §2 belong to the domain of the phrase (third stratum), and hence the problem of opacity cannot be solved.

Furthermore, Kiparsky (2015) argues that opacity should be investigated in obligatory processes only because with optional processes we cannot reliably establish whether the observed opacity effect is genuine or simply a result of not applying an optional process. However, our data show clearly that the opacity effect is caused by the non-application of consonant deletion after vowel apocope has taken place. Since vowel apocope is the undoubtedly optional process and consonant deletion practically always applies phrase-finally, we must conclude that the observed pattern is a genuinely opaque interaction. Taking the surface distributions into account, we can calculate the probability of each option. In words such as *pasos* the probability of [ˈpasos] is 8% while the probability of (transparent) [ˈpa] is 0% and the conditional probability of (opaque) [ˈpas] is 39%. In vowel-final words, the probability of (opaque) [ˈpas] is 61% while (transparent) [ˈpa] surfaces 0% of the time. Thus, mathematically speaking, the zero probability of transparent final C deletion cannot be derived from merely assuming that vowel apocope and final consonant deletion apply optionally at every derivational step: if the latter were the case, we would see at least some occurrences of forms like [ˈpa]. Consequently, we argue that opaque interactions within a stratum are not only possible but also quite productive across languages. Arguments against morphophonological explanations of opacity have been set forth based on examples from Catalan and Bedouin Arabic by McCarthy (2007: 40–41, 196–197). Similarly,

Broś (2016) reports a different case of post-lexical opacity in Spanish, and a recent contribution to the topic by Milenković (2022) shows an interaction of two non-optional lexical processes in a stratum-internal opaque interaction in Gallipoli Serbian. Moreover, the seminal case of fed counterfeeding in Tundra Nenets mentioned in this paper is also argued to be a within-stratum interaction (Kavitskaya & Staroverov 2010: 283). These pieces of evidence taken together make it necessary to adjust the formal mechanisms used to address the opacity problem in phonology and add weight to the discussion of the structural restrictions governing opaque vs transparent process interactions.

The opacity case presented in this paper also adds evidence to the fact that opaque interactions can be very diverse and have different implications depending on the type of processes involved, as well as the type of rules or constraints that can be used as a solution. More specifically, in §3.3 and §5.1 we demonstrated that our case of fed counterfeeding is less problematic in terms of formal analysis compared to previous accounts of similar interactions (e.g. Tundra Nenets).

In addition to the above, it must be stressed that analysing variation, apart from staying true to the actual productions of native speakers, has an additional advantage. As we have shown, certain ordering restrictions and rankings can only be found once variation is taken into account, which is an important contribution to the study of phonological interactions. §3 shows the roles both SM constraints play in our derivations. While SM(*Final-C,*Unstrv) appears to be the only serial markedness constraint necessary to derive each of the individual surface forms, including the fed counterfeeding pattern (Variant A), another high-ranked constraint SM(*Unstrv,*Final-C) is necessary when the quantitative aspect of variation is taken into account (Variant E). Such *latent opacity* effects need to be investigated further. Moreover, the effect of opacity versus different types of rankings necessary to derive surface forms should be mentioned in this context. Note that the different percentages of process application are independent from the attested counterfeeding interaction.[37] The fed counterfeeding opacity, for instance, concerns both C-final and V-final words regardless of the differing rates of process application. Nonetheless, the latter lead to the discovery of the need to construct an additional ranking for the probabilistic grammar. Thus, the existence of a possibility that the two interacting processes apply differentially, i.e. that not only both apply, both fail to apply or one of them fails to apply, but that one of them may apply only if a preceding process did not apply, is a potential challenge for the formal representation of the dialect. The option in which *pasos* is pronounced

---

[37] In fact, it is conceivable that an otherwise transparent pattern might exhibit latent opacity. Suppose that a language has lexically assigned final or penult stress, and there is an optional process of final stress retraction to the penult, which transparently feeds an optional process of unstressed final vowel deletion: /ˈmana/ → [ˈmana ~ ˈman]; /maˈna/ → [maˈna ~ ˈmana ~ ˈman]. In this case, the rate of unstressed final vowel deletion among penult stress forms may differ between underlyingly penult stress words and retracted final stress words, just like it does in Gran Canarian between V-final and C-final words, which would be a latent opacity effect in an otherwise transparent pattern.

['ˈpaso] but *paso* is pronounced [ˈpas] requires a framework that goes beyond mimicking extrinsic rule ordering and derivational steps, with or without reranking. Note that even if we assumed that apocope and consonant deletion apply at different strata, a Stratal OT approach and the like would fail to predict such outputs. Reranking constraints responsible for the occurrence of either of the processes is not enough without a probabilistic component in the grammar. Consequently, our data show that optionality leads to variation that obscures possible analyses, which is of consequence for phonological theory and should be considered in future research linking language variation and change with phonological computation.

Finally, we have seen that a single variety of a language can show not one but two complex opacity cases that are presumably typologically rare. Especially, mutual counterfeeding has been questioned as a linguistic reality. Wolf (2011) discusses one possible case of /ə/-syncope and VN coalescence in Hindi-Urdu, which has been contested in the literature. Exchange rules can be listed as another possible mutual counterfeeding pattern (see Wolf 2011: 103–106 for a review). The present study adds yet another example which, in our opinion, is difficult to dismiss. Thus, the Gran Canarian data contribute to the discussion on the typology of opacity and raise the question of whether more mutual opacity cases might be encountered cross-linguistically as more research is done into optional processes and the resultant variation.

6. Conclusion

In this paper, we have shown a case of advanced lenition in the form of variable phrase-final deletion of both final consonants and vowels in Gran Canarian Spanish. Of the two processes, word-final consonant deletion applies in more environments and is produced by all speakers, while vowel apocope is more restricted, both in terms of context and frequency of occurrence, and in terms of language users. The interaction of the two processes produces a special case of opaque variation in the dialect, which involves fed counterfeeding and a latent effect in the form of mutual counterfeeding. The latter results from a different behaviour of vowels in C-final vs. V-final stems. Against this background, we have shown that the output forms can be successfully generated using Serial Markedness Reduction, without the need for any additional types of constraints (like those proposed by Kavitskaya & Staroverov 2010). Furthermore, we presented a solution for generating variation with *latent opacity*, using simulations in a dedicated Expectation-Driven Learning algorithm (Jarosz et al. 2018). Our results show that complex opaque interactions and variation can be jointly modelled in a probabilistic constraint-based framework. They also show that looking into optional processes and variation may be necessary to uncover latent opacity interactions that encourage further development of theories of opacity.

Appendix 1. Examples of phrases with and without syllable apocope contexts

Speaker: Cr, aged 25

[1] 16.01–34.78 s, sound file 2 – *Bueno me apunté a la academia, pero me he **lesionado** **[le.sjo. ná]** (apocope context countervened by intervocalic *d* deletion and vowel simplification). *Estuve un año y pico allí, para nada porque no salieron **plazas** **[plá.sa]** (apocope context, only consonant deletion occurs). *Estaba perdiendo el tiempo **prácticamente** **[pɾa.ti.ka.mént]** (apocope context, apocope occurs), *y luego intenté hacer un par de ciclos **superiores** (no context, rising intonation), *que no me salieron hasta que encontré en el que estoy, de de energía renovable que es **futuro** **[fu. dú.ɾo̞]** (apocope context, incomplete apocope occurs: devoicing).

'Well, I signed up for the academy, but I got **injured**. I was there for a year and a bit, and for nothing because then there were no **jobs**, no, I was wasting my time **practically.** And then I tried to do a couple of advanced **vocational** courses which did not go well until I found the one I'm doing right now, one on on renewable energy, which is the **future**.'

[2] 36.71–39.74 s, sound file 2 – *Ahora **estamos**...* **[eh.tá.mo]** (rising intonation, no context, only consonant deletion) *el otro día montamos un panel **solar*** **[so.lá]** (final vowel stressed, only consonant deletion).

'Now we are... the other day we fixed a **solar** panel'.

[3] 63.81–69.86 s, sound file 2 – *No hacíamos nada **porque**...* (no context, information incomplete, hesitation) *en la programación estaba puesto de que al final no había **taller*** **[ta.jéɾ̥]** (falling intonation but final vowel is stressed, context for deletion but there is only *r* devoicing), *hasta segundo **año*** **[áɲ]** (apocope context, apocope occurs).[38]

'We weren't doing anything because... in the syllabus it said that all in all there was no **workshop**, until the second **year'**.

[4] 29.89–33.66 s, sound file 4 – *Y es una persona que si dice que el cielo es verde no se lo discutas porque el cielo es **verde*** **[βéɾtʰ]** (apocope and devoicing plus aspiration of the final [d]).

'And it's a person that when he says that the sky is green don't argue with him because the sky is **green'**.

---

[38] Here, note that the palatal nasal is considered a complex segment in Spanish.

Appendix 2. Variable rule analysis of our data.

A variable rule analysis of the Gran Canarian data is possible, but only as long as multiple copies of a rule are allowed in the grammar, and disjunctive blocking (Anderson 1992) is possible. We can model our data as in **Table 10**, where the first copy of vowel deletion (VD) is in a disjunctive blocking relationship with consonant deletion (CD): if you can apply VD, do so; otherwise, apply CD if possible. As shown in this table, choosing these application for each rule yields the attested frequencies in the data. Importantly, if multiple copies of the same rule are not allowed, the same problem arises as with the OT-CC account: there is no way to generate a preference of [ˈpaso] over [ˈpas] for /pasos/ and the opposite preference for /paso/. Since we are interested in an OT-based account, we will not pursue this account any further.

UR	/pasos/			/paso/		
<VD: V → ∅ / _# (p = .36),	--			pas (p = .36)	-- (p = .64)	
CD: C → ∅ / _# (p = .92)>	paso (p = .92)		-- (p = .08)	N/A (disjunctive blocking)	--	
VD: V → ∅ / _# (p = .39)	pas (p = .39)	-- (p = .61)	--	--	pas (p = .39)	-- (p = .61)
SR	[pas] (p = .92 × .39 = .36)	[paso] (p = .92 × .62 = .56)	[pasos] (p = .08)	[pas] (p = .36 + .64 × .39 = .61)	[paso] (p = .64 × .61 = .39)	

Table 10: Derivations for the words *pasos* and *paso* in a probabilistic rule framework.

Appendix 3. Summary of SMR rankings necessary for obtaining all 5 surface variants.

Only the crucial rankings of the constraints used in the simulations are listed for each of the variants.

Variant A. /pasos/ → [ˈpas], /paso/ → [ˈpas]
MAX(V)/INITIAL, CONTIG, SM(*FINAL-C,*UNSTRV) >> *UNSTRV >> *FINAL-C >> MAX(seg)
& *UNSTRV >> SM(*UNSTRV,*FINAL-C)

Variant B. /pasos/ → [ˈpasos], /paso/ → [ˈpaso]
MAX(V)/INITIAL, CONTIG >> MAX(seg) >> *UNSTRV, *FINAL-C
SM(*FINAL-C,*UNSTRV) & SM(*UNSTRV,*FINAL-C) ranked freely

Variant C. /pasos/ → [ˈpaso], /paso/ → [ˈpaso]

Max(V)/Initial, Contig >> *Final-C >> Max(seg) >> *UnstrV

SM(*Final-C,*UnstrV) & SM(*UnstrV,*Final-C) ranked freely

Variant D. /pasos/ → [ˈpasos], /paso/ → [ˈpas]

Max(V)/Initial, Contig >> *UnstrV >> Max(seg) >> *Final-C

SM(*Final-C,*UnstrV) & SM(*UnstrV,*Final-C) ranked freely

Variant E. /pasos/ → [ˈpaso], /paso/ → [ˈpas]

Max(V)/Initial, Contig, SM(*Final-C,*UnstrV), SM(*UnstrV,*Final-C) >> *UnstrV >> *Final-C >> Max(seg)

Appendix 4. Contents of E-step and M-step for Expectation Driven Learning (Jarosz 2015)

The E-step calculates the expected frequency of each pairwise ranking given the current grammar (G) and the data corpus (D): this can be thought of as "the current best estimate of how often this ranking must have been used by the speaker in generating the data corpus". It first estimates, for each pairwise ranking and each attested mapping, how likely this pairwise ranking A >> B is to yield a given mapping d given the current grammar by using a sampling procedure: out of r sample rankings (we chose the standard setting $r = 50$), the algorithm counts under how many rankings the attested mapping wins, which is the number of matches for that ranking and that attested mapping given the current Pairwise Ranking Grammar G: $m(A>>B, d \mid G)$. The sampling procedure is repeated for each attested mapping in the data and each possible pairwise ranking, which then allows the learner to calculate a set of new pairwise ranking probabilities given the data and the current grammar. This is done by first calculating the probability of each pairwise ranking given the attested mapping using Bayes' rule, (17a), plugging in the probability of the pairwise ranking in the current Pairwise Ranking Grammar, G, into the formula as $P(A>>B \mid G)$. Subsequently, the algorithm computes the expected frequency of each pairwise ranking for the entire dataset. $E(A>>B|D,G)$, as in (17b). The A>>B probabilities for each mapping are summed together, weighted by the frequency of that mapping. In our case, it was assumed that each mapping had a frequency of 100 times the proportion of that particular mapping in the corpus (e.g., the frequency of /pasos/ → [ˈpasos] is 100 * 8% = 8). If the same input maps to different outputs at different frequencies, the conflicting ranking preferences of each mapping will each contribute to an overall ranking preference, weighted by each mapping's frequency. Then, during the M-step, new pairwise ranking probabilities given the entire dataset are computed by normalizing the expected frequencies of each pairwise mapping between both rankings of the constraints involved, as in (17c). These probabilities are inserted into the updated grammar (G_{t+1}), and the cycle starts again until the specified stopping criterion is reached (for our simulations, this means reaching a fixed number of iterations, namely 15).

(17)　*Formulas for updating a Pairwise Ranking Grammar from* G *to* G_{t+1} *in EDL*

a.　$P(A \gg B \mid d, G) = \dfrac{m(A \gg B, d \mid G) \times P(A \gg B \mid G)}{m(A \gg B, d \mid G) \times P(A \gg B \mid G) + m(B \gg A, d \mid G) \times P(B \gg A \mid G)}$

b.　$E(A \gg B \mid D, G) = \sum_{d \in D} freq(d) \times P(A \gg B \mid d, G)$

c.　$P(A \gg B \mid G_{t+1}) = P(A \gg B \mid D, G) = \dfrac{E(A \gg B \mid D, G)}{E(A \gg B \mid D, G) + A(B \gg A \mid D, G)}$

Appendix 5. Calculations for obtaining the data

Tables 11–16 summarize the resulting grammars for each model. Inputs that behave similarly (e.g. V-final inputs, the C-final inputs) are grouped together. The frequency of each output candidate is calculated as follows. The learner calculates probabilities for each input-output mapping in the data based on 1000 samples, and lists the probability of every other candidate that was generated in the process. Since the same input occurred in multiple mappings, this means that there are multiple estimates for the probability of every candidate. When estimating the grammar's prediction of the frequency of a particular candidate at a particular run, we averaged all probability estimates of that candidate and multiplied the result by 100. These numbers are the basis of the numerical results in §4. The predicted frequencies column in **Tables 11–16** shows the mean of these numbers, averaged over all 10 runs, as well as the range (minimum and maximum) of the predicted frequency for that mapping across all 10 runs.

Input	Output	Frequency, model's average prediction (95% CI)	Frequency, attested
/paso, paharo/	[ˈpaso,ˈpaharo]	51 (51–52)	39
	[ˈpas,ˈpahar]	20 (20–20)	61
	(other)	29	0
/pasos, paharos/	[ˈpasos,ˈpaharos]	28 (28–29)	8
	[ˈpaso,ˈpaharo]	43 (43–43)	56
	[ˈpas,ˈpahar]	0	36
	(other)	29	0
/metro, oferta/	[ˈmetro,oˈferta]	51 (51–52)	39
	[ˈmetr,oˈfert]	48 (47–48)	61
	(other)	1	0
/metros, ofertas/	[ˈmetros,oˈfertas]	28 (28–28)	8
	[ˈmetro,oˈferta]	43 (43–43)	56
	[ˈmetr,oˈfert]	27 (26–27)	36
	(other)	2	0

Table 11: Full results for noSM model.

	*FINAL-C	*UNSTRV	MAX	MAX(V)/INITIAL	CONTIG
*FINAL-C		0.53	0.65	0.11	0.07
*UNSTRV	0.47		0.95	0	0
MAX	0.35	0.05		0.02	0.01
MAX(V)/INITIAL	0.89	1	0.98		0.44
CONTIG	0.93	1	0.99	0.56	

Table 12: Resulting PRG for 1st of 20 runs for NoSM model.

Input	Output	Frequency, model's average prediction (95% CI)	Frequency, attested
/paso, paharo, metro, oferta/	['paso,'paharo,'metro, o'ferta]	49 (49–49)	39
	['pas,'pahar,'metr,o'fert]	50 (49–50)	61
	(other)	1	0
/pasos, paharos, metros, ofertas/	['pasos,'paharos,'-metros,o'fertas]	11 (11–11)	8
	['paso,'paharo,'metro, o'ferta]	45 (45–45)	56
	['pas,'pahar,'metr,o'fert]	42 (42–42)	36
	(other)	2	0

Table 13. Full results for 1SM model.

	*FINAL-C	*UNSTRV	MAX	SM(*FINAL-C, *UNSTRV)	MAX(V)/INITIAL	CONTIG
*FINAL-C		0.47	0.89	0.05	0.07	0.02
*UNSTRV	0.53		0.93	0.12	0	0
MAX	0.11	0.07		0.03	0.01	0
SM(*FINAL-C, *UNSTRV)	0.95	0.88	0.97		0.4	0.25
MAX(V)/INITIAL	0.93	1	0.99	0.6		0.37
CONTIG	0.98	1	1	0.75	0.63	

Table 14: Resulting PRG for 1st of 20 runs for 1SM model.

Input	Output	Frequency, model's average prediction (95% CI)	Frequency, attested
/paso, paharo, metro, oferta/	[ˈpaso,ˈpaharo,ˈmetro,oˈferta]	41 (41–41)	39
	[ˈpas,ˈpahar,ˈmetr,oˈfert]	57 (57–57)	61
	(other)	2	0
/pasos, paharos, metros, ofertas/	[ˈpasos,ˈpaharos,ˈ-metros,oˈfertas]	12 (12–12)	8
	[ˈpaso,ˈpaharo,ˈmetro,oˈferta]	52 (52–52)	56
	[ˈpas,ˈpahar,ˈmetr,oˈfert]	32 (33–34)	36
	(other)	4	0

Table 15: Full results for 2SM model.

	*FINAL-C	*UNSTRV	MAX	SM (*UNSTRV, *FINAL-C)	SM (*FINAL-C, *UNSTRV)	MAX(V)/ INITIAL	CONTIG
*FINAL-C		0.37	0.9	0.47	0.04	0.06	0.01
*UNSTRV	0.63		0.92	0.56	0.13	0.01	0
MAX	0.1	0.08		0.18	0.02	0.01	0
SM (*UNSTRV, *FINAL-C)	0.53	0.44	0.82		0.21	0.12	0.03
SM (*FINAL-C, *UNSTRV)	0.96	0.87	0.98	0.79		0.39	0.21
MAX(V)/ INITIAL	0.94	0.99	0.99	0.88	0.61		0.37
CONTIG	0.99	1	1	0.97	0.79	0.63	

Table 16: Resulting PRG for 1st of 20 runs for 2SM model.

Appendix 6. Simulations with contextual faithfulness constraints

The parallel OT model with MAX/_V and MAX/_C was run with the same parameters of the same learner as in the SMR simulations above (20 runs of 15 iterations and with all other parameters

being the same as well). The numerical results are similar to the 1SM and 2SM models for the SMR simulations: the MAE is between the values for those two models, though the log-likelihood is lower than that of either model, with non-overlapping CIs.

MAE	4.1 (3.9-4.7)
Log-likelihood	–6.676 (–6.690; –6.632)

Table 17: Numerical results of simulations using contextual faithfulness constraints. Numbers averaged across 20 runs (95% CI in parentheses).

Qualitatively speaking, the resulting grammars match the data distribution very well for inputs with one unstressed vowel (/paso(s), metɾo(s)/), but are further off for inputs with multiple unstressed vowels (/paxaɾo(s), ofeɾta(s)/), for which they predict a significant presence of outputs with vowel deletion outside the attested word-final position (['paxɾ(s), 'feɾt(s)]), and this is even though we include both CONTIG and MAX(V)/INITIAL. The parallel OT setup puts more candidates in direct comparison, which could be one of the contributing factors to this result. However, this shows that a parallel OT account can in principle be considered for these data, as long as there are constraints that distinguish between deleting before an underlying vowel or consonant versus deleting the final segment of the underlying form, provided that there is a mechanism to account for variation. As mentioned above, however, the parallel OT version of Hauser & Hughto's account was not presented as a serious proposal for accounting for opacity in general, which is why we have not presented it alongside our main analysis, while we are currently unable to test the serial analysis due to difficulty implementing Faith-UO (see above). A different Parallel OT account that can capture opacity (e.g., Boersma 2007; Van Oostendorp 2008) might be explored in future research.

(18)   Hasse diagram of rankings found for the Parallel OT learner
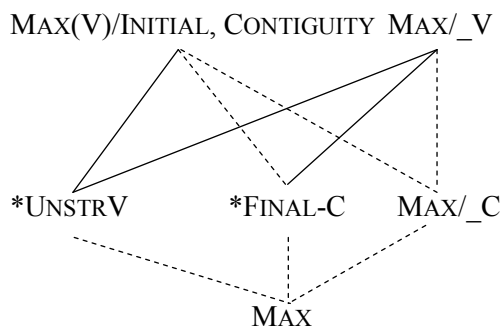
Input	Output	Frequency, model's average prediction (95% CI)	Frequency, attested
/paso, metɾo/	[ˈpaso,ˈmetɾo]	38 (38–38)	39
	[ˈpas,ˈmetr]	61 (61–62)	61
	(other)	1	0
/pasos, metɾos/	[ˈpasos,ˈmetɾos]	13 (13–13)	8
	[ˈpaso,ˈmetɾo]	57 (56–57)	56
	[ˈpas,ˈmetr]	29 (29–29)	36
	(other)	1	0
/pahaɾo, ofeɾta/	[ˈpahaɾo,oˈfeɾta]	38 (38–38)	39
	[ˈpahaɾ,oˈfeɾt]	55 (55–56)	61
	(other)	7	0
/pahaɾos, ofeɾtas/	[ˈpahaɾos,oˈfeɾtas]	13 (13–13)	8
	[ˈpahaɾo,oˈfeɾta]	56 (56–57)	56
	[ˈpahaɾ,oˈfeɾt]	24 (24–24)	36
	(other)	7	0

Table 18: Full results for Parallel OT model.

	*UNSTRV	*FINAL-C	MAX	CONTIG	MAX(V)/ INITIAL	MAX/_V	MAX/_C
*UNSTRV		0.69	0.92	0.05	0.04	0.01	0.4
*FINAL-C	0.31		0.82	0.08	0.1	0	0.31
MAX	0.08	0.18		0.01	0.04	0	0.16
CONTIG	0.95	0.92	0.99		0.41	0.47	0.78
MAX(V)/ INITIAL	0.96	0.9	0.96	0.59		0.4	0.76
MAX/_V	0.99	1	1	0.53	0.6		0.89
MAX/_C	0.6	0.69	0.84	0.22	0.24	0.11	

Table 19: Resulting PRG for 1st of 20 runs for Parallel OT model.

Abbreviations

C – consonant

V – vowel

HS – Harmonic Serialism

OT – Optimality Theory

OT-CC – Optimality Theory with Candidate Chains

SMR – Serial Markedness Reduction

EDL – Expectation-Driven Learning

Funding information

Acknowledgements

Competing interests

The authors have no competing interests to declare.

Author contributions

The first author was responsible for data collection and phonetic analysis. The second author was responsible for running the simulations and overseeing the phonological analyses used in the paper. Both authors prepared the manuscript as well as formal analyses using SMR and other frameworks.

References

Almeida, Manuel, & Díaz Alayón, Carmen. 1988. *El Español de Canarias.* Santa Cruz de Tenerife.

Alvar, Manuel. 1972. *Niveles Socio-culturales en el Habla de las Palmas de Gran Canaria*. Las Palmas de Gran Canaria: Eds. del Cabildo Insular.

Anderson, Stephen R. 1992. *A-morphous Morphology*. (Studies in Linguistics 62.) Cambridge: Cambridge University Press.

Baković, Eric. 2011. Opacity and ordering. In Goldsmith, John A. & Riggle, Jason & Yu, Alan C. L. (eds.), *The Handbook of Phonological Theory, 2nd edition*, 40–67. Wiley-Blackwell, London. DOI: https://doi.org/10.1002/9781444343069.ch2

Beckman, Jill. 1998. *Positional Faithfulness*. Doctoral dissertation, UMass, Amherst.

Bermúdez-Otero, Ricardo. 1999. *Constraint Interaction in Language Change [Opacity and Globality in Phonological Change.]* PhD dissertation, University of Manchester / Universidad de Santiago de Compostela. www.bermudez-otero.com/PhD.pdf.

Bermúdez-Otero, Ricardo. forthcoming. *Stratal Optimality Theory*. The University of Manchester.

Boersma, Paul. 1998. *Functional Phonology*. PhD dissertation. Amsterdam: University of Amsterdam.

Boersma, Paul. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117. 1989–2054. DOI: https://doi.org/10.1016/j.lingua.2006.11.004

Boersma, Paul & Weenink, David. 2019. Praat: Doing phonetics by computer. Version 6.1.03. http://www.fon.hum.uva.nl/praat/.

Broś, Karolina. 2016. Stratum junctures and counterfeeding: Against the current formulation of cyclicity in Stratal OT. In Hammerly, Christopher & Prickett, Brandon (eds), *Proceedings of the Forty-Sixth Annual Meeting of the North East Linguistic Society*, Volume 1, 157–170. Amherst, MA: Graduate Linguistics Students Association.

Broś, Karolina. 2022. Lenition in contemporary speech from Gran Canaria: Two corpus case studies. *Phonica* 18. 60–85. DOI: https://doi.org/10.1344/phonica.2022.18.60-85

Broś, Karolina & Lipowska, Katarzyna. 2019. Gran Canarian Spanish non-continuant voicing: gradiency, sex differences and perception. *Phonetica* 76. 100–125. DOI: https://doi.org/10.1159/000494928

Broś, Karolina & Żygis, Marzena & Sikorski, Adam & Jan Wołłejko. 2021. Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. *Phonology* 38(1). 1–40. DOI: https://doi.org/10.1017/S0952675721000038

Browman, Catherine P. & Goldstein, Louis. 1990. Articulatory gestures as phonological units. *Phonology* 6. 201–251. DOI: https://doi.org/10.1017/S0952675700001019

Byrd, Dani. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57. 3–16. DOI: https://doi.org/10.1159/000028456

Dempster, Arthur & Laird, Nan & Rubin, Donald. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1). 1–38. DOI: https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Goldman, Jean-Philippe. 2011. EasyAlign: An automatic phonetic alignment tool under Praat. *Proceedings of Interspeech 2011*, 3233–3236. DOI: https://doi.org/10.21437/Interspeech.2011-815

Goldwater, Sharon & Johnson, Mark. 2003. Learning OT constraint rankings using a maximum entropy model. In Spenader, Jennifer & Eriksson, Anders & Dahl, Östen (eds), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. Stockholm: Stockholm University.

Hauser, Ivy & Coral Hughto. 2020. Analyzing opacity with contextual faithfulness constraints. *Glossa: a Journal of General Linguistics* 5(1). 82. DOI: https://doi.org/10.5334/gjgl.966

Jarosz, Gaja. 2014. Serial markedness reduction. In Kingston, John & Moore-Cantwell, Claire & Pater, Joe & Staubs, Robert (eds.), *Proceedings of the 2013 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.

Jarosz, Gaja. 2015. *Expectation Driven Learning of Phonology*. University of Massachusetts Amherst manuscript.

Jarosz, Gaja. 2016. Learning opaque and transparent interactions in Harmonic Serialism. In Hansson, Gunnar Ólafur & Farris-Trimble, Ashley & McMullin, Kevin & Pulleyblank, Douglas (eds.), *Proceedings of the 2015 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. DOI: https://doi.org/10.3765/amp.v3i0.3671

Jarosz, Gaja & Anderson, Carolyn & Lamont, Andrew & Prickett, Brandon. 2018. *Hidden Structure Suite: Version 3*. http://github.com/gajajarosz/hidden-structure

Kavitskaya, Darya & Staroverov, Peter. 2010. When an interaction is both opaque and transparent: The paradox of fed counterfeeding. *Phonology* 27. 255–288. DOI: https://doi.org/10.1017/S0952675710000126

Kiparsky, Paul. 1971. Historical linguistics. In Dingwall, William O. (ed.), *A Survey of Linguistic Science*, 577–642. College Park, MD: Linguistics Program, University of Maryland.

Kiparsky, Paul. 2000. Opacity and cyclicity. *Linguistic Review* 17. 1–15. DOI: https://doi.org/10.1515/tlir.2000.17.2-4.351

Kiparsky, Paul. 2015. Stratal OT: A synopsis and FAQs. In Hsiao, Yuchau E. & Wee, Lian-Hee (eds.), *Capturing Phonological Shades*. Cambridge Scholars Publishing.

Labov, William. 2001. *Principles of Linguistic Change.: Vol. 2. External Factors*. Oxford: Blackwell.

Legendre, Geraldine & Miyata, Yoshiro & Smolensky, Paul. 1990. Can connectionism contribute to syntax? Harmonic Grammar, with an application. In Ziolkowski, Michael & Noske, Manuela & Deaton, Karen (eds), *Proceedings of the 26th regional meeting of the Chicago Linguistic Society*, 237–252. Chicago: Chicago Linguistic Society.

McCarthy, John. 2003. OT constraints are categorical. *Phonology* 20(1). 75–138. DOI: https://doi.org/10.1017/S0952675703004470

McCarthy, John. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

McCarthy, John. 2008. The gradual path to cluster simplification. *Phonology* 25. 271–319. DOI: https://doi.org/10.1017/S0952675708001486

McCarthy, John & Prince, Alan. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In González, Mercé (ed.), *Proceedings of the Twenty-Fourth Meeting of the North East Linguistics Society*, Volume 2, 333–379. Amherst, MA: Graduate Linguistics Student Association.

Milenković, Aljoša. 2022. *Stratification versus gradualness: Opaque metrical structure in Gallipoli Serbian*. Paper presented at the 29th Manchester Phonology Meeting, 25–27 May, 2022.

Oftedal, Magne. 1985. *Lenition in Celtic and in Insular Spanish*. Oslo: Universitetsforlaget Oslo.

R Core Team. 2017. *R: A language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. http://www.r-project.org

Staubs, Robert & Pater, Joe. 2016. Learning serial constraint-based grammars. In McCarthy, John & Pater, Joe (eds.), *Harmonic Grammar and Harmonic Serialism*, 155–175. London: Equinox Press.

Tajchman, Gary & Jurafsky, Daniel & Fosler, Eric. 1995. Learning phonological rule probabilities from speech corpora with exploratory computational phonology. *33rd Annual Meeting of the Association for Computational Linguistics*, 1-8. Cambridge, MA: Association for Computational Linguistics. https://aclanthology.org/P95-1001. DOI: https://doi.org/10.3115/981658.981659

van Oostendorp, Marc. 2008. Incomplete devoicing in formal phonology. *Lingua* 118(9). 127–142. DOI: https://doi.org/10.1016/j.lingua.2007.09.009

Wolf, Matthew. 2011. Limits on global rules in Optimality Theory with Candidate Chains. *Phonology* 28(1). 87–128. DOI: https://doi.org/10.1017/S0952675711000042