Bogus Clusters and Lenition in Tuscan Italian: Implications for the theory of sonority[1]

Shanti Ulfsbjorninn
University of Lyon / UCL – London

Abstract

It is widely believed that syllabification is determined by a sonority-driven algorithm like the *Sonority Sequencing Principle* (Selkirk 1984; Clements 1990, Vaux and Wolfe 2009). In this study, I evaluate this claim in light of Tuscan Italian. Using three phonological diagnostics, it will be possible to split the consonant clusters (CCs) of Tuscan into three types: Branching onset, Coda-Onset and Bogus clusters. *Metrical lengthening* and *Gorgia Toscana* filter out Branching onsets leaving behind Coda-Onset and Bogus clusters as remnant. Elsewhere, the process of *Epenthesis* (in non-standard dialects) filters out the Bogus clusters instead leaving Branching onsets and Coda-Onset clusters as remnant. Comparing the two sets of remnant allows for the extraction of the Coda-Onset set. Using a Sonority Differential analysis (Parker 2011), it becomes evident (process by process) that sonority is not the primary (or a preferable) mechanism in determining these sets. In seeking an alternative analysis, *Gorgia Toscana* will be presented in some detail along with its implications for sonority. *Gorgia* underapplies with Bogus clusters. I will provide a suggestive sketch for a competing representational solution based on Strict CV (in particular Lowenstamm 2003 and Brun-Trigud & Scheer 2010). Informed in part by interlude theory (Steriade 2008), it offers an alternative account for the lenition facts: compressible CCs (Branching onsets) are equivalent to a singleton stop, while non-compressible clusters (Coda-Onset and Bogus clusters) are equivalent to geminates. Unlike sonority based analyses, the phonological definition of the clusters offered here has a clear relationship with the phonological processes that occur to them.

Introduction

The structure of the study is as follows. Section one begins by describing the core organisational principle of Tuscan syllabification: *Metrical lengthening*. This process cuts the CCs of Tuscan into two sets. Section one also presents the Sonority Differential analysis and shows that the two sets of CCs defined by *Metrical lengthening* are not predicted by sonority. Section two then introduces the process of *Gorgia*, the distinctive lenition pattern of Tuscan Italian. Tuscan deaffrication is also presented as part of the same process. This section shows that as it is standardly described, the rules that are thought to govern *Gorgia* badly mispredict the underapplication of lenition in a certain group of clusters. Once the context of the process is correctly defined, one arrives at a richer collection of the same two CC classes that were identified in Section one (call them Set A and Set B CCs). Again, sonority is shown to be problematic because it cannot define the sets.

[1] Thanks to John Harris, Andrew Nevins, Michael Kenstowicz, Bridget Samuels, Joaquim Brandão de Carvalho, Noam Faust, Monik Charette, Donca Steriade, Ricardo Bermúdez-Otero, Michela Russo, Ioana Chitoran & Florian Breit. Needless to say they may not agree with all or any of what is written here.

Section three then takes a third process: Epenthesis, found in non-standard variants of Italian, but also in Brazilian Portuguese. This process reveals that the previously identified Set B actually breaks down into two smaller sets. Section four takes stock of the situation and taking an overview of the processes identifies and formally defines the *three* sets of CCs necessary in Tuscan: Branching onsets, Coda-Onset and Bogus clusters. Armed with this tripartite distinction, we end the section with an alternative representational solution to Tuscan syllabification based on compressibility. Branching onsets are reanalysed as essentially singletons and Bogus clusters are immune to *Gorgia* in essentially the same way as geminates. A summary of the CC types identified, their status viz. compressibility and their interaction with the three processes ends the paper.

1 Metrical lengthening restrictions and sonority

Tuscan Italian (Tuscan) is a generous language. It richly supplies the learner with phonological evidence for the syllabification of its words. The form of this evidence is robust and categorical, there are essentially no exceptions.

The central organising principle of Tuscan syllable structure relates to the interaction of quantity and stress. Vowels of stressed open syllables must be long in all but final position: (a) [ví:peɾa] *[vípeɾa] 'viper', (b) [muɾé:na] *[muɾéna] 'moray eel', and (c) [kaɾitá] *[kaɾitá:] 'charity'.[2] This interaction is known as *metrical* or *tonic lengthening* (cf. Hayes 1995). In Tuscan, all vowel length is a product of *Metrical lengthening*. Long vowels are otherwise illicit: [vipeɾ-átʃ:a] *[vi:peɾ-átʃ:a] 'bad viper' (cf. [ví:peɾa]) and [alabá:ma] *[a:labá:ma] 'Alabama' (cf. àlabáma *Eng*).[3]

The effect of heterosyllabic syllable structures on metrical lengthening is demonstrated by geminates. Stressed vowels cannot be long before geminate consonants. The placement in the word and the sonority of the geminate are not variables.

(1) Length is blocked before heterosyllabic structures

	Antepenultimate		Penultimate	
Stop	[dzɔ́k:olo]	'clog'	[maɾmɔ́t:a]	'marmot'
Affricate	[kjɔ́tʃ:ola]	'snail'	[ɾítʃ:o]	'hedgehog'
Nasal	[míɲ:olo]	'little finger'	[kán:a]	'spliff'
Lateral	[libél:ula]	'damselfly'	[koɾál:o]	'coral'
Rhotic	[tɔ́r:iðo]	'scorching'	[vér:e]	'male fox'

Geminates are unambiguously heterosyllabic. They are one feature bundle attached to two positions in the syllable structure (Hayes 1989, Kenstowicz 1994). Specifically, one of their

[2] This is different from Standardised Tuscan, a variety with metrical lengthening only in penultimate position (Savoia 2014).

[3] The two examples here are supposed to give the 'best chance' for lengthening, one is a derived form with vowel length in a related paradigm. The second is a loanword where the source language provides clear evidence of two trochee heads in open syllables.

positions is located in the stressed rime that precedes them (cf. Hayes' (1989) flopped structure). Metrical lengthening is therefore understood as a condition on the size of the stressed rime. Stressed rimes in Italian must contain exactly two positions: VC, VV.[4]

Unlike geminates that never permit vowel length before them, consonant clusters have variable patterning. CCs split into two types. There are those that behave like heterosyllabic geminates: [mósto, ménta, pólpo, tʃɛ́rvo] 'must, mint, octopus, deer', and those that behave like singletons: c[á:]pra, v[é:]tro, s[ɔ:]brio, c[í:]clo 'goat, glass, sober, cycle'. The two full sets are presented in (2) beneath. Immediately after, in (3), it will be shown that sonority cannot be used to distinguish the two sets.

(2) CCs and vowel lengthening[5]

(A) Permit length

(i)	pr, br	c[á:]pra	'goat'	c[ɔ:]bra	'cobra'
(ii)	tr, dr	m[í:]tra	'machine gun'	c[é:]dro	'cedar'
(iii)	kr, gr	[á:]cro	'acrid'	p[í:]gro	'lazy'
(iv)	kl, gl	c[í:]clo	'cycle'		

(B) Ban length

(a) Rhotic

(i)	rp, rb	s[ɛ́]rpe	'snake'	m[ɔ́]rbo	'illness'
(ii)	rt, rd	[á]rte	'art'	s[ó]rdo	'deaf'
(iii)	rtʃ, rdʒ	[á]rci	'association'	[ɔ́]rgia	'orgy'
(iv)	rk, rg	p[ɔ́]rco	'pig'	[ɔ́]rgano	'organ'
(v)	rm	[ó]rma	'print'		
(vi)	rn	k[ɔ́]rno	'horn'		
(vii)	rl	m[ɛ́]rlo	'blackbird'		

(b) Lateral

(i)	lp, lb	[á]lpi	'Alps'	[á]lbo	'alb'
(ii)	lt, ld	c[ó]lto	'educated'	c[á]ldo	'hot'
(iii)	ltʃ, ldʒ	c[á]lce	'chalk'	[á]lgida	'brand'
(iv)	lk, lg	[á]lcol	'alcohol'	[á]lga	'algae'
(v)	lf, lv	[á]lfa	'alpha'	b[é]lva	'beast'
(vi)	rf, rv	sm[ɔ́]rfia	'scowl'	c[ɔ́]rvo	'crow'
(vii)	lm	k[ó]lmo	'full'		
(viii)	ln	[ú]lna	'ulna'		

[4] Another common way of expressing this is: rimes must be bimoraic.

[5] Only two-member clusters are discussed. Three member clusters are of no specific interest to the questions raised by this study. This is probably not a fully exhaustive list of two member clusters, but it must approach exhaustivity.

(c) Nasal

(i)	mp, mb	c[á]mpo	‘field’	g[á]mbo	‘stalk’
(ii)	nt, nd	d[ɛ́]nte	‘tooth’	[ó]nda	‘wave’
(iii)	ntʃ, ndʒ	g[á]ncio	‘hook’	sf[í]nge	‘sphynx’
(iv)	ŋk, ŋg	st[í]nco	‘shin’	v[ó]ngola	‘clam’
(v)	nf	t[ó]nfo	‘heavy fall’		
(vi)	ns	t[ó]nsilla	‘tonsil’		

(d) Fricative

(i)	sp	r[ɔ́]spo	‘toad’
(ii)	st	t[ɔ́]sto	‘tough’
(iii)	sk	d[í]sco	‘disco’
(iv)	fk	k[á]fka	‘kafka’

(e) Stop

(i)	ps	[í]psilon	‘Y’
(ii)	tl	[á]tlas	‘atlas/map’
(iii)	tn	[ɛ́]tna	‘Etna’

While the consonant clusters in (2A) form a natural class from a sonority point of view, the same cannot be said for the set in (2B). I will demonstrate this using the sonority differential analysis inspired by Parker (2011). This is an analytic technique that establishes a metric for sonority in clusters. It works by subtracting the sonority value of C1 from C2. The sonority value is arbitrarily assigned according to an arbitrary scale.

(3) Sonority differential

$$C2 - C1 = \Delta x$$

(4) Sonority Hierarchy

Glide	8
Rhotic	7
Lateral	6
Nasal	5
Fricative	4
Affricate	3
Stop	2

According to the scale and equation shown in (3, 4), the Tuscan CCs permitting lengthening all have the same sonority profile: ‘step rises’. They are exclusively made up of a stop and a liquid. This allows for a natural class that can be pointed at and where that natural class can be defined by sonority. But we will see the problems begin almost immediately after.

(5) Sonority differential of CCs permitting length

(a) Steep Rises:

(i)	Stop - Rhotic	pr	::	r |7| - p |2|	=	Δ5
(ii)	Stop – Lateral	kl	::	l |6| - k |2|	=	Δ4

Conversely, the CCs that do not permit lengthening do not form a natural class, they include sonority plateaus, small sonority rises and even steep sonority rises.

(6) Sonority differential of CCs banning length

(a) Steep falls

(i)	Rhotic – Stop	rp	::	p |2| - r |7|	=	Δ-5
(ii)	Rhotic – Affricate	rtʃ	::	tʃ |3| - r |7|	=	Δ-4
(iii)	Lateral – Stop	lp	::	p |2| - l |6|	=	Δ-4
(iv)	Lateral – Affricate	ltʃ	::	tʃ |3| - l |6|	=	Δ-3

(b) Shallow falls

(i)	Fricative – Stop	st	::	t |2| - s |4|	=	Δ-2

(c) Near flat

(i)	Lateral – Nasal	ln	::	n |5| - l |6|	=	Δ-1
(ii)	Rhotic – Lateral	rl	::	l |6| - r |7|	=	Δ-1

(d) Shallow rises

(i)	Stop – Fricative	ps	::	s |4| - p |2|	=	Δ2

(e) Steep rises

(i)	Stop – Lateral	tl	::	l |6| - t |2|	=	Δ4

This rather simple demonstration is enough to demonstrate that a sonority-based syllabification algorithm simply does not cut the mustard (or more specifically the CCs) into sonority natural classes that match their phonological behaviour.[6]

[6] One way to solve the problem might be to split out the /ps/, /tl/ etc... clusters and add *non-sonority based conditions* to account for the whole system. No doubt this can be done, but this analysis will necessarily be disjunctive. It will use sonority for the parts of the analysis where it works and use

2 *Gorgia Toscana* and sonority

The discussion on metrical lengthening and consonant clusters in section one reveals that sonority cannot make the appropriate cut between those CCs that permit vowel length before them and those that do not.

In this section we will look at a second phonological process of Tuscan: *Gorgia Toscana*. This process also provides clear, categorical, and exceptionless evidence for the language learner, and it also splits the CCs into sets. Yet again, sonority will not be able to define the sets. Crucially, the sets of CCs identified by *Gorgia Toscana* contain the same members as the sets identified by Metrical lengthening. *Gorgia* adds new members to the sets, but there are no mismatched or overlapping CCs across the two sets. This means that two apparently unrelated processes split Tuscan CCs into exactly the same sets. This permits extra confidence in supposing a representational difference underlying the two sets.

2.1 Description of *Gorgia*

Gorgia is a highly distinctive lenition process of Tuscan Italian where stops are spirantised. The nature of the process is uniform across Tuscany, though there are different lenition outcomes in different areas. However, the relationship of *Gorgia* to CCs is essentially uniform.

Transcriptions in this section are of the full Maremmano lenition, but exactly the same discussion could have been framed with reference to the better known Florence/Siena. Current Maremmano and Florence/Siena are essentially identical but for the outcome of the voiceless dorsal stop).

(7) *Gorgia Toscana* distribution and realization across Tuscany

(a) Massa-Carrara/N Lucca	–	none	
(b) Arezzo	–	Restricted to /k/	
(c) Pisa/Livorno	–	Restricted to /k, t, d/	
(d) Florence/Siena	–	Full: /p, b, t, d, k, g/	› /ɸ, β, θ, ð, h or ɦ, ɣ/
(e) Grosseto (Maremmano)[7]	–	Full: /p, b, t, d, k, g/	› /ɸ, β, θ, ð, x, ɣ/ or... /^{p}ɸ, ^{b}β, t^{θ}, d^{ð}, x, ɣ/[8]

alternatives where it stops working. Clearly this analysis cannot not be favoured if there is a competing analysis that accounts for the whole system.

[7] *Pace* Marotta (2008) who claims (based on old data) that Maremmano (spoken in most of the province of Grosseto) patterns with Livorno and Pisa. It is possible that *Gorgia* has been extended in recent decades, but essentially all stops undergo Gorgia in some way.

[8] Marotta (2008) also reports fricativised stops as lenition outcomes in Pisa.

Gorgia is typically described as a post-vocalic process (Marotta 2008) because it applies both to intervocalic singleton stops: /libél:ula/ 'damselfly' › [liβél:ula], and post-vocalic branching onsets: /pregjéɾa/ 'prayer' › [laɸreɣjé:ɾa].

Gorgia applies post-lexically, both within and across word-boundaries: [kré:ma] 'cream' › [laxré:ma] 'the cream' › [na-βél:a-xré:ma] 'a nice cream'. It is exceptionless in neologisms and loanwords: *crackers* › [i-xré:xe(ɾ)] 'the crackers'. In these forms, it has precisely the same phonological environment as native words, *Gorgia* is not triggered post-consonantally: *Merkel* › [laméɾkel] 'the Merkel (German chancellor)'.

As Marotta (2008) notes, *Gorgia Toscana* is intimately connected to another distinctive post-lexical Tuscan process: inter-/post-vocalic de-affrication.

(8) Tuscan de-affrication

(a)	[tʃé:ɾa]	'wax'	›	[la-ʃé:ɾa]	'the wax'
(b)	[dʒóɾno]	'day'	›	[i-ʒóɾni]	'the days'

Geminate stops are inalterable and do not undergo *Gorgia*, just as geminate affricates are immune from de-affrication: [grá**k**:jo] 'alpine chough', [stra**tʃ**:o] 'rag'. This is an observation that will be exploited in section 4.

3.2 Defining the triggering context of *Gorgia*

Because of its application to branching onsets: /grák:jo/ 'alpine chough' › [i-ɣrák:ji] 'the alpine choughs', the post-vocalic *Gorgia* rule could be standardly described as post-vocalic:

(9) *Gorgia Toscana I (to be rejected)*

[+cons, -son, -cont] → [+cont] / [-cons, +cont] __

The rule in (9), however, makes a significant misprediction that has hitherto gone unreported in the literature on *Gorgia*. A whole subset of CCs fail to undergo the process even though the conditions are met.

(10) Failure to apply *Gorgia*

Cluster	Initial?	Example	Counterfactual	Gloss
tl	*no*	[a**tl**éta]	*[a**θl**éta]	'athelete'
tn	*no*	[é**tn**a]	*[é**θn**a]	'Etna'
tm	*only*	[la-**tm**é:si]	*[la-**θm**é:si]	'tmesis'
pn	*yes*	[a**pn**é:a] [lo-**pn**eumá**θ**i**x**o]	 *[lo**ɸn**eumá**θ**i**x**o]	'breathlessness' 'the tire'
ps	*yes*	[ípsilɔn] [lo-**ps**i**x**oanalista]	*[í**ɸs**ilɔn] *[lo**ɸs**i**x**oanalista]	'Y' 'the psychoanalyst'
kt	*no*	[í**kt**us]	*[í**xt**us]	'stroke'
pt	*only*	[lo-**pt**e**ɾ**o**ð**at:ilo]	*[lo**ɸt**e**ɾ**o**ð**at:ilo]	'the pterodactyl'

This is a significant misprediction. One way to salvage the approach might seem to be to redefine the conditioning environment of *Gorgia* so that the rule applies in inter-sonorant-continuant context (rather than post-vocalic).

(11) *Gorgia Toscana II (to be rejected)*

[+cons, -son, -cont] → [+son, +cont] / [+son, +cont] __ [+son, +cont]

This rule would allow spirantisation in 'vowel __ vowel *or* rhotic' sequences but one would have to remove [cons] from the rule. It would also be important to define nasals and laterals as [-cont]. This is not particularly problematic but it is important or laterals and nasals would also trigger spirantisation.[9] Relatedly, however, the rule in (11) makes a further misprediction. If rhotics are special in being sonorant continuants in Tuscan and *Gorgia* applies in an intersonorant context, then the rule predicts that spirantisation should apply symmetrically in 'rhotic__vowel' sequences. However, these contexts do not trigger *Gorgia* (or deaffrication): [kɔ́ɾ**p**o] *[kɔ́ɾ**ɸ**o] 'body', [sɔ́ɾ**tʃ**o] *[sɔ́ɾ**ʃ**o] 'mouse'.

Descriptively then, in addition to the intervocalic context, when it comes to rhotics, it seems that *Gorgia* is only permitted if the rhotic *follows* the stop. It is possible to write this in the rule, but (at this point) the description of the facts makes it clear that a linear rule should be

[9] This actually *does* happen in Spanish for laterals (not nasals) and non-coronal obstruents: [al**ɣ**o] 'other'.

replaced with a syllable structure analysis. *Gorgia* applies to the stops of certain CCs when these are intervocalic.[10,11]

The CCs undergoing and blocking *Gorgia* are split into sets, these correspond to the sets derived by *Metrical lengthening*, with the addition of new members to Set A (cf. 2A).[12]

(12) CCs and *Gorgia*

(A) Undergoing *Gorgia*

		Non-*Gorgia*		***Gorgia***	**Gloss**
(a)	pr	[**pr**é:ɣo]	'pray'	[le**ɸr**é:ɣjé:ɾe]	'the prayers'
(b)	pj	[**pj**é:**n**a]	'full river'	[la**ɸj**é:**n**a]	'the full river'
(c)	pl	[**pl**ak:a]	'plaque'	[la**ɸl**ák:a]	'the plaque'
(d)	br	[**br**ú:xo]	'worm'	[i**βr**ú:xi]	'the worms'
(e)	bj	[**bj**áʃ:ixa]	'chew slowly'	[le**βj**áʃ:ixaθe]	'the chewings'
(f)	bl	[**blát:a**]	'bug/cockroach	[le**βl**át:e]	'the bugs'
(e)	tr	[**tr**á:ʃina]	'weever (fish)'	[leθ**r**á:ʃina]	'the weevers'
(f)	dr	[**dr**úm:e]	'tobacco brand'	[ilmi**ðr**úm:e]	'my tobacco'
(g)	kr	[**kr**e**ɸ**a]	'fold/crack'	[la**xr**e**ɸ**a]	'the crack'
(h)	kj	[**kjó:**ma]	'tree top'	[la**xjó:**ma]	'the tree tops'
(i)	kl	[**kl**ás:e]	'class'	[le**xl**ás:i]	'the classes'
(j)	gr	[**gr**ú]	'crane'	[le**ɣr**ú]	'the cranes'
(k)	gj	[**gj**andá:ja]	'acorn'	[le**ɣj**andá:ja]	'the acorns'
(l)	gl	[**gl**ú:θei]	'glutes'	[i**ɣl**ú:θei]	'glutes'

(B) Resisting *Gorgia*

Same as set in B of *Metrical lengthening* (2B), plus the examples shown in (13).

[10] We have focused on branching onsets with rhotics so far because the branching onsets with laterals iodatised between Latin > Italian. However, Italian did reintroduce very many stop-lateral branching onsets from Latin as 'learned lexicon'. These stop-lateral sequences lenite exactly like stop-rhotic sequences.

[11] The facts could lead to still further modification of the linear rule. However, it's clear that this would simply be a notational variant of a syllable structure.

[12] There is a larger number of cluster types because Gorgia affects all the word's positions, including initial so there is a larger set of contexts to test. But no CC of Set A defined by metrical lengthening is a Set B consonant as defined by *Gorgia*.

Crucially, sonority cannot be used to generate the two sets of clusters. This is because although all the CCs undergoing *Gorgia* have steep rises in sonority, the CCs that do not undergo *Gorgia* do not form a homogenous natural class, at least from a sonority perspective.

(13) Sonority differential and *Gorgia*

(a) Type A - CCs undergoing *Gorgia*

(i)	Stop - rhotic	=	Δ5
(ii)	Stop - lateral	=	Δ4

(b) Type B II - CCs not undergoing *Gorgia*

Cluster	Sonority Differential	Sonority Profile	Example	Gloss
kt	Δ0	Flat	[í**kt**us]	'stroke'
pt	Δ0	Flat	[lo-**pte**roðat:ilo]	'the pterodactyl'
ps	Δ2	Shallow rise	[ípsilɔn] [lo-**psix**oanalista]	'Y' 'the psychoanalyst'
tm	Δ3	Steep rise	[la-**tm**é:si]	'tmesis'
pn	Δ3	Steep rise	[a**pn**é:a] [lo-**pn**eumá**θix**o]	'breathlessness' 'the tire'
tn	Δ3	Steep rise	[é**tn**a]	'Etna'
tl	Δ4	Steep rise	[a**tl**éta]	'athelete'

Sonority is again not managing to cut the CCs into the appropriate sets.

Descriptively, it appears that Type A CCs are a restrictive set and can be defined in syllable structure terms as 'Branching onsets'. Meanwhile, Type B CCs are a sonority-varied 'bin'. On this view, they would be, in effect, coda-onset sequences. In order to have this 'bin' and describe the facts, one must abandon the notion that sonority is chiefly responsible for syllabifying CCs.

(14) Type A – Branching Onsets

(a) Stop – Liquid (lateral or rhotic) (excluding 'tl' 'disjunctively' for different reasons)

(15) Type B – Bin of varying sonorities (to be elaborated on)

(a) Liquid – Stop
(b) Liquid – Affricate
(c) Liquid – Fricative
(d) Liquid – Nasal
(e) Rhotic – Lateral

(f) Nasal – Stop
(g) Nasal – Affricate
(h) Nasal – Fricative

(i) Fricative – Stop

(j) Stop – Fricative
(k) Stop – Liquid
(l) Stop – Nasal
(m) Stop – Stop

Hiving off a subset of CCs as Branching Onsets is the correct approach. However, to define the 'bin' of various sonorities 'codas' would be independently problematic. Making Type B set into coda-onset clusters is unfortunate from a Tuscan perspective because of the high priority the language seems to devote to excluding non-continuant obstruents. A part from geminates, stops are allowed in coda position[13]: *stop* (Eng) › [stɔ́p:e] 'stop', OKTU(M) (Lat) › [ot:o]. Moreover, it seems unpleasant typologically, especially considering Harris (1997) and Ségéral & Scheer (2001), and Scheer & Ziková (2008), to propose lenition processes specifically targeting onsets while leaving codas intact. Worst of all, it fails to model the typological and dialectal observation that the 'bin' set is not made up of one syllable structure type. This evidence come from a third process labelled *Epenthesis*. It is discussed in the next section.

3 Epenthesis and sonority

There is evidence in dialectal/non-standard Italian, and from typological observation of other romance languages, that the Type B set is not uniform from a phonological perspective and instead it further splits into two.

[13] The generalisation could be stated as follows: a non-continuant root node cannot occupy a coda position if the same root node does not also occupy an onset.

Non-Branching onset CCs are broken up by *Epenthesis* in non-standard Italian (Repetti 1993) and also in Brazilian Portuguese (Nevins 2005). However, *Epenthesis* does not target all members of what we have called Type B, only the stop-initial subset.[14]

(16) *Epenthesis* by sound class of 'Type B'

		Epenthesis	Example
(a)	Liquid – Stop	no	
(b)	Liquid – Affricate	no	
(c)	Liquid – Fricative	no	
(d)	Liquid – Nasal	no	
(e)	Rhotic – Lateral	no	
(f)	Nasal – Stop	no	
(g)	Nasal – Affricate	no	
(h)	Nasal – Fricative	no	
(i)	Fricative – Stop	no	
(j)	Stop – Fricative	yes	[pis]icologa
(k)	Stop – Liquid	yes	[ati/e]las
(l)	Stop – Nasal	yes	[ati]mosfera
(m)	Stop – Stop	yes	[pe/Ø]terodattilo

3.1 Phonological processes and CC classes

We can identify three classes of consonant cluster by summarising what has been learned from examining each of the three phonological processes. Each phonological process can be seen to be filtering CCs for the structural change that they cause. The remainder, objects not marked for a structural change are the remnant. Comparing the filtered items from the remnant for each of the three process, shows that in Tuscan Italian there are three separate CC representations.

(17) Process and its rejects

	Target/Filtered	**Remnant**
(a) *Metrical Lengthening*	Branching onset	Rime-onset, Bogus cluster
(b) *Gorgia*	Branching onset	Rime-onset, Bogus cluster
(c) *Epenthesis*	Bogus cluster	Rime-onset, Branching onset

From comparing these processes, their targets and the remnant, one can find the unique sets of CCs. *Metrical lengthening* and *Gorgia* filters the first type: Branching onsets. *Epenthesis* filters the second set, what Harris (1994) calls: Bogus clusters. Finally, comparing the remnant of *Metrical Lengthening* and *Gorgia* against that of *Epenthesis* allows one to finds the unique set of rime-onset sequences.

[14] This seems to suggest that 'Bogus clusters' in Tuscan and Brazilian Portuguese derivationally 'Branching onsets that failed to be compressed'. However, Bogus clusters are *a priori* unrestricted.

4 Defining the cluster sets and explaining Gorgia

Sonority is a useful heuristic. It helps to problematize the resulting CCs sets from our processes. It was very helpful in demonstrating the need to continue splitting the CCs into groups because - for better or for worse - sonority is helpful in anecdotally locating natural classes.

However, from a generative or learner's perspective, the phonological processes that splits the CCs into sets exists independently. Also, the aetiology behind the CC sets is founded in a more basic (and concrete) explanation than sonority. Therefore, sonority is not formally required at any level of the explanation.

The three cluster types now require defining phonologically and, ideally, this definition will be connected to the kinds of phonological processes that identify them. Because ideally there should be a connection between the phonological object's definition and its behaviour (and effect on other phonological objects surrounding it).

4.1 Branching onsets and compressibility

The three types of CC are as follows: Branching onsets, Coda-Onset clusters, and Bogus clusters. Taking the Branching onset first, we notice a convergence of two distinct phonological traditions. One is the treatment of Branching onsets in Strict CV (Lowenstamm 2003), the other is Steriade's (2008) Interlude theory. Both hint at an account where branching onsets are distinctive because they are compressible.

Steriade's interlude theory of weight comes about from a description of Ancient Greek and Latin meter. Steriade observes that weight categories in these languages are based in part on the entire consonant interlude (CI), where CI is defined as the interval between two vowels. Steriade's analysis requires positing two conditions on weight, one syllabic and the other based on CI, independent of syllable structure affiliation. Steriade shows that word-initial consonant clusters split into two kinds those with long CI (which are a natural class with into two types: (a) compressible (stop-liquid sequences and (in Greek) voiceless stop-nasal sequences) and (b) incompressible (the rest).

Applied to Tuscan, preliminary measurements show that consonant clusters are indeed split in terms of their CI. Branching onsets are compressible and bogus clusters are not. As shown in (18) beneath, Branching onsets have essentially the same duration as singletons, while the CI of bogus clusters is like that of geminate consonants. Interestingly, unlike Steriade's finding in Greek, voiceless stop-nasal clusters in Italian have a long CI.

(18) CI by consonant cluster type in Tuscan

Cluster type	Sonority	CI Duration (ms)	CI	Natural class	Token	Gloss
Singleton	n/a	82	short	**A**	trɔ́:**[θ]**a	trout
Geminate	n/a	210	long	B	tʃí**[t:]**o	boy
PR	steep +	98.8	short	A	ká:**[ɸ]**ra	goat
KL	steep +	110	short	A	tʃí**[kl]**i	cycles
KT	flat	240	long	B	í**[kt]**us	stroke
PS	shallow +	270	long	B	í**[ps]**ilon	Y
TN	steep +	220	long	B	é**[tn]**a	Etna[15]
PN	steep +	260	long	B	a**[pn]**é:a[16]	breathlessness
TL	steep +	210	long	B	á**[tl]**as	book of maps

These findings are harmonious with the observation from phonological behaviour. For all intents and purposes, Tuscan Branching onsets behave as if they are monopositional. That is, they appear to have their composite linked to a single skeletal/syllabic position (Lowenstamm 2003, Brun-Trigaud & Scheer 2010). From a *Metrical lengthening* perspective, the distribution of Branching onsets is identical to that of singleton consonants. Likewise, with respect to *Gorgia*, their phonological behaviour is identical to that of singleton stops.

Combining these phonological findings with the durational findings from (18), a picture emerges where the structure of Branching onsets is analogous to singleton consonants. Meanwhile, the Bogus clusters, including ones with rising sonority profiles: [ps, tl, pn, tm, ...] cannot be compressed and so are analogous to geminates. The underapplication of *Gorgia* in Bogus clusters seems to have the same explanation as the underlying causes of geminate inalterability.[17]

[15] It is considering clusters like /tn, tm, pn/ that we realise why this account does not need a special treatment for clusters that violate OCP for Coronal: /tl, tn/. The Bogus clusters listed here with nasals constitute (rather) steep rises of sonority but they cannot nonetheless be compressed as branching onsets. /tl/ also cannot be compressed so compressibility is the first level explanation for lack of lenition in Tuscan (grouping together tl, tn, tm, pn. For that reason the grammar does not have to directly consider an OCP-COR constraint (or equivalent). Spirantisation acts on monopositional stops, however they're derived.

[16] This is not attested post-tonically, unlike the other tokens it was measured pre-tonically

[17] In Strict CV terms, stops are spirantised when they are governed by a rightwards (non-governed) V. The stop of Branching onsets is still local to its nucleus (V) because the 'root nodes' of the branching onset are compressed into a single timing slot.

(19) Singleton and Branching onset vs. Geminate

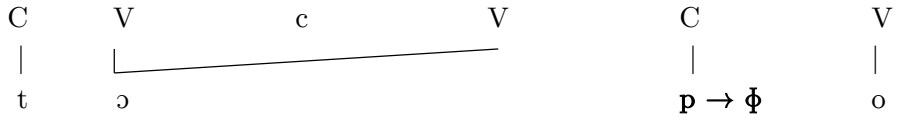
(a) Metrical lengthening and Gorgia

	Singleton		**Geminate**	
	/tɔ́**p**o/	[untɔ́:**ɸ**o]	/pá**p:**a/	[laɸá**p:**a]
	mouse	a mouse	slurry	the slurry
	Branching onset		**Bogus cluster**	
	/prima/	[la**ɸr**í:ma]	/pneo/	[ʎi**pn**é:i]
	first	the first	tire	the tires
Duration (approx.)	100ms		200ms	

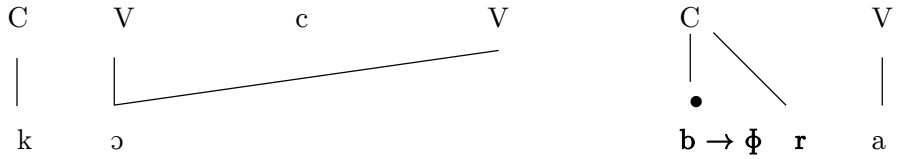
In traditional terms then *Gorgia* is still intervocalic, rather than being an post-vocalic process. Branching onsets are monopositional and essentially count as singletons.

In terms of phonological representation therefore, the members of a branching onsets are shown to slot into a single C position. As is shown in (20-21), it is the phonological structure of a Branching onset (vs. Bogus cluster) that makes inherent sense of its interaction with phonological processes. A sonority-based analysis cannot offer such a match-up between structure and behaviour under a process. For example, the structure of the Branching onset is like a singleton consonant, all its parts are attached to a single timing slot. This allows lengthening before it, just like the singleton. Consequently, both singletons and Branching onsets are monopositional and show no resistance to Gorgia (see (20) vs. (21)). The phonological structure of Branching onsets proposed by Lowenstamm (2003) and adopted by Brun-Trigaud & Scheer (2010), and Lai (2016) is shown in the following representations (small caps c and v stand for phonetically unrealised skeletal positions).

(20) Singleton spirantising with metrical lengthening (*tópo* 'mouse')

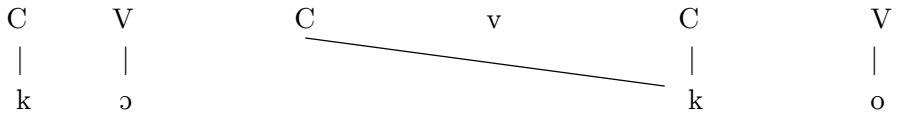


(21) Compressed Branching onset are short like singleton (*cóbra* 'cobra')

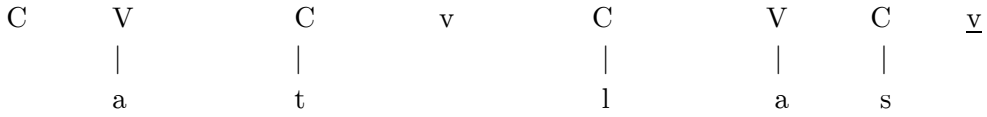


Unlike what is shown above, the representations of geminates and Bogus clusters both show that there is no 'space' before them to host lengthening (so they resist *Metrical lengthening*), and they are both resistant to *Gorgia*.

(22) Geminate length resists lenition (*cócco* 'coconut')



(23) Uncompressed Bogus clusters are 'big' like geminates (*átlas* 'book of maps')



I now turn to explaining the difference between the other two types of CC. The big difference between ordinary CCs like Coda-Onset clusters and Bogus clusters lies in the local interaction (or lack thereof) of the members of the cluster. In Branching onsets and Coda-Onset clusters there is a set of licensing relationships holding between the two members of the cluster. While in Bogus clusters there is no such relationship. They are not even phonologically 'clusters', merely adjacent filled onsets straddling an empty nucleus (V slot).

The precise conditions holding within clusters do not concern us here, they can be found in Harris (1990 and 1994), as well as Charette (1990). The importance is the difference between the two kinds of cluster, because it is what makes the essential difference between Bogus clusters and other CCs with regard the processes of lenition and especially *Epenthesis*.

It suffices, at this point, to say that Coda-Onset clusters are bound together by licensing obligations that are irrelevant to Bogus clusters. In the classic reference to Bogus clusters, this lack of relationship between the members of cluster also has implications for lenition.

The argument is from Harris (1994) and concerns English t-glottaling. Harris notes that t-glottaling is blocked from applying in branching onsets, although the processes' conditions would otherwise be met: petrol › [pɛ́tɹu] *[pɛ́ʔɹu] 'petrol' (cf. [kɛ́ʔəɹɪŋ] 'kettering'). This restriction, however, has no bearing however, on 'tr' sequences formed by syncope: [bǽʔɹi] 'battery'. This is taken to show that a 't' of a Branching onset ([pɛ́**t**ɹu] 'petrol') must remain intact in order to remain a head governing a dependent (and be at least as equally complex as the dependent (Harris 1990). Meanwhile in the Bogus cluster [bǽ**ʔ**ɹi] /bǽʔəɹi/ 'battery', the 't' and 'r' are essentially invisible to each other. Therefore, the 't' can lenite without impacting the status of a dependent.[18]

[18] Considering the behaviour of [tʃ] in the same leniting environment that suggests another possible analysis: [mǽtʃːəu] 'macho' vs. [mətʃízməu] 'machismo'.

In Tuscan, although the actual outcome is the inverse (Bogus clusters are the ones that do not lenite[19]), the presence or absence of a licensing relationship between the members of the clusters is what makes the phonological difference in epenthesis.

Though both cluster types have intervening empty V slots (a) Coda-Onset: ca/rvp/a *carpa* 'carp' vs. (b) Bogus cluster: i/pvs/ilon *ipsilon* 'stroke'. The v-slot of Coda-Onset sequences is firmly silenced by the licensing relationship that holds between the two consonants. The Bogus cluster, having no licensing relationship to tie them together is far more susceptible to being split up by *Epenthesis*: /pvs/icologa › [pis]icologa.

It is therefore the lack of licensing between the members of the 'cluster' that is responsible for its fragility and the epenthesis outcome. For some speakers, and presumably also in Brazilian Portuguese, any cluster without a licensing relationship between the members is subject to Epenthesis. This will leave the Coda-Onset clusters intact, but break the Bogus clusters.

[19] The conditions and outcomes of lenition are very different in the two languages.

4.2 Summary

Overall, the relationship between cluster type and process that they undergo has a natural or formal explanation. These are summarised in the following table.

(24) Cluster type, compressibility and process

			Process		
Cluster type	**Compressible**	**Licensed**	**Metrical Lengthening**	**Gorgia**	**Epenthesis**
Branching Onset pr, br, tr, dr, kr, gr, kl, gl	yes	yes	yes	yes	no
Coda-Onset rp, rb, rt, rd, rtʃ, rdʒ, rk, rg, rm, rn, rl, lp, lb, lt, ld, ltʃ, ldʒ, lk, lg, lf, lv, rf, rv, lm, ln, mp, mb, nt, nd, ntʃ, ndʒ, ŋk, ŋg, nf, ns, sp, st, sk, fk	no	yes	no	no	no
Bogus cluster ps, tl, tn, kt, pt, tm, pn, tn, tl	no	no	no	no	yes

5 Conclusion

Tuscan shows us that a sonority based approach is neither sufficient nor preferable as an explanation for syllabification. Using three phonological processes and diagnostics it was possible to split CCs in Tuscan into three types: Branching onsets, Coda-Onset and Bogus clusters. Throughout the paper it was demonstrated that sonority did not cut the CCs into the appropriate sets. In offering an alternative, this study presented *Gorgia Toscana* in some detail and a competing representational analysis based on compressibility was presented to account for it. Ultimately, *Gorgia* was resisted in the same manner essentially as geminate inalterability. Unlike sonority based analyses, the phonological definition of the clusters has a clear relationship with the phonological processes that occur to them.

References

Brun-Trigaud, G. & T. Scheer. 2010. Lenition in branching onsets in French and in ALF dialects. In: Karlík, P. (Ed.). *Development of Language through the Lens of Formal Linguistics*.15-28. Munich: Lincom.

Charette, M. 1990. License to govern. *Phonology* 7:233-253.

Clements, G. N. 1990. The Role of the Sonority Cycle in Core Syllabification. In: Kingston, J. & M. E. Beckman (Eds.). *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*. Cambridge (MA): Cambridge University Press. 283-333.

Harris, J. 1990. Segmental complexity and phonological government. *Phonology Yearbook* 7:255–300.

Harris, J. 1994. *English Sound Structure*. Oxford: Blackwell

Harris, J. 1997. Licensing Inheritance: an integrated theory of neutralisation. *Phonology* 14:315-370.

Hayes, B. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20:253-306.

Kenstowicz, M. 1994. *Phonology in Generative Grammar*. Cambridge (MA): MIT Press.

Lai, R. 2016. Complex segments in Sardinian: their origin, nature and distribution. Talk at: RFP 2016, Nice.

Lowenstamm, J. 2003. Remarks on Mutae cum Liquida and Branching Onsets. In: Ploch, S. (Ed.). *Living on the Edge, 28 Papers in Honor of Jonathan Kaye*. Berlin, New York: Mouton de Gruyter. 339-363.

Marotta, G. Lenition in Tuscan Italian (Gorgia Toscana). In: De Carvalho, J. B., Scheer, T. and P. Ségéral. (Eds.). *Lenition and Fortition*. Berlin: Mouton de Gruyter.

Nevins, A. (2007). Review of Tobias Scheer's 'A Lateral Theory of Phonology: What is CVCV and why should it be'. Lingua. 118(3):425-434.

Parker, S. 2011. Sonority. In: Oostendorp, M. van., Ewen, C, J., Hume, E. & K. Rice (Eds.). *The Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell. 1160–1184.

Savoia, L. 2014. *Introduzione alla Fonetica e alla Fonologia*. Milano: Zanichelli.

Scheer, T. and M. Ziková. 2010. Coda-Mirror V2. *Acta Linguistica Hungarica* 57(4):411–431.

Ségéral, P. & T. Scheer. 2001. La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris* 96:107-152.

Selkirk, E. 1984. On the major class features and syllable theory. In: Aronoff, M. & Oehrle (Eds.). *Language Sound Structure: Studies in Phonology*. Cambridge (MA): MIT Press. 107-136.

Steriade, D. 2008. Metrical evidence for an interlude theory of weight. Talk at: *Conference on the Syllable*. City University of New York.

Vaux, B. & A. Wolfe. 2009. The Appendix. In: Raimy, E. and C. Cairns (Eds.). *Contemporary Views on Architecture and Representations in Phonology*. Cambridge (MA): MIT Press. 101-143.