Variation and change in Italian phonology: On the mutual dependence of grammar and lexicon in Optimality Theory

Martin Krämer

UiT The Arctic University of Norway

Abstract

In this paper I discuss the influence of language acquisition and borrowing on the reorganisation of grammar and lexicon in the development from Latin into Italian. We will have a look at the historical sequencing of the introduction of new phonological processes, velar palatalization, mid vowel breaking, and lateral palatalization, and how they conspire to create new contrasts or reintroduce contrasts that have been subject to neutralisation. The amphichronic analysis proposed here brings together insights from acquisition and loanword phonology in Optimality Theory to explain the historical development.

Keywords: Italian phonology, diachronic microvariation, language change, palatalization, Optimality Theory, acquisition

1 Introduction

The repeated introduction of new phonological processes with a partial overlapping of targets and outputs leads to an increase in the degree of opacity in the grammar up to a peak at which learners radically restructure both the grammar as well as underlying representations. The selected processes from the history of Italian to be discussed here, show how counter-feeding opacity arises and how it leads to the innovation of contrasts.

Optimality Theory (Prince & Smolensky 1993/2004, McCarthy & Prince 1995 et seq.) sees the differences between languages as an epiphenomenon of different constraint rankings. Likewise, the differences between historical stages of a single language have to be seen in this way (e.g., Cho 1998, Holt 2003, Bermúdez-Otero 2006, 2007). The diachronic stages of a language vary by small differences in constraint ranking, just as dialects synchronically vary in this way. We are thus dealing with microvariation in time here. Unlike with dialectal variation, the question arises how one ranking turns into another, i.e., how (and why) do rankings change, a question that has not been satisfactorily answered yet.

It is by now a widely held assumption that language change at least partially emerges through imperfect learning by new generations of speakers (McMahon 2000, Bermúdez-Otero 2013). We will have a look at

the Biased Constraint Demotion Algorithm (Tesar & Smolensky 2000, Prince & Tesar 2004 inter alia), the currently predominant approach to language learning in OT, and how this can help explain selected historical processes in Italian. In addition, it will be shown here, that some historical changes are caused by adults, who introduce new loanwords, and thereby render the grammar inconsistent for the next generation of learners. Adults and children (i.e., learners) deal with grammatical inconsistency in different ways. While adults can cope with exceptions more easily, children try harder to make a consistent generalisation. Technically this is modelled by the introduction of lexically indexed Faithfulness constraints (Ito & Mester 1999, 2001, Pater 2000, 2006, 2009). This paper is a contribution to the amphichronic program (Kiparsky 2006, Bermúdez-Otero 2014), since it explains diachronic variation by looking at the grammars of individual historic stages and linking them through language acquisition and borrowing.

The paper is structured as follows. In the next subsection I will give some empirical background on the Latin and Italian segment inventories. In 1.2 the phonological processes are introduced and brought into a historical order. Section 2 provides the theoretical background. I will introduce the basic learning algorithm, the idea of phonological opacity as a motor of contrast innovation, and finally the technical tools of constraint indexation and conjunction, which will be utilized in the analysis of

exceptional loanwords and opacity, respectively. Section 3 puts together the results of 1.2 with the machinery of section 2 to provide an explanatory amphichronic (Kiparsky 2006, Bermúdez-Otero 2014) account of the historical development.

1.1 The Latin and Italian segment inventories

Even though Italian is considered the closest offspring of Latin, the differences in both inventories and phonotactics are quite dramatic. While Italian has lost three of the Latin consonants (the laryngeal fricative and the labialized dorsals), it has also extended the voicing contrast to the fricatives and, more importantly, introduced a whole new series of palatals and affricates. Compare (1) and (2). The segments missing in the respective other inventory are given in boldface.

(1) Latin contrastive consonants (Vincent 1988a: 29)

Labial	Alveolar/ Dental	Velar	Labio-velar	Glottal
p b	t d	k g	**k^w g^w**	
f	s			**h**
m	n			
	l			
	r			

(2) Italian contrastive consonants (Krämer 2009)

	Labial	Dental	Alveolar	Palatal	Velar
Stop	p, b	t, d			k, g
Affricate			**ts, dz**	**tʃ, dʒ**	
Fricative	f, **v**		s, **(z)**	**ʃ**	
Nasal	m		n	**ɲ**	
Lateral			l	**ʎ**	
Rhotic			r		

The development is not less fascinating in the vowel and diphthong system. While the inventory of plain vowels has been shrunk considerably, abandoning the length distinction and allowing a tenseness contrast only in

mid vowels, all three Latin diphthongs have been eliminated and a whole new range of diphthongs has been introduced.

(3) a. Latin vowels

iː	ɪ				ʊ	uː	
	eː	ɛ			ɔ	oː	
			aː	a			

b. Latin Diphthongs

ai	(orthographic AE)
au	
oi	(orthographic OE)

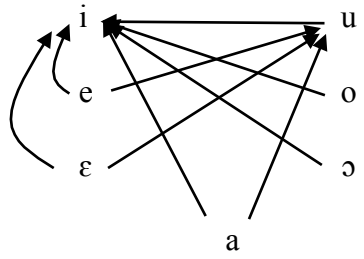
(Vincent 1988a: 29)

(4) Italian simple vowels (Krämer 2009a)

		i				u	
			e			o	
				ɛ		ɔ	
					a		

(5) Italian falling diphthongs (Krämer 2009a)

a.

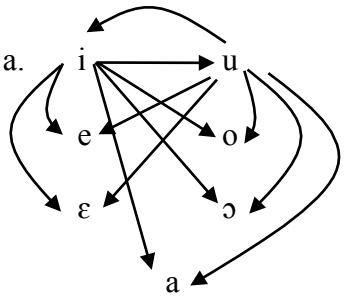


b.

	i	u
i	*ij	*iw
e	ej	ew
ɛ	ɛj	ɛw
a	aj	aw
ɔ	ɔj	*ɔw
o	oj	*ow
u	uj	*uw

(6)Italian rising diphthongs (Krämer 2009a)

a.



b.

	i	e	ɛ	a	ɔ	o	u
j		je	jɛ	ja	jɔ	jo	ju
w	wi	we	wɛ	wa	wɔ	wo	

It is a matter of debate whether all these diphthongs should be analysed as such. Criteria for diphthonghood could be that both vocoids are part of the same syllable constituent. Thus, if, for example, the glides in the rising diphthongs are part of the onset, this reduces the inventory considerably. See the discussion in Krämer (2009a) and references there. However, as

we will see shortly, some of these diphthongs are an important part of the puzzle about the introduction of the palatal series of consonants, regardless of their analysis.

1.2 Phasing processes

Palatalization came in two waves in Italian. The first wave affected all sorts of coronal and dorsal consonants before front vocoids.

(7) First palatalization

/li̯ → ʎʎ/	FILIA	figlia	[fiʎʎa]	‘daughter’
/le̯ → ʎʎ/	PALEA	paglia	[paʎʎa]	‘straw’
/lne̯ → ɲɲ/	BALNEU	bagno	[baɲɲo]	‘bath’
/ti̯ → tts/	UITIU	vezzo	[vettso]	‘habit’
/di̯ → ddz/	MEDIU	mezzo	[meddzo]	‘half’
/ki̯ → ttʃ/	ERICIU	riccio	[rittʃo]	‘hedgehog’
/ge̯ → ddʒ/	FAGEU	faggio	[faddʒo]	‘beech’
/ski̯ → ʃʃ/	FASCIARE	fasciare	[faʃʃaːre]	‘(to) bandage’
/rsi̯ → ʃʃ/	REVERSIARE	rovesciare	[roveʃʃaːre]	‘(to) reverse’

More palatal segments were created by the emergence of the coda condition, which resulted in coalescence of many marked coda consonants with the following onset.

(8) Parasitic onset licensing of features in coda position by fusion

LAXARE	lasciare	[laʃʃaːɾe]	‘(to) let’ (X = [ks])
LIGNA	legna	[leɲɲa]	‘wood’ (GN = [ŋn])

Second, or velar, palatalization affected only the dorsal stops and created more palatal fricatives and affricates.

(9) 2nd palatalization

a.	CIRCULUS	*circolo*	[tʃiɾkolo]	‘circle’
b.	GENTE	gente	[dʒɛnte]	‘people’
c.	PISCE	pesce	[peʃʃe]	‘fish’

Palatalization of dorsals also results in paradigmatic alternations, as illustrated in (10)a. However, in modern Italian we find many exceptions to the pattern. Verb stems in the 1st conjugation class never alternate (while 2nd and 3rd conjugation class stems always alternate) and nominals show mixed behaviour. I will come back to the discussion of the productivity of velar palatalization below in section 3.

(10) Morphophonological consequences of palatalization

a.	giun[dʒ]ere	giun[g]o	‘(to) join - (I) join’
b.	pa[g]are	pa[g]i	‘(to) pay - (you.sg) pay’

Even more palatals emerged from glide fortition. Since in these cases as in the coalescence cases above the trigger disappears and since many forms don’t show alternations that reveal the original or underlying form, as illustrated with a 2nd conjugation verb above, it becomes more and more likely that the palatal consonants are lexicalized.

(11) Glide fortition

a.	PEIUS	peggio	[ˈpeddʒo]	‘worse’
	IANUARIUM	gennaio	[dʒenˈna:jo]	’January’
b.	CIUILEM	civile	[tʃiˈvi:le]	‘civil’
	UINUM	vino	[ˈvi:no]	‘wine’

After the original glides were eliminated by fortition, new glides emerged through the diphthongisation of mid vowels. That modern Italian has mid vowels is owed to the smoothing of Latin diphthongs to mid vowels, as in *poco*. Smoothing of Latin AU (but not AE) thus must have happened after mid vowel diphthongization. The interesting development here is the

formation of front glides after dorsal stops (in cooperation with delabialisation of original /k^w/), which is the context of velar palatalization. A source of mid front vowels is the disappearance of the glide after diphthongization when it coalesced with the preceding stop to form the palatal affricate.

(12) Reorganisation of mid vowels and diphthongs

a.	ˈkwɔːko (COQUUS)	‘chef’	ˈpɔːko (PAUCUM)	‘not much’
b.	ˈnwɔːvo (NOUUM)		noviˈta	‘new / news’
c.	ˈgɔːdo	‘(I) enjoy’	goˈdeːre (GAUDERE)	‘enjoy-inf.’
d.	ˈɔːdo	‘(I) hear’	uˈdiːɾe (AUDIRE)	‘hear-inf.’
e.	ˈʧɛːko (CAECUS)	‘blind’		
f.	ˈkjɛːdeɾe (QUAERERE)	‘ask-inf.’	kjeˈdjaːmo	‘(we) ask’

Further instances of dorsals followed by palatal glides were created through the lenition of laterals in complex onsets. In some dialects, *l*-

weakening and palatalization apply transparently, as shown in (13) on the right. When the lateral is weakened into a palatal glide and preceded by a dorsal stop the context for palatalization is created. In standard Italian, the process underapplies, while in some dialects (e.g., Venetian), the dorsal stop and the palatal glide coalesce to the palatal affricate.

(13) Cl to Cj

Latin	*Italian*	*Dialectal*	*Gloss*
CLAUSTRUM	chiostro		‘cloister’
CLAUIS	chiave	[tʃ]av(e)	‘key’
ECCLESIA	chiesa	[tʃ/j]esa	‘church’
PLUVIA	pioggia		‘rain’
PLENUM	pieno		‘full’

There are several words that are exceptions to lateral weakening, as shown in (14). They are either newer loans from other languages (e.g., *club* ‘club’) or assumed to be re-borrowings from Latin (e.g., *classe*). The re-borrowing hypothesis finds additional support through doublets, such as *chierico* ‘cleric’ versus *clericale* ‘clerical’.

(14) (Re-)borrowing (from Latin)

classe ‘class’

flauto	‘flute’
flotta	‘fleet’
club	‘club’

Not only *l*-weakening creates exceptions to palatalization, borrowing does as well, as attested in words like *chitarra* ‘guitar’, which probably entered Italian from Greek via Arabic in the late middle ages. Compare *chitarra* and *civile* ‘civil’.

Tekavčić (1980) provides the following chronology. Metaphony (another source of diphthongs) applied already in Latin. First palatalization (before /j/) started in the 1st/2nd century, followed by AE smoothing, which was followed by mid vowel breaking in Late Latin. According to Tekavčić , second palatalization didn’t set in before the fifth century and Lateral palatalization (ClV->CjV) begins much later, in the second half of the 10th century.

(15) A chronology

1. 1st palatalization applies.
2. /kw/ reduces to /kʷ/
3. Diphthong AE and AU smoothen to mid vowels
4. Mid vowel breaking
5. 2nd palatalization

6. /k^{w}/ → /k/

7. Lateral palatalization

8. Cl clusters reimported, /ki-/ words reimported

In a word like ˈkjɛːdere (QUAERERE) neither palatalization process could apply, because there was a labial glide between the smoothened vowel and the dorsal stop. In CAECUS *cieco* 'blind', the diphthong turned into a mid vowel and there was no intervening labial glide that could stop palatalization of the dorsal. Thus, the dorsal palatalized either immediately or by the latest when the mid front vowel split into a diphthong again, this time [jɛ], rather than [ai]. Since we still find /k^{w}/, as in *questo* 'this', *qualità* 'quality', *questione* 'question' etc., I suspect that only those labial glides disappeared that had to make room for the palatal glide merging from diphthongization of mid front vowels. The laterals in complex onsets, e.g., ECCLESIA *chiesa* 'church', palatalized even later and didn't cause velar palatalization either. All the words that contain a consonant + lateral onset or a velar before a palatal vowel are assumed to have entered the language (via Latin, Greek and other languages) after all three palatalization processes.

In conclusion, the historical events constitute a case of counterfeeding opacity. Later processes create the environment for earlier processes, which fail to apply in these newly created environments. It has

to be assumed that there were phases in which this was also an opaque process interaction in the synchronic grammar, and that it is this opacification which resulted in a severely restricted environment of velar palatalization in present-day Italian.

Palatalization of the lateral in onset clusters doesn't cause any alternations, due to other constraints on Italian morphophonology. It could thus have applied throughout the whole lexicon and got switched off very fast again. This is supported by the many new words with consonant-lateral onset clusters. Velar palatalization on the other hand, causes regular alternations and non-alternating forms could, at least in some environments, take a free ride on the alternation to keep the palatal affricate out of the lexicon. A form that gives us a (weak) hint that second palatalization was still active even after lateral palatalization stopped is the word CYCLUS *ciclo* [tʃiklo]'circle', which underwent velar palatalization, but not lateral palatalization. It could have entered after lateral palatalization stopped, indicating that velar palatalization was still active and only stopped in word-initial environment in the middle ages, when the guitar *chitarra* appeared in Italy.

2 Theoretical background

In section 1.3 I concluded that successively introduced innovative processes that create the environment for older processes, as well as new borrowings, which tend to be exempt from regular phonological processes before they get fully nativized, create surface exceptions to a formerly productive exceptionless process. The process becomes either inactive or is marginalized to certain morphological contexts. In contemporary Italian, velar palatalization is marginalized to morpheme junctures, mostly in second and third conjugation verbs. Whether velar palatalization is productive in nouns and adjectives seems to be a personal decision every Italian has to make. Krämer (2009a,b) presented native speakers of Italian with nonce-words with root-final dorsal stops, which they had to use in a plural context (changing the ending from *–o* to *–i*, creating a palatalization context). Half of the group applied palatalization, the other half didn't.

2.1 Learning as constraint demotion

Bermúdez-Otero & Hogg (2003) discuss the actuation problem in OT, i.e., how does language change start or what is it that triggers language change. One answer to this question, though not an exhaustive one, is incomplete or imperfect acquisition of the grammar by a new generation of speakers. A classic example is r-intrusion in English. What started as r-dropping got

confused or reinterpreted by a new generation as r-insertion. This led to a surface change. Originally only prepausal and preconsonantal /r/ was dropped, which lead to paradigmatic alternations, as in *snore – snoring*. Once learners assume that non-alternating forms take a free ride on alternating forms and don't discover the changes this assumption causes in surface patterns early enough (as in *saw – saw*[r]*ing*), an insertion grammar emerges (see Krämer 2012). This is a case of reanalysis. However, dramatic changes can also arise by simple minor errors in constraint reranking.

In OT, language acquisition is a matter of finding the correct constraint ranking for the target language. There is relatively broad consensus that this development of the constraint hierarchy is brought about by constraint demotion (Tesar 1997, Tesar & Smolensky 2000, Prince & Tesar 2004).

(16) The Constraint Demotion Algorithm

Learners group the candidate set into winner-loser pairs (or mark-data pairs). For each such pair the learner registers for each constraint whether it supports the winner (W) or the loser (L), after all violation marks shared by both candidates have been cancelled out.

All L-marked constraints have to be dominated by at least one W-marked constraint. This is achieved by demotion of the L-marked constraints.

After every mark-data pair has been used and no ranking argument is left over the algorithm terminates.

The learning algorithm is error-driven. A learner only does this kind of operation on the grammar once she detects a mismatch between her own output and the perceived target output.

Furthermore, a learner is assumed to arrange the constraints in such a way that the grammar is as restrictive as possible. A way of measuring restrictiveness of a grammar in OT, is to count how many Markedness constraints dominate every Faithfulness constraint, introduced as the r-measure by Prince & Tesar (2004).

(17) The r-measure (Prince & Tesar 2004)

The r-measure for a constraint hierarchy is determined by adding, for each faithfulness constraint in the hierarchy, the number of markedness constraints that dominate that faithfulness constraint. (Tesar & Prince 2004: 252)

One would assume that the Richness of the Base (i.e., the learner's goal to exclude all candidates supplied by GEN[1] except for the desired winner) already makes sure that the grammar is maximally restrictive, however,

[1] GEN = The Generator function, which generates the set of candidate output forms.

constraint demotion alone doesn't lead to maximum r-measures, as Prince & Tesar note. To achieve maximum restrictiveness the algorithm has to be biased against Faithfulness constraints.

(18) Faithfulness Delay (Prince & Tesar 2004)

On each pass, among those constraints suitable for membership in the next stratum, if possible place only *markedness constraints*. Only place faithfulness constraints if no markedness constraints are available to be placed in the hierarchy.

As it happens, sometimes a learner has several Faithfulness constraints to choose from. This is the case when there is a positionally restricted clone of a general Faithfulness constraint. For example, when learning distinctions of place of articulation (PoA), and detecting the difference between, e.g., *pat* and *cat*, a learner has to choose between demoting the Markedness constraint *PoA below IO-IDENT-PoA or below IO-IDENT-PoA/Onset, the latter only militating against unfaithfulness to segments in a syllable onset. Ranking *PoA below the positional Faithfulness constraint results in a more restrictive grammar than ranking it below the general Faithfulness constraint. The r-measure of the two grammars, though, is the same.

(19) Ranking options with equal r-measures

a. H_0: *PoA ›› IO-IDENT-PoA, IO-IDENT-PoA/Onset

b. H_n: IO-IDENT-PoA/Onset ›› *PoA ›› IO-IDENT-PoA

c. $H_{n'}$: IO-IDENT-PoA ›› *PoA ›› IO-IDENT-PoA/Onset

We can thus formulate an addition to Faithfulness Delay.

(20) Least Impact Strategy (freely interpreting Prince & Tesar 2004):
If a learner has the choice between two W-marked F constraints she ranks the one with the least impact. (i.e., rank the more specific constraint first).

A legitimate question to ask is if this should be extended to Markedness constraints: If some Markedness constraint has to be ranked below a Faithfulness constraint, only using the Markedness constraint with the widest scope has an effect, using a specific Markedness constraint wouldn't show any change in the choice of output candidate. Consider the same scenario on PoA with a positional Markedness constraint instead. Detecting a PoA contrast in word-initial position doesn't affect the ranking of *PoA/coda. Only *PoA incurs L marks in this scenario. *PoA/coda becomes relevant only once a learner detects words, such as *blog* and *blob*. Thus, in a language that doesn't have this kind of words, the positional Markedness constraint is left on top of the hierarchy and should show an

effect in second language acquisition, which is borne out, see e.g., Broselow et al. (1998). See also the discussion of default rankings of Markedness constraints in section 3.

In the following I illustrate the connection between learning and historical change in OT with the emergence of the Coda Condition (Itô 1988) in Italian. The basic idea is that this is a case of incomplete acquisition. A generation of Italians just stopped reranking constraints too early.

Latin displayed a wide range of consonants in word-final and in word-internal preconsonantal position. This disjunction is usually unified in the assumption of the coda as a syllable constituent.

(21) Latin syllable-final consonants

a. Internal codas (mostly from Tekavčić 1980:149)

LŪCTUM	‘mourning’
CLAUSTRUM	‘bolt, bar, prison, cloister’
FARCTUS	‘stuffed’

b. Word-final codas

CAMPUM	‘field’
SATIS	‘enough’
CAPUT	‘head; top; leader’
SOL	‘sun’

NOMEN	‘name’
PATER	‘father’
AB, OB	prep.
ISTUD	
REX [ks]	‘king’

In comparison to Latin, the range of coda consonants is severely restricted in Italian. There is an increasing number of consonant-final loanwords, as shown in (22)a. Words directly inherited from Latin however, show several different types of modifications which all conspire to avoiding word-final consonants and non-coronal consonants in word-internal codas. Distinct PoAs are only allowed in word-internal codas, if the consonant is a geminate, i.e., also linked to the following onset, as illustrated in (23).

(22) Italian word-final consonants

a.	ˈkampo	'field'	b.	ˈbaɾ	‘bar’
	ˈre	‘king’		ˈbus	‘bus’
	ˈnome	‘name’		bɾiˈɔʃ	‘brioche’
	virˈtu	‘virtue		ˈklub	‘club’
	tʃivilˈta	‘civilisation		ˈspɾajt	‘Sprite’
				ˈfilm	‘film’

Words directly inherited from Latin however, show several different types of modifications which all conspire to avoiding word-final consonants and non-coronal consonants in word-internal codas. Distinct PoAs are only allowed in word-internal codas, if the consonant is a geminate, i.e., also linked to the following onset, as illustrated in (23)a. Other than that we find the coronal fricative /s/ (23)b, nasals that share PoA with the following onset (23)c, and the two alveolar liquids (23)d. There are a couple of exceptions, some of them in free variation with a more restricted form (23)e.

(23) Italian word-internal codas

a.	ˈpatto	‘pact’	c.	ˈampjo	‘ample’
	ˈfaʃʃa	‘band, bandage’		tɾiˈoɱfo	‘triumph’
	ˈdʒɛmma	‘gem’		ˈpoⁿⁿte	‘bridge’
	ˈkaɾɾo	‘cart’		ˈʃentsa	‘science’
	...			ˈlintʃe	‘lynx’
b.	ˈkaspita	‘good gracious!’		ˈkaŋkɾo	‘cancer’
	ˈpasta	‘pasta’	d.	ˈkɔɾpo	‘corpse, body’
	ˈkasko	‘helmet’		ˈkolpo	‘blow, stroke’
			e.	ˈkɔpto / ˈkɔtto	‘Coptic’
				atˈlante	‘atlas, book of maps’

ˈɛtna / ˈɛnna	‘Etna (name of a volcano)’
ˈkaktus	‘cactus’

The development from Latin to Italian codas, disregarding the recent loanwords, can be analysed as incomplete learning.

A learner of Proto-Romance started with all Markedness constraints outranking all Faithfulness constraints and was faced with the ranking dilemma discussed above, when discovering a PoA contrast in onsets.

(24) Incomplete acquisition

gato	*PoA	IO-Ident-PoA	IO-Ident-PoA/Onset
gato > tato	L	W	W

A careful learner chooses to demote the Markedness constraint only below the positional Faithfulness constraint, to keep the grammar as restrictive as possible.

Once the learner is confronted with more complex inputs she automatically produces the innovative Italian pattern.

(25) Accidental coda neutralisation

/akto/	IO-Ident-PoA/Onset	*PoA	IO-Ident-PoA
a. akto		*!	
☞ b. atto			*

Learning the Latin pattern would thus involve a second ranking operation, which is left out in Italian. This historical change is thus an effect of incomplete learning. To complete the acquisition of Latin a learner would have had to demote *PoA one further step, below the general Faithfulness constraint.

(26) The Latin target ranking

/akto/	IO-Ident-PoA/Onset	IO-Ident-PoA	*PoA
☞ a. akto			*!
b. atto		*	

The emergence of lateral palatalization constitutes a case of incomplete learning in which a Markedness constraint was left at the top that should have been demoted. Two Markedness constraints against laterals are relevant here. The general *Lateral militates against all laterals, and the very specific *Cl doesn't allow laterals in complex onsets. In Italian, Obstruent + lateral clusters are the onsets with the flattest sonority rise,

compared to obstruent + rhotic and obstruent + glide onsets. The latter two classes of sonorants are generally assumed to be of higher sonority than laterals. *Cl might thus be one of the constraints that add up to the Sonority Sequencing Principle (Clements 1990). We can formalize these as stringent constraints, as shown in (27).

(27) Sonority Sequencing constraints on onsets

a. *C+lateral (*Cl): Assign a violation mark for every obstruent+lateral onset.

b. *C+liquid (*CL): Assign a violation mark for every obstruent+liquid onset.

c. *C+sonorant (*CS): Assign a violation mark for every obstruent+sonorant onset.

At an early stage, at which she doesn't master complex onsets yet, due to high ranking *COMPLEXOnset, the learner is concerned with learning laterals. To this end she demotes *Lateral below Faithfulness, but doesn't touch *Cl. When she finally learns complex onsets she again demotes the more general constraints first, i.e., first *CS, then *CL. Medieval Italian learners didn't finish demotion operations and "forgot" *Cl on top of the hierarchy. As history shows, forgotten constraints can have a visible impact.

(28) Learning laterals – H_0

/lato/ 'side'	*Cl	*CL	*CS	*Lateral	FAITHFULNESS
a. lato				*!	
☞ b. jato					*

(29) Learning laterals – H_1

/lato/ 'side'	*Cl	*CL	*CS	FAITHFULNESS	*Lateral
☞ a. lato					*
b. jato				*!	

(30) Learning complex onsets – starting hierarchy

/prato/ 'meadow'	*Cl	*CL	*CS	FAITHFULNESS	*Lateral
a. prato		*!	*		
☞ b. pato				*	

From this starting ranking, a learner most probably first demotes the most general Markedness constraint, *CS, and then the next general one, *CL. I conflate both steps here into one tableau.

(31) Learning complex onsets after demotion of two M constraints

/prato/ 'meadow'	*Cl	FAITHFULNESS	*CL	*CS	*Lateral
☞ a. prato			*	*	
b. pato		*!			

After demotion of the two more general Markedness constraints, a learner should also have demoted the last Markedness constraint. Apparently, this step was left out by several generations of Italians. Learning remained incomplete, as shown in the next talbeau.

(32) Learning terminated before completion

/pleno/ 'full'	*Cl	FAITHFULNESS	*CL	*CS	*Lateral
a. pleno	*!		*	*	*
b. preno		*	*!	*	
☞ c. pjeno		*		*	

In these cases we see Emergence of the Unmarked Effects (TETU, see McCarthy & Prince 1994). However, recall the comparison of the Latin and the Italian segment inventories. The consonant and the diphthong inventories expanded considerably in the course of history, increasing the

number of marked segments and segment combinations. How does more marked structure emerge historically?

2.2 Opacity as a motor of lexical innovation

Marked segments emerge as the result of assimilation or other phonological processes, e.g., coalescence. However, as long as the phonological process is transparent and the emergent marked segments predictable one doesn't consider this as the introduction of new contrasts. The novel segments displayed in table (2) are contrastive, since we find minimal pairs, such as those in (33), or we find these segments in unpredictable environments, as also demonstrated by these forms.

(33) Unpredictable palatals

a.	*palla*	[pal:a]	'ball'	b.	*ricco*	[rik:o]	'rich'
	paglia	[paʎ:a]	'hay'		*riccio*	[rit͡ʃ:o]	'hedgehog'
c.	*cinema*	[t͡ʃinema]	'cinema'				
	chitarra	[kitara]	'guitar'				

Bermúdez-Otero (2007) explains the introduction of new contrasts as the opaque interaction of successively innovated phonological processes. He illustrates this with the law of palatals in Sanskrit. Proto-Indo-Iranian is

assumed to have had /k/ before all five vowels and no /c/. A palatalization rule then changed /k/ before non-low front vowels into [c]. A subsequently introduced vowel lowering rule removed the trigger of palatalization by lowering the mid vowels to [a]. Despite the absence of the context for the palatalization rule, the palatal consonants were retained. Thus, the allophones [k] and [c] of /k/ split into the two contrasting phonemes /k/ and /c/, at least before [a].

(34) The Law of Palatals (Bermúdez-Otero 2007: 506)

a) Proto-Indo-Iranian	*-ki-	*-ke-	*-ka-	*-ko-	*-ku-
b) Palatalization	***-ci-**	***-ce-**	*-ka-	*-ko-	*-ku-
c) Lowering /e,o/ → [a]	*-ci-	***-ca-**	*-ka-	***-ka-**	*-ku-
d) Sanskrit distribution	-ci-	-	{k, c}	a-	-ku-

This is an instance of historical phonological counter-bleeding opacity (see Kiparsky 1973 on the characterization of opacity, or the more recent discussion in McCarthy 2008). The context of a phonological rule is removed after its application.

Above we observed that weakening of laterals in complex onsets and mid vowel breaking created the context for velar palatalization, but the process didn't apply. This is an instance of historical counter-feeding opacity: the context for a phonological process is created after it has applied. One can describe the two respective cases as well in more neutral terms as overapplication and underapplication of a process, respectively. Palatalization in Sanskrit overapplies before some low vowels, while velar palatalization underapplies in Italian velar stops followed by palatal glides derived from laterals.

In the case of Sanskrit, an abstractly minded reader might insist that all instances of surface [c] can be derived from underlying /k/ if one assumes vowel lowering as a synchronically active process and the two allophones thus didn't split into two contrasting segments yet. In the same spirit one can say for Italian that one can still regard all instances of postalveolar affricates as derived from underlying velar stops, as long as one doesn't find a surface velar stop in an environment that is expected to cause palatalization. As soon as words such as *chiaro* 'clear' develop or words such as *chitarra* 'guitar' enter the language, the two segments have to be assumed to be contrastive at least in this position and all non-alternating surface postalveolar affricates that historically derive from velar stops have to be regarded as underlying postalveolars/palatals rather than underlying velars/dorsals.

Here we are dealing with two sources for a phonemic split, the introduction of a new phonological process that creates a situation of counter-feeding opacity, and the import of new words from other languages, which display the input configuration for the older phonological process at the surface.

2.3 Constraint cloning

Above in section 2.1 I discussed the r-measure as an evaluation metric for the restrictiveness of OT grammars. In principle, the hypothesis that all constraints are universal and languages only differ in terms of their ranking, yields only grammars of the same level of complexity. In rule-based phonology one can count phonological rules and the symbols used in the formulation of each rule to evaluate complexity. In OT, an analysis of a process can make use of cloned constraints, and we can count the number of such cloned constraints to assess how complex a grammar is. A high number of clones makes a grammar more difficult to learn, since they introduce idiosyncrasies, and thus such a grammar is expected to be diachronically unstable, i.e., prone to change.

There are three ways of cloning constraints. Pater (2009) used the term in connection with lexical indexing of constraints. Constraints can be

indexed and the copy of the constraint with the index is ranked higher in a hierarchy than its un-indexed original. Some lexical items are also indexed, and it is only the output candidates of these inputs that are sensitive to the higher ranked indexed constraint clone.

Arbitrary indexing of Markedness constraints explains rules that only apply in certain morphemes, while arbitrary indexing of Faithfulness constraints explains exceptional underapplication of a phonological process in certain morphemes (see Krämer 2009a,b for the use of both in the analysis of velar palatalization in Italian nouns).

A non-arbitrary type of indexing is positional Faithfulness: There is a general Faithfulness constraint (e.g., IO-IDENT-PoA above) and a more restricted version that is only active in a certain environment (e.g., IO-IDENT-PoA/onset above), that is a proper subset of the scope of the more general constraint. There are two differences between the two forms of indexing. First, positional restriction is not arbitrary, as lexical indexing can be. It refers to well-defined positions or classes, such as stressed syllables (i.e., prosodically defined) or stems (i.e., morphologically defined). The boundary between the two types of indexing already becomes blurry with morphologically defined classes/domains. Second, for prosodically defined positions one could say that positional Faithfulness is defined over surface categories, while indexing is defined over input properties.

The third cloning option is Local Constraint Conjunction (LCC; Prince & Smolensky 1993/2004, Łubowicz 2002, Smolensky 2006). Two (or more) constraints join forces in a domain (e.g., the segment) and every instance of that domain in which each of the two constraints is violated constitutes a violation of the local conjunction of the two constraints. The LCC only has an effect on output forms if it dominates at least one of the two constraints involved. The idea is thus that the LCC and its component constraints are present in the hierarchy independently. We can thus subsume this kind of constraint interaction under cloning. In the next section I will use constraint cloning, i.e., indexation and LCCs to analyse exceptionality and indirect mapping. We can thus measure the level of grammatical complexity caused by lexical exceptions and opaque rule interaction. The more constraint clones a grammar contains the more instable it becomes, since it is built on contradictory ranking information. An instable grammar is instable in the sense that a learner can neither infer the correct ranking nor the correct underlying forms and is therefore going to reanalyse both in a way distinct from the previous generation.

Before I move on to the actual analysis it should be instructive to set up the LCC part of the analysis schematically. Counter-feeding opacity can be considered as a chain shift (see Łubowicz 2004). In a chain shift there is a change from/A/ to [B] and from /B/ to [C]. Crucially the direct map from /A/ to [C] does not happen. For OT this is a problem, because

the map /B/ to [C] has to be caused by a Markedness constraint against B, i.e., *B. This Markedness constraint should also turn the mapping of /A/ to [B] less optimal than /A/ to [C], unless the two unfaithful maps are caused by different constraints and the map of /A/ to [C] is banned for an independent reason.

The idea to be advanced here is based on Kirchner's (1996) proposal of LCCs of Faithfulness constraints as an explanation for synchronic chain shifts. Extending this approach to counter-feeding opacity boils down to the following. If the triggering environment for a process is created by an unfaithful mapping, the application of this process would result in an even more unfaithful mapping. While /A/ to [B] involves one violation of Faithfulness constraint F1, i.e., for changing one feature, the mapping of /A/ to [C] must incur an additional violation of a second Faithfulness constraint, F2, i.e., for changing a second feature. While both constraints are ranked below some Markedness constraints which cause the mappings of /A/ to [B] and /B/ to [C], e.g., *A and *B, the LCC of both F1 and F2 outranks both *A and *B.

(35) Counter-feeding opacity as a chain shift as an LCC effect

	F1&F2	*A	*B	F1	F2
/A/ → [A]		*!			
☞ /A/ → [B]			*	*	
/A/ → [C]	*!			*	*
/B/ → [A]		*!			
/B/ → [B]			*!		
☞ /B/ → [C]					*

LCC adds constraints to the grammar. Likewise does Constraint indexation. With constraint indexation we can add an exception to the above schematic pattern.

(36) Counter-feeding opacity as a chain shift as an LCC effect

	F1&F2	$F1_L$	*A	*B	F1	F2
/A/ → [A]			*!			
☞ /A/ → [B]				*	*	
/A/ → [C]	*!				*	*
/B/ → [A]			*!			
/B/ → [B]				*!		
☞ /B/ → [C]						*
☞ $/A_L/$ → [A]						
$/A_L/$ → [B]		*!			*	
$/A_L/$ → [C]	*!	*			*	*

The forms tagged with a lexical index L are loanwords, which defy application of otherwise regular neutralisation patterns.

Acquiring language that contains both /A/ and $/A_L/$, and the indexed constraint clone, is difficult for a learner. First of all it is a non-trivial task to figure out whether this is a case of a lexical exception to a productive process or whether the form undergoing the process is the exception, and, correspondingly, it is the Markedness constraint, e.g., *A, that is co-indexed.

To be able to establish that [B] is a correspondent of /A/ and that [C] is a surface correspondent of /B/ these mappings have to be context-specific and at least some instances of type /A/ have to map to [A] in other environments, and some instances of type /B/ have to map to [B]. Otherwise, the mappings /A/ to [B] and /B/ to [C] become undetectable and the inputs irrecoverable. A learner would have no other choice but to resort to a faithful map and a different (less restrictive) grammar.
In the next section we will see how this works in detail.

3 To input irrecoverability ... and beyond

The chronological scenario for the changes from Latin to Italian was set up in section 1.2 and is recapitulated in simplified form here.

1. Palatalization – (causing alternations, food for Free Rides)
2. Mid vowel breaking – (causing alternations, food for Free Rides)
3. L-weakening – (Sweeping the lexicon, causing no alternations)
4. ClV reintroduction via (re-)borrowing

In this section, we will look into the mechanical details of grammatical and lexical change. I will give an account of how velar palatalization and then lateral gliding/palatalization have emerged in acquisition, how the latter

process created exceptions to the former and how both processes were further moved towards unproductivity by the introduction of loanwords which triggered subsequent grammatical reorganization.

3.1 The emergence of velar palatalization

In the analysis of palatalization I will follow Krämer (2009a) in assuming the affricate to be defined by the presence of two place features, [coronal] and [dorsal], the former contributed by the high vowel or glide and the latter by the velar stop. As indicated in the following tableau, already prior to learning, the grammar contains all sorts of Markedness constraints. The central one here is PAL, an abbreviation for a co-occurrence or co-articulation constraint that doesn't tolerate dorsal obstruents before coronal vocoids. This constraint can be satisfied by fusing the two segments, as here, or by spreading [coronal] from the vocoid to the preceding dorsal, and thereby changing the precedence relation between the features (and by other strategies which will not be discussed here, see Collins & Krämer 2015 and Kochetov 2015).

(37) Introducing palatalization: Step 1, no ranking

/rikio/ 'hedgehog'	PAL	*Dors	*Dors&*Cor	*CPLX	FAITH
a. ɾittʃo		*	*		*
b. ɾikjo	*	*		*	

A learner of an archaic form of Romance would be expected to arrive at the ranking shown in the next tableau.

(38) Expected ranking in Latin

/rikio/	*Dors&*Cor	PAL	FAITH	*Dors
a. ɾittʃo	*		*	*
☞ b. ɾikjo		*		*

However, as discussed in section 2.1, the learner has a bias towards keeping Markedness constraints high in the hierarchy, which include PAL, and some generations of learners must have missed the target ranking and instead ordered the constraints in a different way.[2]

[2] As a reviewer rightly points out, palatalization doesn't emerge directly from reranking. It is usually assumed to start as phonetic coarticulation. Dorsal/velar obstruents are generally slightly fronted before front vowels. This fronting has to be exaggerated by a generation or several and/or misinterpreted as phonological by some learners.

(39) Achieved ranking in Late Latin/Early Romance

	/rikio/	PAL	*Dors&*Cor	FAITH	*Dors
☞ a.	ɾittʃo		*	*	*
b.	ɾikjo	*			*

This ranking has two striking properties. First, the ranking of *[dorsal] below Faithfulness is warranted by any learner's confrontation with words containing velar stops. Second the ranking of all the Markedness constraints follows one simple principle: More specific ranks above more general. PAL is more specific than *Dors&*Cor since it refers to a sequence of segments, whereas the latter only refers to single segments, i.e., any segment that contains both features violates the constraint. *Dors finally, is less specific than *Dors&*Cor since it only refers to one feature. We are thus dealing with some kind of default ranking, an instance of the Least Impact Strategy: the constraints with the smallest scope are ranked highest.

Accordingly, a new and quite complex (double PoA specification) segment is introduced by simply not changing the default ranking of constraints. However, at this stage the segment isn't contrastive yet. It emerges predictably and a learner could still have all instances of surface [tʃ] take a free ride on alternating forms that unambiguously have

underlying /k/ (such as *vinciamo – vinco* '(we) win – (I) win'; the same holds for the voiced counterparts, e.g., *leggiamo – leggo* '(we) read – (I) read'). The emerging affricates aren't lexicalised yet even though the trigger of palatalization, the front glide, disappears in the process.

3.2 Constructing and breaking a chain

On the one hand we have palatal affricates that alternate predictably, as well as affricates that don't alternate and don't show a triggering environment, such as in *riccio* 'hedgehog'. On the other hand we encounter triggering environments in which the process underapplies once the laterals start to palatalize. (e.g., Latin CLAUSTRUM turns into *chiostro* [kjostro] 'cloister', or *okklo* becomes *occhio* [okkjo] 'eye'. An important question is why not all varieties just applied palatalization transparently in the newly derived environments.

Palatalizing the [kj] sequence in 'eye' to [tʃ] would be a two-step chain, changing the underlying segment twice:

(40) Interrupted chain of change: okklo → okkjo *→ otʃo

In parallel OT, such changes are expected, since each of the changes is an optimization in response to a surface Markedness constraint. However,

such double unfaithfulness seems to be suboptimal synchronically, especially in language acquisition, resulting in chain shifts (see Łubowicz 2011 on chain shifts) in child language, such as the famous *puzzle–puddle–puggle* shift observed in children acquiring English (Smith 1973). Children that try to say *puzzle* end up saying *puddle* instead. However, when targeting *puddle* they realize *puggle*. Since *puddle* is as unacceptable as *puzzle* one would expect that both *puzzle* and *puddle* are realized as *puggle*. To realize *puzzle* as *puggle*, however, a child would have to change both manner and place of articulation. Thus, relative markedness seems to be less important than faithfulness.

We observe the same in our case. To change historical /klV/ into [tʃV] at least one more change is necessary than for changing it into [kjV]. We can assume /kl/ --> |kj| -> [tʃ] as a trajectory of change on which |j| is representationally intermediate between the lateral and the affricate. However, depending on the way one looks at it we get a different result. From a purely historical perspective we just see that velar palatalization must have happened before lateral palatalization. However, there is no reason yet to assume that palatalization has stopped to be an active process in the language any time before lateral palatalization started. A derivational analysis of an early stage of Italian captures that by ordering both processes in their diachronic order in a synchronic grammar: The velar palatalization process happens before lateral palatalization within a

synchronic derivation at the historical stage at which lateral palatalization was still active. In the parallelist view of OT, however, we can drive home the result that velar palatalization is first blocked in what would be a phonologically derived environment in a derivational approach. A segment has to violate two Faithfulness constraints to undergo both velar palatalization and lateral palatalization.

(41) Historical, derivational, and parallel interaction

a. Historical: Palatalization happens before lateral gliding

b. Derivational: Palatalization happens before lateral gliding

c. Parallel: Two F violations is too much unfaithfulness

The individual constraints violated on the way from the lateral via the glide to the affricate have to be ranked below the involved Markedness constraints, for short, *ClV (maybe an instance of the Sonority Sequencing Principle) and the already introduced PAL constraint, since for one thing laterals are subject to change and for the other, glides disappear in affricates.

For a lateral to turn into a glide we could assume that the feature [lateral] has to change (and most likely the place feature as well), which violates a specialized Faithfulness constraint, IO-IDENT[lateral]. For the glide to coalesce with a dorsal stop into an affricate, it at least has to give

up its status as a sonorant. We can thus identify IO-IDENT[sonorant] as the involved constraint that is violated in the disappearance of the glide in velar palatalization. As inspection of the tableau below reveals, ranking any of the two Faithfulness constraints above any of the two involved Markedness constraints doesn't yield the right results. Furthermore, we encounter a ranking paradox: (a.ii) supplies information to rank IDENT(son) above PAL, while (bi) provides the opposite information.

(42) Lateral palatalization and opacity I: The conundrum

a. /okklo/	*ClV	PAL	IDENT(lat)	IDENT(son)
a.i. okkjo > okklo	W		L	
a.ii. okkjo > ottʃo		L		W
b. /rikkjo/				
b.i. ɾittʃo > ɾikkjo		W		L
b.ii. ɾittʃo > ɾikklo	W		L	

Hence the two Faithfulness constraints join forces in a Local Constraint Conjunction, forming a third constraint that has a more narrow violation profile than the individual constraints and can be ranked at a higher stratum. This LCC is violated only by segments, which violate each of the two conjoined constraints. Tableau (43) shows the ranking information a learner has available. PAL favours the loser in (a.ii) and should be demoted

below either IDENT(son) or the LCC, which both favour the winner. Demotion of PAL below Ident(son) is blocked by (b.i), since for this winner-loser pair, PAL favours the winner and IDENT(son) the loser. The LCC is neutral with respect to this pair.

(43) Lateral palatalization and opacity II: LCC

a. /okklo/	*ClV	PAL	IDENT (lat)	IDENT (son)	ID(lat)&ID(son)
a.i. okkjo > okklo	W		L		
a.ii. okkjo > ottʃo		L		W	W
b. /rikkjo/					
b.i. ɾittʃo > ɾikkjo		W		L	
b.ii. ɾittʃo > ɾikklo	W		L		

We thus arrive at the following grammar.

(44) Early Italian: Glide formation from post-consonantal laterals and opacity

	*ClV	ID(lat)&ID(son)	PAL	ID(lat)	ID(son)
a. /okklo/ - okklo	*!				
☞ b. /okklo/ - okkjo			*	*	
c. /okklo/ - ottʃo		*!		*	*
d. /rikkjo/ - ɾikkjo			*!		
☞ e. /rikkjo/ - ɾittʃo					*

As mentioned above, some dialects go the transparent road, as in words such as [tʃ]esa ‘church’. For such dialects we can assume that the Markedness constraint hasn’t been demoted below this LCC.[3]

For Italian, the question arises as to where that LCC comes from. We could attribute it to the Least Impact Strategy, introduced above, exemplified by the emergence of the Coda Condition.

[3] The situation is admittedly a bit more complex, since the same speaker might produce church with palatalization, but eye without. Thus, additional constraints, such as Contiguity have to be recruited in a full analysis of dialectal variation of these patterns.

(45) Least Impact Strategy:

If a learner has the choice between two W-marked F constraints she ranks the one with the least impact. (Specific › General)

However, a LCC is not a primitive constraint. It is a construct of two primitives and one normally would assume that it isn't part of the constraint set universally. On the other hand one could make the same assumption about positional Faithfulness constraints. As discussed above, a positional Faithfulness constraint is a clone of a general Faithfulness constraint, tagged for a certain domain or category or class. It is thus also not a primitive constraint and could be assumed to be a language-specific construct. However, as discussed with chain shifts above, positional effects emerge spontaneously in language acquisition. Smith (1973) also observes final devoicing in his English child language data, a process that isn't expected to occur in English learning children, since it isn't part of English phonology. The only plausible explanation of such phenomena is that the constraints that shape these patterns are there universally and that the respective rankings are an effect of the transition towards the target ranking or emerge because of some ranking biases (such as the Least Impact Strategy for the ranking of F constraints).[4]

[4] Whether "universal" means genetically predetermined or inductively learned by every human being doesn't matter for the current discussion. However, such constraints cannot be learned by exposure to data from the target language. Thus

3.3 Borrowing with an already unstable grammar

The next historical step is the introduction of new words that violate otherwise top-ranked phonotactic constraints. The earliest of these loans are borrowed from Latin, e.g., *flauto* 'flute' or *classe* 'class'. Presumably these words were introduced by educated adults. Adults deal with new words in a way different from children learning their first language. Adults have their constraint ranking in place. Any new form that is inconsistent with the ranking is either assimilated or stored as an exception. Pater (2009) proposes indexed Faithfulness constraint as a means to store exceptional morphemes. Thus, adults clone a constraint and add it to their hierarchy. The following tableau shows our crucial forms *occhio* 'eye', *riccio* 'hedgehog', *ciclo* 'cycle'.

(46) Glide formation from post-consonantal laterals and opacity

/okklo/	ID(lat)&ID(son)	$ID(lat)_L$	*ClV	PAL	ID(lat)	ID(son)
a. okklo			*!			
☞ b. okkjo				*	*	
c. ottʃo	*!				*	*

inductive grounding of constraints has to be a more abstract process than simple behaviourist exposure to surface data.

	/ɾikkjo/						
d.	ɾikkjo				*!		
☞ e.	ɾittʃo						*
	/kiklo$_L$/						
☞ f.	tʃiklo			*			
g.	tʃikjo		*!		*	*	
h.	tʃittʃo	*!	*			*	*

The attentive reader might have noticed that *ciclo* is given here as underlying /kiklo/ rather than /tʃiklo/. We thus see that loanword phonology is selective. While lateral palatalization is deactivated, velar palatalization shows an effect. However, since the same processes (lateral palatalization, diphthongisation and borrowing) have an obfuscating effect on initial position as well, palatalization does not endure in this position either (the affricates in words like *ciclo* or *cinema* might actually be an effect of orthographic rules rather than productive phonology).

3.4 Restructuring of the input leads to reorganization of the grammar

A learner of Italian is faced at this stage, i.e., after mid-vowel breaking, lateral palatalization and several instances of borrowing, with the following data.

(47) Anything goes

a. *cena – chiesa – clerico – cherubino – chilo*

[tʃe] – [kje] – [kle] – [ke] – [ki]

'dinner – church – cleric – cherub – kilo (35.27396195 oz)'

b. ciaspola – chiave – classe

[tʃa] – [kja] – [kla]

'snowshoe – key – class'

Velar palatalization still causes alternations, at least at the stem-suffix juncture, and so does mid-vowel breaking, under stress shift (e.g., *nuovo – novità* 'new – news'*; tiene – teniamo* 's/he keeps – we keep'). Lateral palatalization doesn't and never did. However, at this stage, even the velar-palatalization alternations cannot cause Free Rides of non-alternating affricates anymore, as for *ciclo*. For, if *ciclo* has an underlying /k/ that is turned into [tʃ] by velar palatalization, so should the underlying /k/ in *chilo*.

(48) Lexicon and grammar have to be restructured.

- Previous generation's /klave/ becomes /kjave /
- Previous generation's /kiklo/ becomes /tʃiklo/

In the absence of evidence for coherent generalizations and corresponding free ride options for non-alternating forms, the indexed constraint for loanwords isn't cloned anymore and even if Markedness constraints are demoted initially only below the Faithfulness LCC, they will have to be demoted further, below general Faithfulness, in a further step when the learner has a fuller picture of the surface phonotactic possibilities.

To cast more light on the situation we go through the learning procedure, step by step. Assume a learner first learns the form [okkjo]. The ranking information is given in (49). The PAL constraint favours the candidate with an affricate, while the constraint against palatals, *Dors&*Cor, favours the winner in this competition (b). Thus, the learner is tempted to rank these two constraints accordingly. Changing the glide in the winner to a lateral doesn't improve on any Markedness constraint, it just adds a violation of *ClV. No additional ranking is required for this form.

(49) Faithful parse 1: Rank *Dors&*Cor ›› PAL

	*ClV	*Dors& *Cor	PAL	IDENT (manner)	ID(lat)& ID(son)
a. okkjo › okklo	W			W	
b. okkjo › ottʃo		W	L	W	

However, once the learner encounters a form like [ɾittʃo], a new ranking argument comes up (50). This time it is the reverse ranking of the two Markedness constraints that were ranked in the previous operation. For designated winner [ɾittʃo] to beat designated loser [ɾikkjo], PAL has to dominate *Dors&*Cor. The learner could now use Ident(manner) and demote *Dors&*Cor . The learner has two options, either revise input forms to e.g., /okklo/ and /rikkjo/ respectively, or consider further constraints to rank. The LCC of IDENT(lateral) and IDENT(sonorant) doesn't help in this situation if inputs aren't revised. The most efficient Faithfulness constraint to solve the problem without changed inputs would be IDENT(manner). To avoid too many steps here we assume that the learner keeps the decision on hold and I directly jump to the third informative form to be considered, one with a *kl* cluster, and, as a bonus information source, another affricate, *ciclo*, see (51).

(50) Faithful Parse 2 Problem! PAL ›› *Dors&*Cor, *ClV ›› *Dors&*Cor

	*ClV	*Dors& *Cor	PAL	IDENT (manner)	ID(lat)& ID(son)
a. ɾittʃo › ɾikkjo		L	W	W	
b. ɾittʃo › ɾikklo	W	L		W	W

Such forms are evidence to demote *ClV below PAL, as well as *Dors&*Cor below PAL, as shown in the following tableau. Pair (a) requires Pal above *Dors&*Cor. (b) and (c) require demotion of *Cl below PAL and *Dors&*Cor, respectively.

(51) Faithful Parse 3: PAL ›› *Dors&*Cor ›› *Cl

	*ClV	PAL	*Dors& *Cor	ID (manner)	ID(lat)& ID(son)
a. tʃiklo › kiklo		W	L		
b. tʃiklo › tʃikjo	L	W		W	
c. tʃiklo › tʃitʃo	L		W	W	W

Once we put the ranking information from the three sets of winner loser pairs together the inconsistencies become obvious (52). (a) contradicts (b), (c) contradicts (d), and (e) conflicts with combinations of (a, b) and (b, c), i.e., if, according to (e) Pal outranks *Cl, then (a) and (c) can't both be correct.

(52) Inconsistent ranking information

a. okkjo › *ottʃo = *Dors&*Cor ›› PAL

b. ɾittʃo › *ɾikkjo = PAL ›› *Dors&*Cor

c. ɾittʃo › *ɾikklo = *ClV ›› *Dors&*Cor

d. tʃiklo › *tʃitʃo = *Dors&*Cor ›› *ClV

e. tʃiklo › *tʃikjo = PAL ›› *ClV

The LCC doesn't help resolve any of these ranking paradoxes. It doesn't even help in connection with revised inputs. [kj] in *occhio* can't be /kl/ because then the grammar needs to produce the mapping of /kl/ to [kj], which would also affect [kl] in *ciclo* , which has to be changed to something else or indexed with some constraint. Changing /tʃ/ to /kj/ and /kj/ to /kl/ and indexing in addition to selecting the local conjunction would do, as we have seen in the historical run-up to the situation. However, the learner needs to figure out the chain shift, which inputs have to be changed to what and select the right constraints for the LCC, as well as the right constraint and lexical items for indexing. The chances are high that a learner simply selects the only constraint that can be ranked to overrule all conflicting information on the ranking of Markedness constraints, the Faithfulness constraint IDENT(manner), and decides to settle on an identity mapping for all items involved. While this reduces the r-measure of the grammar it greatly reduces its complexity, dispensing with arbitrary indexation and the use of LCCs.

This solution, however, only works as long as the learner doesn't consider the alternations caused by velar palatalization. Once a learner decomposes morphologically complex forms, the situation looks different again.

3.5 On the productivity of velar palatalization and the life cycle of constraint rankings

In this subsection we consider the issue whether velar palatalization is productive at all in present-day Italian. A potential partial answer to this question could provide further evidence for Bermúdez-Otero's "life cycle". In the life cycle, which is couched within Stratal Optimality Theory, it is assumed that innovative phonological processes enter the grammar at a late stratum. They apply across the board (as, e.g., flapping in American English). The older and less productive a process becomes the more its domain is narrowed, which is reflected in its percolation up to earlier strata of the grammar (from phrase to word to stem level) until it tolerates lexical exceptions and eventually becomes completely unproductive (as velar softening, i.e., spirantisation of /k/ in pairs like *electric – electricity*). Thus, velar palatalization could be still fully productive in Italian and its prosodic conditioning is just very complex. It could be partially active,

with lexical exceptions, or it could be inactive with some exceptional cases which show a lexically marked alternation. Alternatively it could have retreated to an earlier stratum.

Present-day Italians divide into two groups according to Krämer (2009 a,b), speakers with palatalization in nouns/adjectives as the exception and speakers with blocking of palatalization as the exception. Giavazzi (2008, 2012) conducted a nonce-word test (as Krämer 2009a,b) and arrives at the conclusion that palatalization is productive dependent on prosodic structure. Words with antepenultimate stress show palatalization, while words with penultimate stress display blocking. Apparently, nonce-word tests seem to produce variable results, depending on all sorts of external factors.

Giavazzi's reason to assume foot dependence is that if the consonant in triggering position can be assumed to be outside the main stress foot or at its edge, it is more likely to undergo palatalization. However, the reverse is not the case: Just because a velar can be assumed to be in a foot doesn't mean it is immune to palatalization. Furthermore, as the example *analogo* - *analoghi* with stress on the antepenult, in (54) shows, being outside a foot doesn't mean either that palatalization has to apply — it is just more likely. In example (53), feet are indicated by brackets.

The derived forms show that whether a velar palatalizes in a derivational context, that is, in a potentially earlier stratum of grammar, depends on its form and behaviour in inflection, not on the foot structure in the derived form. Furthermore, these foot parsings are at odds with recent analyses of Italian stress (see the discussion in Krämer 2009a,c).

(53) Palatalization, derivation and feet

a.	cattolico	cat(toli)ci	cattoli(cesi)mo	'catholic (sg/pl) / catholicism'
			cattoli(cissi)mo	'very catholic'
b.	sporco	(sporchi)	spor(chissi)mo	'dirt (sg/pl) / very dirty'
	turco	(turchi)	turchiz(zare)	'Turkish (sg/pl) / make Turkish'
c.	greco	(greci)	greciz(zare)	'Greek (sg/pl) / make Greek'

If one looks at further stems and derivational affixes, a more reliable conclusion seems to be that (a) some stems block palatalization in inflection, but don't do so with certain derivational affixes, as illustrated in (54)a, while other stems block palatalization as well in derived forms, as shown in (54)b. The reverse, i.e., blocking in derivation and application under inflection, doesn't seem to occur.

(54) Velar palatalization with derivational affixes

a.	pedagogo	pedagoghi	pedagogia	‘pedagogue/s pedagogics’
	opaco	opachi	opacità	‘opaque (m.sg/m.pl) / opacity’
			opacissimo	‘very opaque’
	analogo	analoghi	analogismo	‘analogue (m.sg/m.pl) analogism’
	pedagogo	pedagoghi	pedagogismo	‘educationism’
	--	--	pedagogista	‘educationist’
			pedagogico	‘pedagogic’
b.	fuoco	fuochi	fuochista	‘fire (sg/pl) / pyrotechnist’
	antico	antichi	antichità	‘antique (sg.pl) / antiquity’
			antichissimo	‘very antique’

Morpheme-initial/-internal unpredictability coupled with alternations at root-affix junctures will persuade learners to swap the above IDENT constraints with a constraint like CONTIGUITY(F) (‘No changes inside strings’). In addition, Base-Output Faithfulness as well as arbitrary indexation are required.

Thus, palatalization is retreating to the stem level, as can be inferred from the implicational relation among blocking contexts for stems. There are no stems that block palatalization in derivation but palatalize in inflection. Most of the derivational affixes exemplified above are usually considered stem-forming, we are thus dealing with stem level phonology, while the inflectional affixes could be argued to be subject to word level phonology. The pattern also tolerates exceptions (stems that block palatalization only in inflection or both in inflection and derivation), as assumed in Bermúdez-Otero's theory of the life cycle. Though, it doesn't do this in the straightforward way one would expect in Stratal OT. It is completely switched off in inflection of 1st conjugation verbs, but still completely regular in 2nd and 3rd conjugation inflection. And it is variably productive in nominal inflection, while in derivation its application depends on the involved morphemes. The most straightforward life cycle situation had been if all inflection behaved like 1st conjugation verbs and some or all derivational affixes triggered the process. However, as just sketched, the situation isn't as straightforward.

4 Conclusions

In this paper I have provided an amphichronic analysis of central facts of diachronic variation in the history of Italian phonology. Rather than simply

tracking change, the analysis provided several synchronic stages and, crucially, the trajectory from one stage to the next via intergenerational transmission, i.e., taking heed of the different roles of adults and language learners, couched within a theory of acquisition.

In this endeavour we have seen that some changes are driven by errors in acquisition and governed by basic principles of acquisition, while others are rooted in adults' additions to the lexicon leading to pattern inconsistency and therefore learning problems that result in radical lexical and grammatical restructuring.

The Italian history of change analysed here confirms Kiparsky's (2014) claims that "[c]hange occurs when some aspect of the target language is never acquired. [...]Changes which in the end simplify the language can pass through quite messy intermediate stages." As one can see from the current state of velar palatalization, Italian hasn't quite cleaned up its mess yet. This shows that it can take quite a while for a process to become inactive.

Acknowledgements

This paper benefitted from the many constructive comments I received at the 40th Incontro di Grammatica Generativa in Trento, at the talk I gave in

Edinburgh a week later, and at the workshop 'Historical Phonology and Phonological Theory' at the SLE 2015 in Leiden. I would like to thank the IGG organizers for inviting me, two anonymous reviewers for their helpful and inspiring feedback, and especially Birgit Alber and Patrick Honeybone for discussing various aspects of the ideas presented here with me. All errors and inconsistencies are the author's responsibility.

Please send comments and questions to martin.kramer@uit.no.

References

Allen, W. Sidney 1970. *Vox Latina*. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo 2006. Phonological change in Optimality Theory. In *Encyclopedia of language and linguistics*. 2nd edn. vol. 9, Keith Brown (ed.), 497-505. Oxford: Elsevier.

— 2007. Diachronic phonology. In *The Cambridge Handbook of Phonology*. Paul de Lacy (ed.), 497-517. Cambridge: CUP.

— 2014. Amphichronic explanation and the life cycle of phonological processes. In *The Oxford Handbook of Historical Phonology*. Patrick Honeybone & Joseph C. Salmons (eds.). Oxford: OUP.

— (forthcoming). *Stratal Optimality Theory*. Oxford: OUP.

Bermúdez-Otero, Ricardo & Richard M. Hogg 2003. The actuation problem in Optimality Theory: Phonologization, rule inversion and rule loss. In *Optimality Theory and Language Change*. D. Eric Holt (ed.), 91-119. Dordrecht: Kluwer Academic Publishers.

Bertinetto, Pier Marco, and Loporcaro, Michele 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association* 35: 131-151.

Broselow, Ellen, Su-I Chen & Chilin Wang 1998. The Emergence of the Unmarked in Second Language Phonology. *Studies in Second Language Acquisition* 20:261-280.

Calabrese, Andrea 1993. Palatalization processes in the history of Romance languages: A theoretical study. In *Linguistic perspectives on the Romance languages*. William J. Ashby, Marianne Mithun, Giorgio Perissinotto & Eduardo Raposo (eds.), 65-83. Amsterdam, Philadelphia: John Benjamins.

Clements, George N. 1990. The Role of the Sonority Cycle in Core Syllabification. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. John Kingston & Mary Beckman (eds.), 283-333. Cambridge: CUP.

Collins, Joe & Martin Krämer 2015. Crazy rules and grounded constraints. In *Proceedings of CLS 51*.

Fikkert, Paula 2007. Acquiring Phonology. In *The Cambridge Handbook of Phonology.* Paul de Lacy (ed.), 537-554. Cambridge: CUP.

Giavazzi, Maria 2008. On the Application of Velar Palatalization in Italian. *Proceedings of WCCFL 27*. UCLA Working Papers in Linguistics. 20-29.

— 2012. *Stress-conditioned palatalization in Italian*. Ms. MIT.

Hayes, Bruce & Donca, Steriade 2004. Introduction: the phonetic bases of phonological markedness. In *Phonetically based Phonology*. Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), 1-33. Cambridge: CUP.

Holt, D. Eric (ed.) 2003. *Optimality Theory and Language Change*. Kluwer.

Ito, Junko 1988. *Syllable Theory in Prosodic Phonology*. New York: Garland.

Ito, Junko & Armin Mester 1999. The structure of the phonological lexicon. In *The Handbook of Japanese Linguistics*. Natsuko Tsujimura (ed.), 62-100. Oxford: Blackwell.

— 2001. Covert generalizations in Optimality Theory: the role of stratal faithfulness constraints. In *Proceedings of the 2001 International Conference on Phonology and Morphology.* The Phonology-Morphology Circle of Korea, (ed.), 3-33.

Kiparsky, Paul 1973. Phonological representations. In *Three Dimensions of Linguistic Theory.* Osamu Fujimura (ed.), 3-136. Tokyo: TEC.

— 1988. Phonological change. In *Linguistics: The Cambridge Survey. Vol. 1: Linguistic Theory: Foundations*. Frederick J. Newmeyer (ed.), 363-415. Cambridge: CUP.

— 1995. The phonological basis of sound change. In *The Handbook of Phonological Theory*. John A. Goldsmith (ed.), 640-670. Oxford: Blackwell.

— 2006. The Amphichronic Program vs. Evolutionary Phonology. *Theoretical Linguistics* 32: 217-36.

— 2014. New perspectives in historical linguistics. In *Handbook of Historical Linguistics*. Claire Bowern (ed.). London: Routledge.

Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27: 341-350.

Kochetov, Alexei, 2015. *Palatalization and glide strengthening as competing repair strategies: Evidence from Kirundi.* Ms. University of Toronto.

Krämer, Martin 2009a. *The Phonology of Italian*. Oxford OUP.

— 2009b. Crypto-variation in Italian velar palatalisation. In *Romance Linguistics: Structures, Interfaces, and Microparametric Variation*. Pascual J. Masullo, Erin O'Rourke & Chia-Hui Huang (eds.), 193-208. Amsterdam: John Benjamins.

— 2009c. Main stress in Italian nonce nouns. In. *Romance Languages and Linguistic Theory 2006*. Danièle Torck & W. Leo Wetzels (eds.), 127-141. Amsterdam: John Benjamins.

— 2012. *Underlying Representations*. Cambridge: CUP.

— in press. Metaphonic chain shifts, vowel height, and markedness. In *Approaches to metaphony in the Languages of Italy*. Katrin Linke, Marc van Oostendorp, Francesc Josep Torres-Tamarit (eds.). Berlin: Mouton de Gruyter.

Łubowicz, Anna 2002. Derived environment effects in Optimality Theory. *Lingua* 112: 243-280.

— 2004. Counter-feeding opacity as a chain shift effect. *Proceedings of the WCCFL* 22. 315-327.

— 2011. Chain Shifts. In *Blackwell Companion to Phonology*. Marc van Oostendorp, Colin Ewen, Beth Hume and Keren Rice (eds.). Hoboken: Wiley-Blackwell.

McCarthy, John 2005. Taking a free ride in morphophonemic learning. In, 'Morphology in Phonology', Maria-Rosa Lloret & Jesus Jiménez (eds.), thematic issue of *Catalan Journal of Linguistics* 4: 19-55.

— 2007. *Hidden Generalizations. Phonological Opacity in Optimality Theory*. London: Equinox.

McMahon, April 2000. *Lexical Phonology and the History of English*. Cambridge: CUP.

Pater, Joe 2000. Nonuniformity in English stress: the role of ranked and lexically specific constraints. *Phonology* 17: 237-74.

— 2006. The locus of exceptionality: morpheme-specific phonology as constraint indexation. *UMOP 32 – Papers in Optimality Theory III*. Amherst: GLSA.

— 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In *Phonological argumentation. Essays on evidence and motivation*. Steve Parker (ed.), 123-154. London: Equinox.

Prince, Alan & Bruce Tesar 2004. Learning phonotactic distributions. In *Constraints in Phonological acquisition*. René Kager, Joe Pater & Wim Zonneveld (eds.), 245-291. Cambridge: CUP.

Smith, Neil V. 1973. *The acquisition of phonology: a case study*. Cambridge: CUP.

Smolensky, Paul 2006. Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In *The harmonic mind: From neural computation to optimality theoretic grammar, vol. 2: Linguistic and philosophical implications.* Paul Smolensky & Géraldine Legendre (eds.), 27-160. Cambridge, MA: MIT Press.

Tekavčić, Pavao 1980. *Grammatica storica dell'italiano. I. Fonematica*. Bologna: Il Mulino.

Tesar, Bruce 1997. *Multi-Recursive Constraint Demotion*. Ms. Rutgers University.

Tesar, Bruce & Paul Smolensky 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Vincent, Nigel 1988a. Latin. In *The Romance Languages*. Martin Harris & Nigel Vincent (eds.), 26-78. New York: Routledge.

— 1988b. Italian. In *The Romance Languages*. Martin Harris & Nigel Vincent (eds.), 279-313. New York: Routledge.