

*The Inadequacy of Filters and Faithfulness in Loanword Adaptation**

Carole Paradis
Laval University

1 Introduction

The objective of this paper is twofold. First, it aims to show that segment deletion in borrowings is largely predictable, and that this predictability might be problematic for a filter-based framework since, as we will see, it entails that phonological processes are visible to phonological constraints. For instance, it is at odds with a filter-based framework such as Optimality Theory (OT; Prince & Smolensky 1993 and McCarthy & Prince 1993), where constraints are in fact “filters”, i.e. surface constraints, which do not have access to the processes of the phonological component or the intermediate forms they generate since, by definition, filters deal with final outputs only.

Second, this paper will demonstrate that ill-formed segments contained in borrowings are adapted, i.e. recast into a different shape (85.2% of cases), or left unadapted (10.7% of cases) — if these are imports — instead of being deleted. Phonologically-induced segment deletion represents only 2.3% of cases, a fact which is attributed to a principle of the Theory of Constraints and Repair Strategies (TCRS; Paradis 1988a,b), the Preservation Principle. It will be shown that, although the Preservation Principle and Faithfulness (Parse and Fill) in OT seem comparable, they make indeed different predictions. In the view of Faithfulness, the optimal output (candidate) is the one which has undergone the least (segment) deletion, i.e. the one whose segments are all “parsed” (realised), in addition to being the one which has undergone the least (segment) epenthesis (Fill). In other words, the Faithfulness constraints (Parse and Fill) establish that the best candidate is the one which is as close as possible to its input. OT treats both Faithfulness constraints on a par in the sense that their ranking with respect to one another — as is the case with any other constraint in this framework — is determined on language-specific grounds, not universal ones. This means

that, statistically, segment deletion in loanwords across languages should be observed as often as segment insertion, a prediction which is not empirically supported.

The remainder of this article will be divided essentially into four parts. In 2, I will briefly show how strong the Preservation Principle is in our four corpora of loanwords (2.1), provide an overview of the functioning of TCRS (2.2), and present my assumptions concerning borrowings in general (2.3). In 3, we will see in which specific cases segment deletion occurs in Fula (3.1) and Kinyarwanda (3.2). In 4, I will address why segment deletion in borrowings is problematic for OT. Three conceivable solutions will be examined: 1) free constraint reranking (4.1); 2) mechanical counting of changes between inputs and outputs (4.2); 3) counting of changes within “constraint domains” (4.3). It will be shown that these three solutions all face theoretical and empirical problems. Finally, the Preservation Principle and Faithfulness will be compared in 5. We will see that the latter, as it stands, cannot capture the high rate of segment preservation in loanword adaptation.

2 Preliminary Information

2.1 Statistics regarding borrowings

Statistics regarding borrowings are presented in (1). They report results from the analysis of four corpora of loanwords studied by Paradis & LaCharité (1995a,b), which group together 12635 malformations (from the viewpoint of the borrowing language (L1)) contained in altogether 4031 borrowings in four different languages: English in Quebec French and French in Moroccan Arabic, Kinyarwanda and Fula.

(1) Number of	English loans in	French loans in:			Total
	Quebec French	Moroccan Arabic	Kinyarwanda	Fula	
borrowings	1183	1547	756	545	4031
forms ¹	2030	3512	2143	1036	8721
malformations	3673	3660	4444	858	12635
adaptations	2688 (73.2%)	2981 (81.4%)	4313 (97.1%)	785 (91.5%)	10767/12635 (85.2%)
deletions	130/3673 (3.5%)	230/3660 (6.3%) ²	104/4444 (2.3%)	57/858 (6.6%)	521/12635 (4.1%)
phonological deletions	71/3673 (1.9%)			32/858 (3.7%)	103/4531 (2.3%) ³
non- adaptations	855 (23.3%)	449 (12.3%)	27 (0.6%)	16 (1.9%)	1347/12635 (10.7%)

The general statistics at the end of the table indicate that segment deletion is rare (4.1% of cases). Segment deletion due to the presence of an ill-formed segment, i.e. due to phonological reasons, is even rarer with 2.3% of cases (see Paradis & LaCharité 1995a,b for details). I attribute this fact to a principle of TCRS, the Preservation Principle.

2.2 TCRS

The Preservation Principle, presented in (2), states that segmental information is not arbitrarily destroyed.

(2) PRESERVATION PRINCIPLE: Segmental information is maximally preserved, within the limits of the Threshold Principle.

The rescue of a segment (i.e. at least its root node) fails only when its retention would be too costly. The Threshold Principle in (3) posits that all languages impose a limit to segment preservation — the crucial part in (3a) — which, as indicated in (3b), does not exceed two processes within a given constraint domain (to be defined later).⁴

(3) THRESHOLD PRINCIPLE:

- a) All languages have a tolerance threshold to segment preservation.
- b) This threshold is set at two steps (or two repairs) within a given constraint domain.

When this principle is not respected, we will see that segment deletion applies. However, segment adaptation in borrowings, i.e. recasting of a segment into another shape, occurs in the vast majority of cases, i.e. 85.2% of cases as shown in (1). Adaptation is due to the application of a repair strategy, defined in (4).

- (4) REPAIR STRATEGY: A universal, non-contextual phonological operation that is triggered by the violation of a phonological constraint, and which inserts or deletes content or structure to ensure conformity to the constraint.

Repair strategies apply according to the Minimality Principle in (5).

- (5) MINIMALITY PRINCIPLE: Repairs
 a) apply at the lowest phonological level to which the violated constraint refers and
 b) involve as few strategies (steps) as possible.

The lowest phonological level is determined by the Phonological Level Hierarchy in (6). This scale simply reflects the phonological organisation required independently of TCRS, where the metrical level is the most important or say “organisational” level, and the terminal feature level the least one.

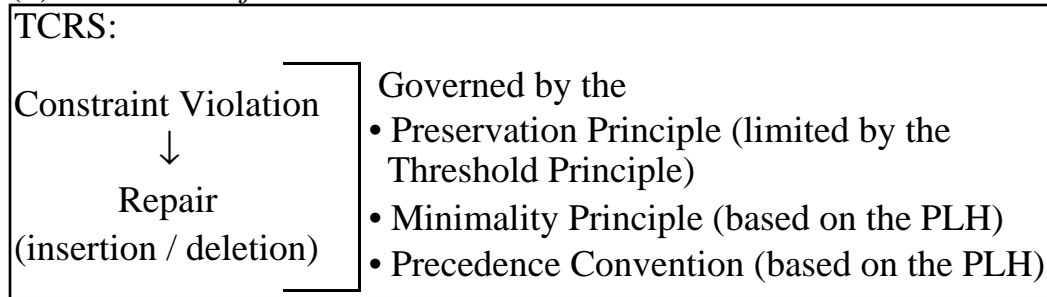
- (6) PHONOLOGICAL LEVEL HIERARCHY (PLH): Metrical level > syllabic level > skeletal level > root node > non-terminal feature > terminal feature.

The PLH also plays a crucial role in determining what the highest phonological level is in the Precedence Convention in (7), whose purpose is to establish which constraint has priority in a constraint conflict.

- (7) PRECEDENCE CONVENTION: In a situation involving two or more violated constraints, priority is given to that constraint referring to the highest phonological level of the PLH.

An overview of the organisation of TCRS is provided in (8).

(8) Overview of TCRS



2.3 Assumptions regarding borrowings.

Following Paradis & LaCharité (1995a,b), it will be assumed that borrowings are introduced by bilinguals, who have access to the phonology of the source language (L2), a claim supported by sociolinguistic studies such as Haugen (1950), Mougeon, Beniak & Valois (1985) and Poplack, Sankoff & Miller (1988), among many others. Loanwords are introduced by bilinguals through what sociolinguists call “code-switches”, “nonces” and “idiosyncrasies”. Phonological patterns of adaptation are also imposed by bilinguals, and are community-wide, especially in mid and high-community bilingualism stages. A definition of “borrowing” is provided in (9).

(9) *Definition of “borrowing”*: An individual word, or compound functioning as a single word, from L2 that phonologically conforms to (at least) the outermost peripheral phonological constraints of the target language (see Paradis & Lebel 1994), and that is incorporated into the discourse of L1.

It is also assumed that the borrowed form is not the phonetic output of L2 but its lexical or syntactic output (see Paradis & LaCharité 1995a,b for arguments). Finally, it will be taken for granted that borrowers remove from borrowings the information they perceive as redundant from the point of view of L1, unless this information is distinctive in L2. The distinctive feature combinations of L2 are all introduced into the L1 dictionary, the input to the lexicon.

3 Affordable and unaffordable repairs

This section aims to provide an overview of how the principles of TCRS work when applied to borrowings, and what kind of predictions they can make (see also Paradis & LaCharité 1995a,b). I will first present examples

c) <i>French</i>				<i>Fula</i>	<i>gloss</i>
bandit	[b ã d̥]	→	b an ⁿ di	gangster ⁶	
canton	[k ã t O]	→	k an t ON	canton	
marin	[mar E]	→	mar EN	sailor	

The examples in (11c) represent cases of isolated ill-formed segments. However, borrowings also very often contain sequences of ill-formed segments as in (12).

(12) $*v \rightarrow w$, $*\mathcal{J} \rightarrow s$, $*v\mathcal{S} \rightarrow VN$, $*S \rightarrow s$:

<i>French</i>			<i>Fula</i>	<i>gloss</i>
changer	[S ã J e]	→	sans -u-dE	to change
gendarme	[J ã darm]	→	san ⁿ darma	policeman
ventilateur	[vã S ilat,, r]	→	wan tilatOr	fan

For example, French *changer* contains three ill-formed segments in a row, i.e. \mathcal{S} , \mathcal{J} , and \tilde{a} which are prohibited respectively by the constraints in (13a,b) and the one we have already seen in (11a). Nonetheless, none of the French ill-formed segments of the sequences in (12) is deleted, and French *changer* yields Fula [sansud̥E]

- (13) a) *Parameter*: Voiced fricatives? French: yes
 (*[+continuant] [+voice] C) Fula: NO (*v, *z, *J)
- b) *Parameter*: Non-anterior fricatives? French: yes
 (*[+continuant] [-anterior]) Fula: NO (*J, *S)

Multiple violations can also be found within a single segment as is the case with the French voiced alveolar \mathcal{J} in the two first examples in (12). Additional examples are provided in (14), where we can see that \mathcal{J} is adapted in s in all cases.

(14) * \mathcal{J} → \mathcal{S}

<i>French</i>			<i>Fula</i>	<i>gloss</i>
barrage	[bara \mathcal{J}]	→	bara \mathcal{S}	dam
collège	[kO \mathcal{L} ɛ \mathcal{J}]	→	kO \mathcal{L} ɛ \mathcal{S}	school
journal	[\mathcal{J} urnal]	→	\mathcal{S} uurnal	newspaper

The examples in (15) clearly show that constraint (13a), against voiced fricatives, and constraint (13b), against non-anterior fricatives, are two separate constraints. Constraint (13b) is enforced in the Core AND in the Periphery (see Paradis & LaCharité 1995a,b on these notions in loanword adaptation), while constraint (13a) is weakened in the Periphery of Fula speakers, which explains why \mathcal{z} is sometimes tolerated.

(15) *Periphery*

<i>French</i>			<i>Fula</i>	<i>gloss</i>
agence	[a \mathcal{J} A \mathcal{S} s]	→	a \mathcal{Z} A \mathcal{S} s	agency
bougie	[bu \mathcal{J} i]	→	buu \mathcal{Z} i	candle
janvier	[\mathcal{J} A \mathcal{S} vje]	→	\mathcal{Z} anviye	January

Syllabic malformations also yield preservation of the existing segments through insertion of an additional segment, not deletion of an existing one. As can be seen in (16), Fula does not allow branching onsets or branching codas, which is expressed by the negative parameter setting in (16b).

(16) *Parameters on syllabic constituents:*

- | | |
|----------------------------------------|---------------------------|
| a) i. branching nuclei? | French: yes |
| | Fula: yes |
| ii. all types of branching nucleus? | Fula: NO |
| -e.g. rising diphthongs? | Fula: NO (*wi, *yi, etc.) |
| b) branching non-nuclear constituents? | French: yes |
| | Fula: NO |

In the case of an initial or final CC cluster, deletion of one of the consonants would satisfy the constraint (16b) but it would violate the Preservation Principle in (2), which requires segmental information to be preserved. I maintain that this is why a nucleus is inserted between the two consonants of such sequences, as in (17), instead of one of the consonants being deleted. The nucleus is subsequently filled by spreading of the closest surrounding vowel-like segment, i.e. a vowel ((17a)) or a glide ((17b)).

(17) #CC → #CVC:

	<i>French</i>		<i>Fula</i>	<i>gloss</i>	
a)	place	[pl [□] as]	→	pal [□] as	place
	tracteur	[tr [□] akt,, r]	→	tar [□] aktOr	tractor
	crayon	[kr [□] Ejϕ\$]	→	kEr [□] EjON	pencil
b)	boisson	[bw [□] asõ]	→	buw [□] asON	drink
	coiffer	[kw [□] afe]	→	kuw [□] a:f-a:-dE	coif (one's hair)
	lieutenant	[lj [□] fltnã]	→	lij [□] etinaN	lieutenant

Word-final CC clusters also yield nucleus insertion, not consonant deletion, which occurs after the CC cluster if the sonority of the cluster is rising ((18a)), or in between the two consonants if the sonority is falling ((18b)).

(18)	<i>French</i>		<i>Fula</i>	<i>gloss</i>	
a)	<i>rising sonority: CC# → CVC#</i>				
	filtre	[fil [□] tr]	→	fil [□] tir	filter
	mètre	[mE [□] tr]	→	mEE [□] tEr	meter
	table	[ta [□] bl]	→	taa [□] bal	table
	lettre	[lE [□] tr]	→	lEE [□] tEr	letter
b)	<i>falling sonority: CC# → CCV#</i>				
	carte	[ka [□] rd]	→	ka [□] rda	card (comb)
	carte	[ka [□] rt]	→	ka [□] rt-al	card
	force	[fO [□] rs]	→	fO [□] rsO	force

3.1.2 Constraint domains

This raises the question of what is exactly scanned by constraint (16b) or how deep and large is its domain. The examples in (18) suggest that segmental information ([±consonantal], relative sonority, etc.), however expressed within the segment, is visible to syllabic rules and their governing constraints (here (16b)). If this were not the case, it would be impossible to explain why the point of insertion of the epenthetic nucleus in (17) and (18a) is systematically located between the two consonants of the problematic CC cluster, while it is located after the cluster in (18b), and might be in

principle located before the sequence in (17); e.g. French *place* [plas] could yield **aplas*, a non-existing form in Fula — incidentally, this latter alternative is found in the Fula borrowing [Estati] from French *statue* [staty] the only example of our corpus where the sonority of a word-initial cluster is not rising. This leads me to express the following hypothesis.

- (19) VERTICAL SCANNING HYPOTHESIS: Scanning for syllabic and segmental malformations proceeds in parallel, i.e. when a cluster is examined for syllabic purposes, segmental constraints are also activated, and the cluster is then scanned for segmental purposes as well.

Therefore, if there is a segmental malformation within the cluster, it will be detected at the same time the sequence is scanned for syllabic purposes. The examples in (17) and (18) also reveal that it is not just the segmental content of a consonant candidate to syllabification which is scanned but its environment as well, i.e. the adjacent non-syllabified material. This is expressed by the following hypothesis.

- (20) HORIZONTAL SCANNING HYPOTHESIS: When a consonant within a cluster is unsyllabifiable, it is the whole consonant cluster which is identified as problematic for the syllabic constraint (here (16b)).

In other words, the non-nuclear syllabification rule makes sure before applying that the adjacent unsyllabified segments (i.e. those between two nuclei or a nucleus and a morpheme boundary) will be properly syllabified. This is why, jointly with constraint (16b), it scans the whole string of unsyllabified segments, in the same way the nucleus syllabification rule scans an entire word in search of sonority peaks. Unless one is ready to accept a costly resyllabification stage, it must be admitted that, if a problem is detected within a CC cluster, syllabification of the whole string is blocked until a repair decision is made, despite the fact, for instance, that the very first consonant of a word-final CC cluster could be syllabified within a coda. This suggests that the whole sequence, comprising the candidate to syllabification (C1) and the non-syllabifiable consonant (C2) is visible to constraint (16b), and identified as a domain. Hypotheses (19) and (20) lead me to formalise a “constraint domain” as in (21).

- (21) **CONSTRAINT DOMAIN:** A “constraint domain” is the “scope” of a constraint, i.e. the material which is horizontally and vertically scanned by a constraint before structure-building rules are authorised to apply.

Both hypotheses, (19) and (20), which contribute to define the notion of “constraint domain”, play a crucial role for the Threshold Principle in (3) since the cost of a repair is established in terms of the number of steps or processes required within the domain of the violated constraint to obtain constraint satisfaction.

We can now understand why the two consonants of the ill-formed CC clusters in (17) and (18) are both preserved, i.e. why the cluster undergoes vowel insertion, instead of consonant deletion. Vowel insertion, motivated by the Preservation Principle in (2), does not violate the Threshold Principle, since it requires only two steps, as shown in (22), i.e. nucleus insertion, as in (22a), and segmental filling of the nucleus, as in (22b).

(22) *Unsyllabifiable CC clusters in (17): 2 repairs within the CC domain.*⁷

#CCV → a) nucleus insertion b) vowel spreading or glide spreading

N	N	N	N	N	N
X	X	X	X	X	X
C		C	V	C	∅
				o	
					α

N	N	N	N
X	X	X	X
C	∅	C	V
			u
			α

3.1.3 Unaffordable repairs in Fula

Segment preservation would not be as economical in the examples in (23), which contain a segmental malformation embedded in a syllabic one. For instance, French *voyou* contains a problematic initial CC cluster, **vw*, which violates constraint (16b), in addition to a voiced fricative, **v*, which is prohibited by constraint (13a). As can be seen, the Fula output *waju* displays a segmental loss, i.e. that of **v*. The same is true of the other clusters within squares, which all underwent a segmental loss, i.e. **v* in (23b), *w* in (23c) and **æ* in (23d). It is noteworthy that the segment deleted is the one which is problematic for the two constraints, unless it constitutes a better onset or coda in terms of sonority than the other consonant of the cluster as in the case of *chewing gum*, where *w*, a permitted segment in Fula, is lost, not **S*.

(23) *Loss of a consonant: Violations of the Threshold Principle in (3)*

	<i>French</i>		<i>Fula</i>		<i>gloss</i>	
a)	voyou	[vwaju]	→	waju	*wuwaju	bum
	voyage	[vwajaʒ]	→	waja:s	*wuwaja:s	trip
	voiture	[vwatyr]	→	wati:ri	*wuwati:ri	car
b)	cuire	[kœivr]	→	kir	*kuwiri	copper
	pieuvre	[pi„vr]	→	piu:ri	*piu:wiri	octopus ⁸
c)	chewing gum	[Swi gOm]	→	siNgOm	*siwiNgOm	chewing gum
d)	cuire	[kœivr]	→	kiri	*kuwiri	copper
	biscuit	[bis kœi]	→	bisk	*biskuwi	biscuit
	tuyau	[tœi j o]	→	tijo	*tiwijo	pipe ⁹

Compare now the examples in (23) with those in (24).

(24)	(de l') huile	[d«lœi l]	→	dilw il	*dilil	oil
	minuit	[mi nœ i]	→	minw i	*mini	midnight

There is no segmental loss here because none of the clusters in (24) contains an unsyllabifiable consonant, although they do contain an ill-formed segment, i.e. *æ. The ill-formed consonant is simply adapted in *w* (one repair) after the first consonant is syllabified within the coda of the first syllable and the second consonant within the onset of the next syllable (default processes). By contrast, segment deletion applies in (23) because segment preservation would require too many steps for the Threshold Principle ((3)) within the domain of constraint (16b), i.e. CC. As shown in (25) and (26), it would require 1) nucleus insertion; 2) segmental filling of the inserted nucleus, and 3) adaptation of the ill-formed segment, that is more than two repairs, the limit imposed by the Threshold Principle. Note that the process order is determined by the Precedence Convention in (7), which requires that syllabic constraints be served before segmental ones.

(25) *vw → wuw

a) nucleus insertion b) w spreading c) *v adaptation ([+voice] delinking)

N		N		N
X X X		X X X		X X X
v G		v ɔ G		w n G
		u		j
		α		α

(26) *Cæ → Cuw

a) nucleus insertion b) æ spreading c) ̄æ adaptation [+anterior] delinking)

N		N		N
X X X		X X X		X X X
C G		C ɔ G		C n G
		u		j
		α		β

The same analysis applies to the borrowings in (27), where the nasality of the French nasal vowel is lost in Fula. These examples contrast with those in (11c), where a nasal vowel systematically yields a VN sequence.

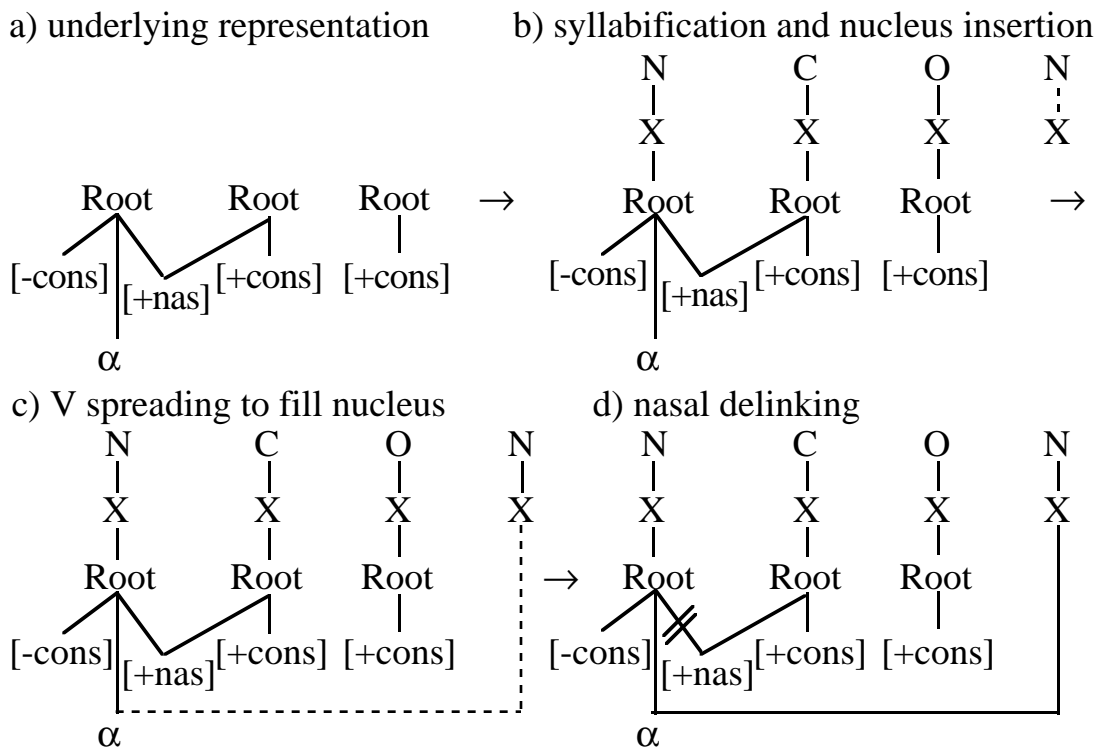
(27) *Loss of the nasality: Violations of the Threshold Principle*

<i>French</i>		<i>Fula</i>		<i>gloss</i>
propagande	[prɔpɑ̃A\$ɸ]	→ pɔrpaɔgn	* <td style="padding: 5px;">propaganda</td>	propaganda
éponge	[epɔ̃sɸ]	→ epɔs	*epɔnsɔ	sponge
balance	[balA\$ɸ]	→ balas	*balansa	scale
dimanche	[dimA\$ɸ]	→ dima:s	*dimansa	Sunday
essence	[EsA\$ɸ]	→ Esa:s	*Esansa	gasoline
vacances	[vakA\$ɸ]	→ wakkas ¹⁰	*wakkansa	vacation
dépense	[depA\$ɸ]	→ deppas	*deppansa	expense
conférence	[kɔ̃fɛrA\$ɸ]	→ kɔ̃feras	*kɔ̃feransa	conference

The constraint involved in (27) is the same as that involved in the examples in (23), i.e. the syllabic constraint against branching non-nuclear constituents in (16b). Since all of the examples in (27) already end with a consonant, nasal vowel unpacking (v\$ → VN) would yield a VNC sequence, i.e. a problematic word-final CC cluster which would violate constraint (16b). The French v\$C# inputs in (27) can be dealt with in two possible ways: 1) by

unpacking the nasal vowel as usual and inserting a nucleus after the VNC sequence (in order to syllabically break the CC cluster) or 2) by deleting one of the consonants of the CC cluster. As can be observed in (27), the option selected is the second one, i.e. deletion of one of the consonants, which infringes upon the Preservation Principle but satisfies the Threshold Principle in (3). Segment preservation would require more than two repairs, the limit established by the Threshold Principle. As shown in (28), it would necessitate: 1) nucleus insertion as in (28b), 2) segmental filling of the inserted nucleus as in (28c), and 3) delinking of the feature [+nasal] from the vowel, as in (28d) (the process order is determined here too by the Precedence Convention in (7)). This is why segment deletion applies instead. Since nasals constitute better codas than stops in terms of sonority, it is *d* which drops in the case of *propagande*. In the other cases, which all involve a fricative, the consonant deleted is the nasal, a consonant which is problematic for two constraints, (11a) and (16b), and which does not constitute a much better coda than *s* from a sonority point of view.

(28) *Derivation of French balance* [balɑs] → Fula *ba* (Fula *ba* [ba])



As shown in (28a), nasal vowels are analyzed as sequences of two segments in underlying representations, i.e. as a vowel followed by a nasal consonant to which it is linked. Many arguments have been recently brought in favor of

this representation in Paradis & LaCharité (1995a,b), and in particular in Paradis & Prunet (1995). The unique behavior of nasal vowels in borrowings constitute one of the main external evidence in favor of this representation: Nasal vowels are the only segments which unpack into two segments in our three corpora of loanwords where French is the source language and in loanword studies in general. Phonemic nasal vowels contained in borrowings systematically yield VN sequences in the target languages that prohibit nasal vowels (provided VN sequences are allowed). This is what we observe in Fula, as we have already seen, in Moroccan Arabic and in English, as shown in (29), and in Kinyarwanda, as we will see.

(29) *French nasals vowels in Moroccan Arabic and English* → VN

Fr. bande	[bA\$ɔ̃]	→	MA banda	*bada	‘gang’
Fr. béton	[betO\$]	→	MA bæt̪ɛn	*bæt̪ɛ	‘concrete’
Fr. bombe	[bo\$b]	→	MA bumba	*buba	‘bomb’
Fr. entrée	[A\$tre]	→	Engl. Ant ^α Ø	*At ^α Ø	
Fr. sangfroid	[sA\$frwA]	→	Engl. sA:nf ^α wA:	*sA:f ^α wA:	
Fr. sans doute	[sA\$ɔ̃dut]	→	Engl. sAnd̪ɛ:t	*sAd̪ɛ:t	

It is noteworthy that the behavior of the nasal vowels contrasts with, for instance, that of the French central vowels *y* and *ɛl*, which do not yield *iu* or *eo* sequences when introduced into languages that do not allow central vowels. The behavior of nasal vowels also contrasts with that of other segments spelled with two characters in French, such as *ou* for [u], *au* for [o], *gu* for [g], *ch* for [ʃ], etc., which, as shown in (30), never unpack into two segments. Thus if one wanted to argue that “unpacking” of nasal vowels into two segments is due to the influence of orthography, he or she would have to face with this discrepancy.

(30) *Examples of sounds written with two characters*

bigraphemes	sounds	pronunciation in:		
		<i>Fula</i>	<i>Kinyarwanda</i>	<i>Moroccan Arabic</i>
<ou>	[u]	[u]	[u]	[u]
<au>	[o]	[o]	[o]	[u]
<gu>	[g]	[g]	[g]	[g]
<ch>	[ʃ]	[ʃ]	[ʃ]	[ʃ]

number of non-adaptation:	0	0%
number of adaptations:	644	99.4%
v\$C... → V ⁿ CV	400/409	97.8% (preservation of N)
other:	9/409	2.2%
v\$# → V(V)#	235/235	100% (deletion of N)

3.2.1 Affordable repairs in Kinyarwanda

As indicated in (32c), nasality in non-word-final position is preserved in 97.8% of cases, i.e. in 400 cases out of 409, since the nasal consonant which constitutes French nasal vowels can attach to the following consonant in Kinyarwanda, and prenasalise it as shown in (33).

(33) v\$C... → VⁿCV: *Cases of nasality preservation*

French		Kinyarwanda	
ampoule	[A\$ <u>p</u>] →	a ^m h uuru	bulb
ambassade	[A\$ <u>bsad</u>] →	a ^m b asadi	embassy
compte	[kO\$ <u>t</u>]	(i)k ^w oo ^{>} hi	account
industrie	[E\$ <u>st</u> ʃi] →	ind uusitiri	industry
fanfare	[fA\$ <u>fr</u>] →	f ^a a ^M faari	band
convoyeur	[kO\$ <u>waj</u> , r] →	k ^w oo ^M v u wajeer	convoy
consul	[kO\$ <u>ʃ</u>] →	k ^w oo ^ʃ iiri	consul
transistor	[tA\$ <u>zist</u> ʃr] →	t ^a a ^{>} zisitoori	transistor
vidange	[viA\$ <u>ʃ</u>] →	vi ^a a ^{>} ʃi	garbage

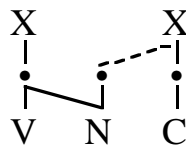
Vowel denasalisation and consonant prenasalisation are illustrated in (34), where it can be seen that it is accomplished in only two steps, i.e. 1) nasal spreading, as in (34b), and 2) nasal delinking, as in (34c), thus respecting both the Preservation and the Threshold principles.

(34) *Vowel denasalisation and consonant prenasalisation (2 repairs)*

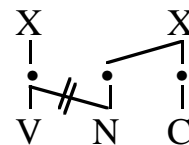
a) French input



b) nasal spreading



c) nasal delinking



3.2.2 Unaffordable repairs in Kinyarwanda

However, the Threshold Principle would not be complied with if the vowel nasality were preserved in the French loanwords in (35) since 1) nucleus insertion, 2) segmental filling of the nucleus, and 3) nasal delinking (from the vowel) would be required, which would represent a repair in more than two steps. I maintain that this is why French borrowings such as *camp* in (35) yields segment deletion, here Kinyarwanda *ka* — where the final nasal is lost — not **kani* with preservation of the nasal consonant (see also Rose 1994 and Paradis & Rose 1995).

(35) $v\$/\# \rightarrow V(V)\#$: Cases of nasality loss

<i>French</i>		<i>Kinyarwanda</i>		<i>gloss</i>
camp	[k \overline{A} \\$/]	→ k \overline{a}	*kan(i)	camp
avion	[avj \overline{O} \\$/]	→ avij \overline{o}	*avijon(i)	plane
coussin	[kus \overline{E} \\$/]	→ k ^w us \overline{e}	*k ^w usen(i)	cushion

The Threshold Principle is supported by other evidence in Kinyarwanda. We have just seen that a nasal vowel in French borrowings yields a prenasalised consonant provided the following consonant is an obstruent. Now the question is: What would happen if the following consonant were not an obstruent but a liquid, since prenasalised liquids are ruled out on a universal basis? There are five underived words which contain such a sequence in French: *genre* [JA\\$/r ‘gender’, *Henri* [A\\$/ri] ‘Henry’, *denrée* [dA\\$/re] ‘food stuff’, *banlieue* [bA\\$/ljfl] ‘suburb’ and *branler* [brA\\$/le] ‘to shake’. To my knowledge, these French words were not borrowed in Kinyarwanda but Paradis & Rose (1995) elicited them with their informants. The pronunciation of the informants is reported in (36), where we can observe that the nasality of the vowel is lost.

(36) *Elicited forms* (cf. Paradis & Rose 1995 for more details)

<i>French</i>			<i>Kinyarwanda</i>
genre	[ʒA\$ʁ]	→	Jaari
denrée	[dA\$eʁ]	→	daare
branler	[brA\$ã]	→	Buraale
banlieue	[bA\$jɛf]	→	Baaliije
Henri	[A\$ʁi]	→	ari

Nasal deletion occurs in spite of the fact that there are other alternatives such as hardening of the liquid. As a matter of fact, this alternative is used in native words: When a nasal + liquid sequence arises in native words, the liquid is simply hardened and yields *nd* as in (37).

(37) *Examples of /nr/ → [> d] in Kinyarwanda*

a) prefixation of the plural class marker *n-*

<i>Singular</i>		<i>Plural</i>
/uru-rabɔ/ [ururaɓo] ‘flower’	/ /	/n-rabo/ → [> daBo] ‘flowers’
/uru-rimi/ [ururiɓi] ‘tongue’	/ /	/n-rimi/ → [> dimi] ‘tongues’

b) prefixation of the 1st sing. pronoun *n-*

<i>Infinitive</i>		<i>1st Sing. Ind. Pres.</i>
/ku-riha/ [k ^w uriha] ‘to reimburse’	/ /	/n-riha/ → [> diha] ‘I reimburse’
/ku-reeba/ [k ^w ureeba] ‘to look’	/ /	/n-reeba/ → [> deeba] ‘I look’

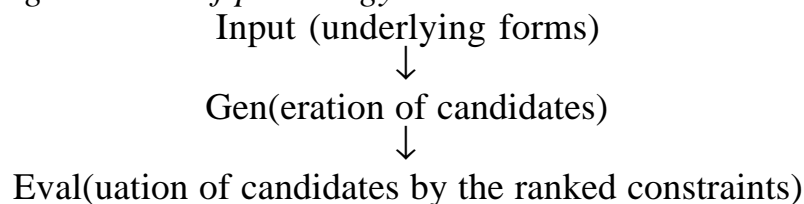
Therefore why French *genre* is not similarly pronounced by our Kinyarwanda informants **Jaãndi* i.e. with a prenasalised hardened liquid, instead of *Jaari*, where the nasal is deleted? In other words, why does nasal deletion apply in (36) instead of liquid hardening? The answer is again the cost that it would entail. Preservation of the nasal in (36) would require too many steps. It would necessitate: 1) nasal spreading to the following consonant, i.e. prenasalisation; 2) delinking of [+nasal] from the preceding vowel (a step which is irrelevant in (37)), and 3) the adaptation of the liquid itself, i.e. its hardening. This would obviously violate the Threshold Principle.

4 Principled deletion in borrowings: A problem for Optimality Theory

Hitherto we have seen that the Threshold Principle yields segment deletion in Fula in 1) unsyllabifiable sequences with an ill-formed segment and 2) nasal vowels before a word-final consonant, and in Kinyarwanda, in 1) word-final nasal vowels and 2) elicited nasal + liquid sequences, i.e. each time more than two processes are required within a given constraint domain. The effect of the Threshold Principle — which crucially requires constraints to be activated in the course of phonological derivations — is persistent in Fula as well as in Kinyarwanda, thus indicating that it is a genuine and important generalisation in linguistics.

Now the question is: How can a filter-based theory such as OT, where constraints are typically uninvolved in the generation of forms, account in a principled way for segment deletion in borrowings? In OT, phonological constraints apply at each morphological level but not after each phonological process since phonological processes in OT are not viewed as serial (Prince & Smolensky 1993). As shown in (38), Eval is the set of universal phonological filters which evaluate the whole candidate set generated by Gen, the place where phonological processes apply.

(38) *The organisation of phonology in OT*



Eval, being a set of filters, has by definition no control over the phonological processes applying in Gen, and therefore cannot count them. Eval does not govern inputs either, even though it does have access to the segmental information contained in inputs, as required by “Containment”.¹²

4.1 Free constraint reranking

Since the filters of Eval are ranked on a language-specific basis, one can first think of positing language-specific constraint reranking to handle segment deletion in borrowings. As shown in (39), it might be proposed that the cases where ill-formed segments are preserved (although sometimes modified), instead of being deleted, are cases where the constraint “Parse segment”

(which requires a segment to be realised) dominates the constraint “Fill” (which prohibits the insertion of a segment) ((39a)), whereas the cases where a segment is deleted are actually cases where “Fill” dominates “Parse segment” ((39b)).

- (39) a) segment preservation cases: Parse segment > Fill
 b) segment deletion cases: Fill > Parse segment

(Parse seg = no segment deletion; Fill = no segment insertion)

However such reranking of the Faithfulness constraints (i.e. Parse and Fill) would not be independently motivated, even though one ranking (39a) could be declared universally less marked than the other (39b) within the Faithfulness family. Such reranking would still be totally descriptive, and would consequently make no prediction concerning segment deletion in borrowings on either universal or even language-specific grounds. Therefore, free reranking is not really an option for a constraint-based theory.

4.2 Mechanical counting between inputs and outputs

An alternative might be to count the number of changes between inputs and outputs. However if we take the Fula and Kinyarwanda borrowings in (40), we notice that they have undergone many changes in comparison with their input — which is at the segmental level, the lexical form of the source language as assumed in 2.3.

- (40) a) Fr. *ingénieur* [Eŋjⁿ„r] → Fula [EnŋjO] ‘engineer’
 delinking (/ unparsing) of [+nasal] from the first vowel
 delinking (/ unparsing) of [-anterior] and [+voice] from the fricative
 delinking (/ unparsing) of [-ATR] from the second vowel
 delinking (/ unparsing) of [-back] from the last vowel
- b) Fr. *industrie* [E^sdys^tr] → Kin. [eⁿdisitⁱndi] ‘industry’
 delinking (/ unparsing) of [+nasal] from the first vowel
 delinking (/ unparsing) of [-ATR] from the first vowel
 delinking (/ unparsing) of [-round] from the second vowel
 insertion of a syllable between *s* and *t*
 insertion of a syllable between *t* and *r*
 filling of the two empty nuclei

For example, French *industrie* [Eʃdys] which yields Kinyarwanda [eⁿdisitiri] ‘industry’ in (40b) has undergone at least six changes (or repairs), and yet no segment deletion has occurred. Thus OT would not want to count the differences between inputs and outputs mechanically, because it would make false predictions, i.e. it would predict segment deletion where it does not occur.

4.3 Counting within constraint domains

Another alternative for OT might be to count the number of changes between inputs and outputs within a given “constraint domain” as in TCRS. This alternative raises a number of questions though.

First, assessment with “domains” means that not only would constraints have access to the segmental information contained in inputs in order to assess the distance between an input and an output (containment), but also that they would divide up the input into domains in order to subsequently count the number of changes within such domains. This means that the role of OT filters would not be limited to assessing the degree of well-formedness of an output in comparison to its input but that they could also detect problems at the input level, a rather unexpected behavior for surface constraints. Furthermore, Eval would presumably still have no control over the processes which apply in Gen, in spite of the fact that the problems would have already been detected at the underlying level, and would subsequently be solved (at least partly) at the surface level — here by the selection of good candidates. Such a move would greatly complicate the framework and entail conceptual contradictions.

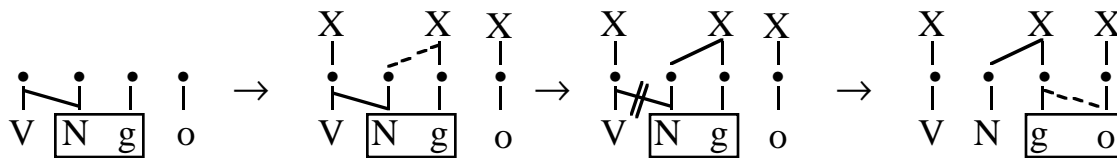
A possible solution would be to maintain that constraint domains of inputs are established only when the output candidates are already generated. Nonetheless, Eval would still have the power of assessing the adequacy of the processes which have applied in Gen vis-à-vis the problems detected in a given input. This would be tantamount to positing that Eval selects candidates which, at the end, have undergone what can be seen as “(minimal) repairs” in Gen.

A second question is: How would domains be established? Why would a segmental constraint be systematically dominated by or embedded within a syllabic one in OT? If this ordering is universal, as suggested by TCRS, OT would have to posit a sort of Phonological Level Hierarchy ((6)). However, is such a scale compatible with current analyses in OT? This universal scale implies that a segmental constraint such as “Parse segment” can never be ranked higher than a syllabic or metrical one.

Finally, what about subsequent changes within a constraint domain occurring after the constraint has been satisfied, which increase the number of changes within the domain but not for the purpose of the constraint? Suppose that a French word such as *lingot* [lɛŋɡo] ‘ingot’ or any similar word (e.g. Fr. *enculer* [ɑ̃kylylɛsodomise], *encoche* [ɑ̃kɔʃ]) is introduced in Kinyarwanda. The adapted form would be [lee^Ng^wo], since as seen in (33), it is systematic for French nasal vowels to prenasalise the following consonant word-internally in Kinyarwanda, which is represented in (41b). Recall that Kinyarwanda disallows codas ((32a)) thus preventing the nasal consonant from surfacing as a segment on its own. As for round vowels, they always labialise a preceding velar, as shown in (41d).

(41) French *lingot* [lɛŋɡo] ‘ingot’ → Kin. [lee^Ng^wo]* (elicited form)¹³

a) French input b) nasal spreading c) nasal delinking d) labial spreading



Given the Precedence Convention in (7), the violated constraint to be taken care of first is the syllabic one, i.e. the one against NC clusters (i.e. against codas), which induces nasal spreading, as shown in (41b). As for the second violated constraint ((32b)), the segmental one embedded in the syllabic one, it triggers nasal delinking, as shown in (41c). Labialisation occurs because of a third constraint, a segmental one, whose domain partly overlaps that of the syllabic constraint as indicated in (41d). In a phonological-derivational framework, this does not cause any problem since a constraint domain disappears as soon as the violated constraint is complied with. However, such an option is not available in a filter-based framework such as OT since there are no phonological stages. Changes are assessed globally as outputs go out of Gen. Therefore, how many changes would OT count in the “Ng” domain in (41)?

These are the issues which lead me to conclude that phonologically induced segment deletion in borrowings is a serious problem for OT, which requires modifications of the framework as significant as abandoning the ban on derivational constraints in phonology. I have endeavored to show that, even though derivational constraints increase the complexity of a framework (they require extra machinery such as the notion of “constraint domain”), they are necessary from the point of view of explanatory adequacy.

5 The Faithfulness constraints vs. the Preservation Principle

In this section, I would like to point out that loanword adaptation is problematic for OT in another respect. One might be tempted to equate the Faithfulness constraints of OT with the Preservation Principle of TCRS in (2), since one of the Faithfulness constraints, i.e. Parse, has similar requirements to those of the Preservation Principle. Both conventions prevent segment deletion from applying to inputs. I would like to show however that Faithfulness and the Preservation Principle make in fact different predictions. This is due to the fact that Faithfulness stands not only for Parse but also for Fill, a constraint which prevents segments from being inserted into inputs. Parse segment prevents segment deletion whereas Fill prevents segment epenthesis, as indicated in (39).

If we return to the statistics in (1), we observe that the effect of Parse in loanword adaptation is much greater than that of Fill since ill-formed segments are very seldom deleted in loanwords whereas segment epenthesis are common-place. As already mentioned, ill-formed segments from the point of view of the borrowing language are adapted (i.e. recast into another shape) in 85.2% of cases, instead of being deleted. Phonologically-induced segment deletion occurs in only 2.3% of the cases. In the case of unsyllabifiable clusters, for instance, it is clear that segment deletions are greatly outnumbered by segment insertions in the four corpora of loanwords. This indicates that Parse segment is systematically ranked above Fill in loanword adaptation, as shown in (42).

(42) *Loanword adaptation*

Parse segment > Fill (universal)

As far as I can see, this ordering is universal. However, there is no internal device in OT which would allow the framework to handle this generalisation on universal grounds, since constraint rankings are conceived as being inherently language-specific in the version of OT proposed by Prince & Smolensky (1993) and McCarthy & Prince (1993); only constraints themselves are universal not their ranking with respect to each other. Thus the generalisation captured by the Preservation Principle in TCRS cannot be straightforwardly dealt with by the Faithfulness constraints of OT, since these constraints can only treat universal segment preservation in loanwords as an accident. In contrast to Faithfulness, the Preservation Principle predicts a universally strong preference for segment preservation over segment deletion. Note that this principle is not restricted to loanwords. Béland &

Paradis (1995) show that it also allows us to capture important generalisations in aphasia.

Some recent attempts to propose universal hierarchies of constraints in OT must be mentioned though. Itô & Mester (1995) posit that some constraint rankings are impossible on a universal basis, whereas Smolensky (1995) proposes a universal markedness hierarchy for segmental constraints. However the question is: Would such markedness hierarchies hold for the two main constraints of OT, i.e. Parse and Fill? If this were the case, it seems that OT would have to abandon the core idea that constraint rankings are (or might be) language-specific, and look for a kind of Phonological Level Hierarchy as that of TCRS in addition to incorporating a theory of markedness similar to that of principles-and-parameters approaches in order to rank constraints on a universal basis.

NOTES

- * I would like to thank the audience at the Royaumont conference, in particular Paul Smolensky and Moira Yip for challenging questions. I am indebted to Peter Avery, Outi Bat-el, Renée Béland, Anna Bosch, Barbara Bullock, Carrie Dyck, John Goldsmith, Larry Hyman, John Ohala, Jean-François Prunet and Caroline Wiltshire for illuminating comments and / or useful discussions on borrowings. I would like to offer special thanks to Sharon Inkelas for very detailed written comments and stimulating discussions, and of course to Darlene LaCharité. I am also grateful to the audience at the MOT Conference on Contrast in Phonology, held at the University of Toronto in February 1994, those at the Department of Linguistics of the University of California at Berkeley, Santa Cruz and Los Angeles, Stanford University, to that at the Trend Colloquium, held at Berkeley in the fall 1994, and that of the 1995 Chicago linguistic Society Meeting, where related papers were presented. These presentations allowed me to benefit from the comments of many people, in particular from those of Young-Mee Yu Cho, Bruce Hayes, Larry Hyman, Junko Itô, Wil Leben, Armin Mester, Paul Kiparsky, Janet Pierrehumbert, Donca Steriade and their students. I acknowledge SSHRC grants # 410-90-0575 and # 410-94-1296 and FCAR grants # 90-NC-0383 and 95-ER-2305
- [1] A “form” is a borrowing pronounced by an informant, i.e. the concrete realisation of an abstract entity, the borrowing. This distinction between “form” and “borrowing” is essential since informants do not all realise borrowings in the same way.
- [2] Deletion applies mainly to vowels at the beginning of words — Moroccan Arabic does not allow onsetless syllables — which contain over three syllables. We suspect the influence of a metrical constraint here since, when the word is shorter, other strategies such as the insertion of a glottal stop are resorted to in order to preserve the word-initial vowel.

- [3] These general statistics report the results from the French-Fula corpus and those from the English-Quebec French one since the two other corpora have not been thoroughly analyzed yet.
- [4] This is at least the limit we have observed in the four corpora studied by Paradis & LaCharité (1995a,b). Should it be different in other languages, the last part of the Threshold Principle would have to be parametrised.
- [5] The alveo-palatals *c* and *ç* are in fact affricates which have traditionally been transcribed as plain palatal stops in the literature. For more convenience, the traditional notation is retained here.
- [6] Most data are from our Futa Toro informant, i.e. a primary source of data. The statistics in (1d), however, gather different sources, primary and secondary.
- [7] It could be that what is in fact inserted is a [-consonantal] root node, not a nucleus, which is syllabified by default. However, the repair count would be the same: two repairs.
- [8] The final *i* in *piu:ri* is actually a class marker (/piu:r-i/) since the definite article is *ⁿdi* (e.g. *piu:ri ⁿdi* ‘the octopus’).
- [9] There are also elicited words such as French *circuit* ‘circuit’ and *aujourd’hui* ‘today’ which were pronounced respectively [sirki] and [OsOrdɪ] by our Futa Toro informant.
- [10] Voiceless stops have a strange distribution in Fula. They are extremely rare word-initially, and tend to be geminated word-internally (see Paradis 1992: 117).
- [11] Prenasalised, palatalised and labialised consonants, which are phonologically derived in Kinyarwanda, are excluded from the table.
- [12] *Containment*: “No element may be literally removed from the input form. The input is thus contained in every candidate form”. (Prince & Smolensky 1993: 20)
- [13] Vowel lengthening is ignored in the schemes since it is irrelevant here.

REFERENCES

- Béland, Renée & Carole Paradis (1995). Principled Phonological Dissolution in a Primary Progressive Aphasia Case. Ms, Montreal University and Laval University.
- Haugen, Einar (1950). The Analysis of Linguistic Borrowings. *Language* 26: 210-231.
- Itô, Junko & Armin R. Mester (1995). The Core-Periphery Structure of the Lexicon and Constraints on Reranking. *UMass Occasional Papers in Linguistics*. (forthcoming)
- Kimenyi, Alexandre (1979). Studies in Kinyarwanda and Bantu Phonology. *Current Inquiry into Languages and Linguistics* 33. Edmonton: Linguistic Research Inc.
- Mougeon, Raymond, Edouard Beniak & Daniel Valois (1985). *Variation in the Phonological Intergration of Loanwords in a Bilingual Speech Community*. Toronto: Center for Franco-Ontarian Studies.

- Mugesera, Léon (1987). Principes et méthodes de la recherche terminologique au Rwanda. Essai d'établissement d'un modèle de création terminologique en langue nationale (ikinyarwanda). PhD dissertation. Laval University.
- McCarthy, John & Alan Prince (1993). Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms, University of Massachusetts at Amherst and Rutgers University.
- Paradis, Carole (1988a). Towards a Theory of Constraint Violations. *McGill Working Papers in Linguistics* 5: 1-43.
- ____ (1988b). On Constraints and Repair Strategies. *The Linguistic Review* 6: 71-97.
- ____ (1992). *Lexical Phonology and Morphology: The Nominal Classes in Fula*. New York: Garland Press.
- Paradis, Carole & Darlene LaCharité (1995a). Preservation and Minimality in Loanword Adaptation. Ms, Laval University and University of West Indies.
- ____ (1995b). Saving and Cost in Loanword Adaptation: Predictions of the TCRS-Phonological Model. *McGill Working Papers in Linguistics*. (forthcoming)
- Paradis, Carole & Caroline Lebel (1994). Contrasts from Segmental Parameter Settings in Loanwords: Core and Periphery in Quebec French. *Proceedings of the MOT Conference on Contrasts in Phonology*. Toronto Working Papers in Linguistics 13: 75-94.
- Paradis, Carole & Jean-François Prunet (1995). Nasal Vowels as Two Segments. Ms, Laval University and University of Quebec at Montreal.
- Paradis, Carole & Yvan Rose (1995). Préservation et perte segmentale dans les emprunts français en kinyarwanda. *Proceedings of the 1995 Canadian Linguistic Association*. Toronto Working Papers in Linguistics. (forthcoming)
- Poplack, Shana, David Sankoff & Christopher Miller (1988). The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation. *Linguistics* 26: 47-104.
- Prince, Alan & Paul Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms, Rutgers University and University of Colorado at Boulder.
- Rose, Yvan (1994). L'adaptation des voyelles nasales en kinyarwanda: l'effet du principe de préservation. Pp. 161-165 in Belyazid *et al.* (eds.) *Actes des 8e Journées de linguistique*. Quebec: CIRAL.
- Smolensky, Paul (1995). On the Internal Structure of CON in UG. *GLOW Newsletter* 34: 70.