

**Diachronic Evidence in Word-Final Empty Structure and its Effects:
RS and CL in Maremmano Italian and Lhasa Tibetan**

Shanti Úlfsbjörninn

Department of Linguistics

The School of Oriental and African Studies

Thesis Submitted to the University of London in Partial Fulfilment of the
requirements for the degree of Bachelor of the Arts (hons.)

May 2007

Contents

0. Abstract
1. Introduction
2. Part One
 - 2.1. The Flavour of the Framework
 - 2.1.1. The tiers
 - 2.1.2. The clusters
 - 2.1.3. The geminates and long vowels
 - 2.2. The Epenthetic Glottal Stop (Kisar)
 - 2.2.1. (1) Amending the OCP (Selayarese, Indonesian)
 - 2.2.2. (2a) 'Words must begin in consonants' (Tamil)
 - 2.2.3. (2b) 'Words must end in consonants' (Cupeño)
 - 2.2.4. (3) None of the above with epenthetic glottals (Thai, Vietnamese)
 - 2.3. The Constraint and 'Pointed ON-Pair Repair' Parameter (Turkish)
 - 2.3.1. Related side note on Charette (forth.) (Dinka)
3. Part Two
 - 3.1. Lhasa Tibetan
 - 3.2. Maremmano Italian
 - 3.3. Cross-Linguistic Comparisons, sections one and two
4. Part Three
 - 4.1. Raddoppiamento Sintattico vs. Compensatory Lengthening
 - 4.1.1. Maremmano Italian RS
 - 4.1.2. Lhasa Tibetan CL
5. Part Four
 - 5.1. Problems with Lowenstamm's Analysis of RS and CL (Biblical Hebrew)
 - 5.2. The Solution to Lowenstamm's (1996; 1999) Problem
 - 5.3. The 'Pointed ON-Pair Repair' Parameter Refined
 - 5.4. Why are Biblical Hebrew and Maremmano Italian Different?
 - 5.4.1. Maremmano Italian with no branching onsets
 - 5.4.1.1. Evidence from Maremmano cluster lenition
6. Conclusion
7. Bibliography

**Diachronic Evidence in Word-Final Empty Structure and its Effects:
RS and CL in Maremmano Italian and Lhasa Tibetan**

Shanti Úlfsbjörninn
The School of Oriental and African Studies

0. Abstract

Part one will outline the theoretical approach to the paper and explain the current theoretical dealings with empty structure; it will also explain the basic theory behind epenthetic glottal stops and motivate the justification of the constraint: a p-licensed empty nucleus cannot license an empty pointed onset. Part one will go on to explain that this constraint is actually nothing other than a straightforward Empty Category Principle violation although due to its specific location within the domain (final) other factors are involved, such as the word-final parameter. Part one shows that different languages behave in one of two different ways in repairing this specific ECP violation and as such these strategies can be parametised.

Part two shows the diachronic changes in Lhasa Tibetan and Maremmano Italian as regards word-final segmentally-empty syllabic structure. These languages historically show the loss of segmental material although retaining its syllabic structure. It is also shown, following the UG tradition, that totally disparate languages can behave identically with regard their repair strategies.

Part three then shows how the utterance-final empty structure interacts with elements placed after it within their domain. Theoretically speaking this has the effect of taking our ECP violation away from domain-final to domain-medial position. What we would expect is that Maremmano Italian and Lhasa Tibetan, two languages which behave identically domain finally, would react the same way to similar circumstances. This however isn't borne out in the facts and in Maremmano Italian the ECP violation repair strategy is Raddoppiamento Sintattico¹, while in Lhasa Tibetan it is Compensatory Lengthening. The data shows us that to properly parametise the repair strategies we must have a differentiation between domain-final and domain-medial applications. These two locations for the same parameter are independent of each other and so if Maremmano Italian and Lhasa Tibetan select for the same setting domain-finally they can (as we will see) select a different setting domain-medially.

Having re-defined the parameters responsible for the repair strategies which superficially we call RS and CL there is an obvious port of call. Lowenstamm (1999) claims that RS and CL in Biblical Hebrew are motivated not by word-final diachronically visible segmental loss and syllabic retention rather by an empty ON-pair (for him 'CV-site') located word-initially. This, he argues, is universal. This study then goes to show that Lowenstamm's rules for generating RS and CL with only one CV-site doesn't output the correct data (as found still within Lowenstamm (1999)). The justifications for his structure and outputs are shown to be falsified a priori. Part four will then go on to show that actually the CV-site is essential to Biblical Hebrew RS and CL but that one CV-site is not enough. Through diachronic

¹ Henceforth RS and compensatory lengthening: CL

*segmental loss and syllabic retention however, we can propose that the RS and CL trigger ‘ha’ being originally *han provides us with justification for a word-final empty-ON pair just as with the Maremmiano Italian and Lhasa Tibetan. With this study’s word-final empty structure and Lowenstamm’s (1999) initial CV-site it is possible to output the correct data (in Lowenstamm 1999).*

The initial-CV site is important to understand Biblical Hebrew and supposedly universal, at least within languages with real consonant clusters (Kula 2006), so Lhasa Tibetan is immune to the following question: why doesn’t Maremmiano Italian, supposedly a branching onset language, behave like Biblical Hebrew. If Kula (2006) is right we could understand this problem by Maremmiano Italian not having true branching onsets and thus not requiring a word-initial CV-site. This would still leave us with the empty word-final structure and the correct outputs. To back up this radical position, there is lenition evidence proving this hypothesis exactly correct. Branching onsets in Maremmiano Italian are most probably best analysed as bogus clusters seen that, in post-vocalic lenition, the supposed governor reduces to a lower elemental complexity than its supposed governee thus violating the complexity condition of government (a pivotal notion in Government Phonology to account for branching onset typology (Kaye et al. 1985; Charette 1990, Harris 1990))².

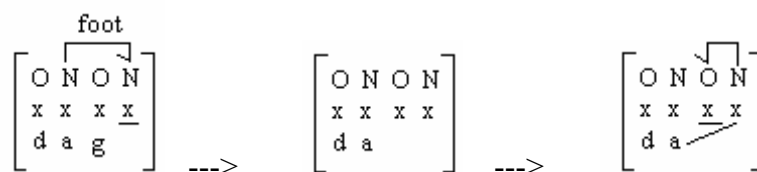
1. Introduction

This paper will explore a subset of languages for which diachronically observable segmental loss has resulted in certain words and/or whole major categories ending in segmentally empty ON-pairs. Utterance-finally these ON-pairs can surface with long vowels ie. Turkish (Charette 2006) or an epenthetic glottal stop eg. Lhasa Tibetan or Maremmiano Italian (DeLancey 2003). The following argument is also identical for languages which without specific historical loss of segments adopt the ‘Han-Template’ ie. Beijing Mandarin (Goh 1997; Kaye 2000; Xu 2001).

Turkish: ‘dag’ /da/ [da:] (Charette pc.)

Mandarin Chinese: ‘ma’ /ma/ [ma:] (Goh 1996)

Structures from Charette (pc.)



In the above we can see how segmental loss and syllabic retention results in an empty ON-pair word-finally. Although Turkish allows for a p-licensed word-final empty nucleus we see that forms like *[da] and *[ba] are disallowed. The argument, goes that in order to have an empty pointed onset, its adjacent nucleus cannot itself be p-licensed. Therefore, we have vowel spreading from N1 to N2 which leaves N2’s

² This is even true of strict-CV where directionality is different but complexity maintained (Scheer 2004)

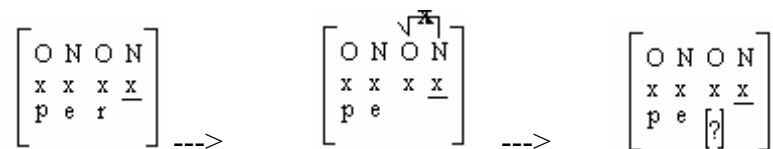
licensing potential unexhausted. Phonetically, the surface result, is a long vowel for every CV syllable.

The other logical option is also attested and will be the focus of this study. It is possible for N2 to remain p-licensed and for there to be a pointed onset although in such cases the ECP forces the pointed onset to undergo phonetic interpretation. In these instances we can see how the glottal stop, which in this study will be understood to be an underspecified consonant (cf. Lombardi 2002), is a common alternant with zero and the consonant equivalent to the schwa in Government Phonology (Charette 1988 et seq.).

Lhasa Tibetan: 'bod' /pö/ [pö?]

Maremmanno Ita: 'per' /pe/ [pe?]

Structures are similar to Denwood (1999) for Thai 'dead syllables'



This paper will then focus on how the above illustrated empty structure motivates Raddoppiamento Sintattico in Maremmanno Italian and Compensatory Lengthening in Lhasa Tibetan. In both, this is accompanied by a glottal-zero alternation in which the words utterance finally surface with a word-final glottal stop which in domain medial environments goes to zero.

Maremmanno Italian (Úlfsbjörninn 2006b)

Standard Ita: [sono] [per] 'are' and 'for'
Maremmanno: [so?] /soxx/ [pe?] /pexx/

1. [i fjo:ri so **p:e** l:a don:a] (the flowers are for the woman)

Lhasa Tibetan (DeLancey 2003)

Classical Tibetan: [bod] [red] 'Tibet' and 'be'
Lhasa Tibetan: [phö?] /phöxx/ [re?] /rexx/

2. [k_ho phö:-pa re:-pä:] (Is he Tibetan?)

2. Part One

2.1. The Flavour of the Framework

2.1.1. The tiers

This paper is going to be written in a modified form of strict-CV³ phonology (Lowenstamm 1996; Scheer 2004) which itself is a branch of Government Phonology (Kaye, Lowenstamm and Vergnaud 1985; Kaye 1990; Charette 1990). I will assume the skeletal tier to be essential to our understanding of certain phonological processes and thus its representations. Therefore, what strict-CV would see as (s1) this study will present as (s2).

$$\begin{array}{cc}
 \text{(s1)} & \text{(s2)} \\
 \left[\begin{array}{c} \text{c} \quad \text{VCV} \\ \text{e} \quad \text{r} \quad \text{o} \end{array} \right] & \left[\begin{array}{c} \text{ONON} \\ (\text{x}) \text{xxx} \\ \text{e} \quad \text{r} \quad \text{o} \end{array} \right]
 \end{array}$$

The optionality of any certain skeletal points adds necessary complexity to the representations and appropriately reflects diachronic loss of segments and their interference with synchronic phonological processes, such as h-aspiré in French (Charette 1991).

Charette shows the hiatus trigger for deletion in French applies above the melodic tier, that is melodically speaking [ero] ‘hero’ and [ami] ‘friend’ are identical. However they don’t behave the same way with one reacting to an OCP violation while the other is immune to it. Government phonology doesn’t allow for the belief that there could be OCP at the constituent tier as ‘onset licensing’ (Kaye et al. 1985) requires every nucleus to have at least a pointless onset. This onset constituent in French is essential to explain liaison phenomena: [peti(t) ami] → [peti t ami] (Charette p.c). The floating consonant has a buffer to which it can attach; if the theory didn’t have an empty onset at the beginning of the word [ami] we would have to posit some form of structural epenthesis which would be a catastrophic violation of the *projection principle* (Kaye et al. 1990) leading to a vastly over-generating theory. Therefore, the constituent tier is always alternating O, R(N) and melodically [ero] and [ami] are essentially identical. The problem is solved by assuming that eventhough both [ero] and [ami] start with an onset the former’s onset is pointed contrary to the latter’s. This has the effect of creating an OCP violation in the form of two adjacent nuclear points: [lə ami]. [ero] however wouldn’t have these two adjacent nuclear points and as such not trigger OCP:

$$\begin{array}{cc}
 \text{(s2a)} & \text{(s2b)} \\
 \left[\begin{array}{c} \sqrt{} \\ \text{ONONON} \\ \text{xxx} \text{---} \text{xxx} \\ \text{l} \quad \text{a} \quad \text{e} \quad \text{r} \quad \text{o} \end{array} \right] & \left[\begin{array}{c} \text{ocp} \\ \text{ONONON} \\ \text{x} \text{ } \boxed{\text{x}} \text{ } \text{x} \text{ } \text{x} \text{ } \text{x} \\ \text{l} \quad \text{a} \quad \text{a} \text{ } \text{m} \text{ } \text{i} \end{array} \right]
 \end{array}$$

Without the skeleton a sCV analysis experiences some major problems. sCV collapses the constituent and the skeletal tier creating onsets (C) and nucleic points (V).

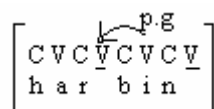
³ Henceforth sCV

Consequently the independence of the two autosegmental tiers of sGP is lost. In the previous paragraph, however, we showed how it is this very independence of the two autosegmental tiers that allows an analysis of liaison and vocalic hiatus deletion in French. If we posit that [ami] has an empty C at the beginning of the word we can keep our liaison facts but loose the reason why [ami] is different to [ero] in hiatus terms. Conversely, if we posit that [ami] doesn't have an empty C beginning the word we can keep the hiatus data but loose the liaison data which would have to be explained in a manner totally disharmonious with the *projection principle*. Some may argue that a little and a big C would be sufficient to explain the facts. This, however, would be a total push into arbitrary abstraction in which 'little Cs' are somehow transparent to one process (hiatus OCP) while visible to another (floating consonant attachment). This would be too permissive a notion to be scientifically useful. Thus in sCV we can see that these French facts cannot be satisfactorily accounted for. The solution, however, is rather obvious. One can keep the strings of antisymmetrical, non-branching, lateral relationships between contrastive constituents (Lowenstamm 1996; Scheer 2004) but have those contrastive constituents be the O and the N with a mediating skeletal tier between themselves and the melodic tier.

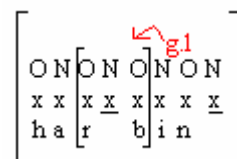
2.1.2. The clusters

Identically to sCV true consonant clusters can be represented without branching and in a manner identical to Cyran (2003)⁴ which uses Government Licensing (Charette 1990) to close domains: [CvC]←V with the singular rule for p-licensing by which a nucleus intervening in onset-to-onset government is automatically p-licensed (cf. Cyran 2003, Charette (in press)). This nucleus, however, would still be projectile (contra Scheer 1998b) in which the p-licensed nucleus would still project into the *nuclear projection* explaining its visibility to stress systems such as Japanese pitch accent (Yoshida 1999). In this study the closed domain hypothesis is also used for 'coda'-onset consonant sequences⁵ (cf. Cyran 2003; Charette in press, contra Scheer 1998b, 2004) The word [harbin] therefore looks like (s3) to standard sCV and (s4) to the modified view of sCV:

(s3)



(s4)



⁴ Identical but for the fact that Cyran has chains of CV pairs while I have chains of ON pairs with a skeletal tier beneath it.

⁵ Which are clearly not just beautifully systematic bogus clusters and are only acquired after proper-government (the licensing needed for bogus clusters) (Pan and Snyder 2003; Úlfssbjörninn 2005)

2.1.3 The geminates and long vowels

Geminate vowels and consonants are seen as empty ON-pairs⁶ which are spread into by a surrounding element producing their long sound, the structure of (s5) is essentially from Lowenstamm (1996):

$$(s6) \quad \left[\begin{array}{cccccc} \text{O N} & \text{O N} & \text{O N} & & & \\ \text{x} & \text{x} & \text{x} & \text{x} & \text{x} & \text{x} \\ (\text{C})\text{y} & & & & & \text{z}(\text{V}) \end{array} \right]$$

In (s6) a long vowel is formed if ‘y’ spreads into the empty structure and a geminate forms if ‘z’ spreads. In order for a long vowel to spread the empty ON sequence must be licensed by an un-p-licensed nucleus to its left (Lowenstamm 1996): *[ka:tpi] vs. [ka:tupi]⁷.

2.2 The Epenthetic Glottal Stop

This paper doesn’t have the scope to explore whether or not there can be such a thing as a lexical glottal stop encoded with an element (?) (Kaye 2000). Recently and continuously since Jensen (1994) the element (?) has been under attack in GP, most notably by Bachmaier, Kaye and Pöchtrager (2006), Úlfsbjörninn (2006), Pöchtrager (2006). The following data from Kisar shows that however theoretically handles the glottal stop there is a difference between its use as a lexically encoded segment and a surfacing of licensing restrictions.

Kisar (Christiensen & Christiensen 1992)

We see in Kisar reduplication that lexically encoded glottal stops remain in the representation while epenthetic glottal stops go to zero:

- | | | | | |
|-------------|---------|------|------------|-----------------|
| (5) /?umu/ | [?umu] | ---> | [?um?umu] | (very dense) |
| (6) /?alam/ | [?alam] | ---> | [?al?alam] | (early morning) |
| (7) /eni/ | [?eni] | ---> | [?enieni] | (this one here) |

In (7) we see how when the beginning of the word is utterance-initial it has a glottal stop it could either be lexical or not. If it is not lexical then the glottal stop goes to zero once it is removed from utterance-initial position. This data evidences a distinction between lexical and epenthetic glottal stops which is also found in a number of other languages, ie. Mayan: Mam (England 1983) to the Paleo-Siberian: Koryak (Kenstowicz 1976).

Further to the above data, glottal stops have often been analysed as epenthetic. In Lombardi’s (2002) paper they are compared with coronals and shown to be the more primitive epenthetic consonant (along with /h/).

⁶ in which the onset is optionally pointed

⁷ This though is open to parametric variation there are bucket-fulls of languages which allow [CV:C] in which a long vowel is followed by a consonant and an empty vowel.

Epenthetic glottals can be seen to participate in three main phonological activities:

1. Amending an OCP violation
2. Satisfying 'phonosyllabic rules' such as: words must begin/end with consonants
3. Surfacing as failure of an p-licensed nucleus to license an empty pointed onset

I will show at the end of this section of part one that the underlying phonological motivation for (2a/b) is actually (3). Languages which behave in the third manner clinch it by showing the same behaviour as (2a/b) but without any possible 'phonosyllabic rule'. I.e. Thai behaves like (2b) *and* (3) and this cannot possibly be due to a rule such as 'words must end in consonants' as a great many words end in long vowels (Denwood 1999). Thus languages which behave like (2a) can be given an underlying representation with their glottal stop surfacing a symptom of the same licensing failure that forces the glottal to surface in (3). Essentially, (3) allows us to remove the arbitrary rules such as 'words must end in Cs', of course this statement is undeniably true, however, languages like (3) allow us to produce a theory.

2.2.1. (1) Amending OCP

Salayarese (Mithun & Basri 1986)

Has an OCP constraint by which Vx-Vy sequences are tolerated but Vx-Vx are interceded by an epenthetic glottal.

- | | |
|----------------|-----------------------|
| 1a. ri-uraNi | (I.hon accompany him) |
| 1b. ku-?-uraNi | (I accompany him) |
| 2a. ri-?-inuNi | (you.hon drink it) |
| 2b. ku-inuNi | (I drink it) |

Indonesian (Cohn & McCarthy 1994)

Which usually has V[high]-V sequences broken up by their respective glide has Vx-Vx sequences intervened by an epenthetic glottal.

- | | | | |
|---------------|-------|-------------|----------------------|
| 3. /baca-/ | /-an/ | [bacaʔan] | (reading) |
| 4. /m@mandi-/ | /-i/ | [m@mandiʔi] | (bathing repeatedly) |

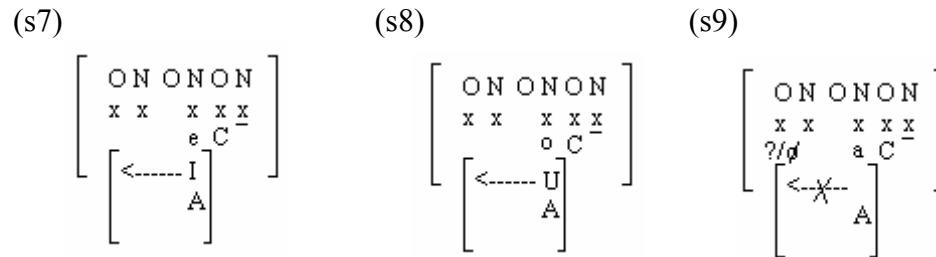
Although the glottal alternation ? → ∅ as demonstrated in Selayarese and Indonesian are not the concern of this essay the facts of these languages shows us how glottal stops can be simple phonetic realisations of onsets ie. underspecified consonant. In the above cases these glottals singularly surface to satisfy OCP.

2.2.2 (2a) 'Words must begin with consonants'

Tamil (Christdas 1988)

Words obligatorily begin with a consonant. If the word-initial lexical vowel is a high vowel then the element spreads from the nucleus to the onset forming [y] from [i] &

[e] and [w] from [u] & [o]. If, however, the initial vowel is [a] then an epenthetic glottal has to be inserted:



[A] is frequently not a spreading element (Turkish, Mongolian, Finnish, Icelandic, Honkeui To...⁸) and as such the initial ON-pair remains segmentally empty. However, in Tamil this empty ON-pair surfaces with phonetic material in the initial onset. We are shown here, for the first time, that when a pointed ON-pair is found and its nucleus is p-licensed, its preceding onset *must* be given phonetic interpretation. That is, Tamil shows us that a p-licensed nucleus cannot license a pointed onset to be empty. Therefore, due to the phonological ECP (Charette 1988, Kaye 1990) the unlicensed empty onset must receive phonetic interpretation.

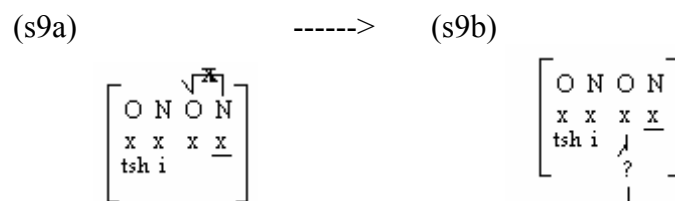
2.2.3. (2b) ‘Words must end in consonants’

Cupeño (Crowhurst 1994)

Mayanists talk of a template for Mayan words which is bi-moraic and comprised of [CVC] templates. This lead to the generalisation that in Cupeño words which are lexically only [CV] have an epenthetic glottal fill the missing C# slot.

5. /tSi/ ---> [tSi?] (gather)
6. /hu/ ---> [hu?] (fat)
7. /kwa] ---> [kwa?] (eat)

‘Coda’ licensing (Kaye 1990) convincingly claims that the word-final consonants at the end of the Mayan /CVC/ template and (5)-(7) are onsets which precede the word-final empty nucleus. Which in Cupeño is p-licensed by the word-final parameter. We can analyse Cupeño epenthetic glottal as being a segmentally empty pointed onsets followed by the required word-final empty nucleus. This is identical to the Tamil case, only this time, the ON-pair is word-final:



Just as in Tamil we can see that when a word-final nucleus is p-licensed it is itself unable to license a pointed onset to be empty. Thus the ECP forces the unlicensed empty category to be realised with phonetic material. This is the identical structure as

⁸ A notable exception is ChiChewa (Harris and Linsey 1995).

with Turkish and Beijing Mandarin *[CV] words, however, the latter two languages opted to license the pointed onset to be empty. In order to do this the languages must suspend p-licensing from the word-final nucleus hence filling it with material from N1 to N2. N2, now a non-p-licensed nucleus, is able to license the pointed onset to be empty.

Tamil and Cupeño in contrast to Turkish and Beijing Mandarin both show that p-licensing and the licensing of a segmentally pointed onsets are linguistically in opposition with languages choosing one strategy over the other.

Despite the credible analysis presented just above it could still be argued that the epenthetic glottal in these languages was surfacing due to *ad hoc* phonosyllabic rules such as ‘words can’t end in Cs’. The following data from Vietnamese and Thai show that, in fact, we get the same phenomena as with Tamil and Cupeño without, however, any possibility for recourse to phonosyllabic ‘rules’. In both Vietnamese and Thai, which exhibit word-initial and word-final epenthetic glottals respectively, words can just as respectively begin and end in vowels. As such we will be able to see that the motivation for a constraint such as: p-licensed nuclei not licensing pointed onsets to be empty’ will be the underlying reason for the epenthetic glottal in both (2a/b) and (3) thus eliminating the need for phonosyllabic rules altogether.

2.2.4. (3) None of the above with epenthetic glottals

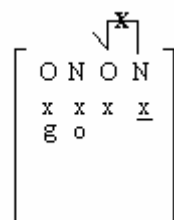
Thai (Denwood 1999)

Thai has a two syllable types [CV:] to [CVC]. There are, however, syllables with only [CV] lexically filled Denwood (ibid.). Denwood demonstrates that the form [CV], as in many Chinese languages, is disallowed and every lexical item fits into a [CVCV] template identical to the Han template (Goh 1996). Thai also has a large class of syllables which the literature calls ‘dead syllables’.

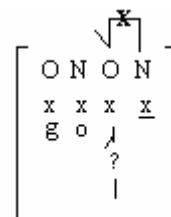
8. /go/ ----> [go?] (Island)
9. /nit/ ----> [ni?] (a little)⁹

One theoretical option is to spread the vowel from N1 into N2 as Honkeui To does (Xu 2001) making all underlying [CV] into [CV:] (idem. Turkish and Beijing Mandarin). Thai however, opts for the other theoretical possibility, leaving N2 p-licensed it phonetically realises O2 which thus surfaces as an elementally underspecified glottal stop (s12b).

(s12a)



(s12b)

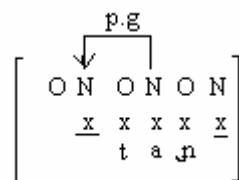


⁹ The consonants here ending the word are historical relics and orthographic only.

Vietnamese (Úlfsbjörninn 2006)

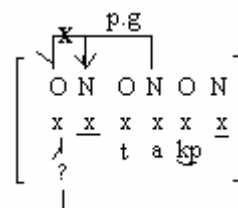
Vietnamese behaves like Tamil in the sense that although not all words must begin with glottal stops one finds glottal stops exclusively word-initially and even before the main consonant of the word (Thompson 1959). This implies that the lexical template for Vietnamese cannot be just the ‘ONON’ *Han template* (Goh 1996; Xu 2001) but there exists an extra ON-pair to the beginning of each word¹⁰. This allows us to understand all Vietnamese words as products of a unified template in which the difference between a glottal-initial word vs. a non-glottal-initial word is the presence or absence of a skeletal point of the first onset and its effects on licensing.

(s10)



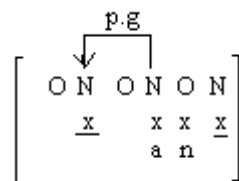
[taJ]

(s11)



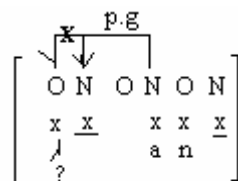
[?takp]

(s10a)



[an]

(s10b)



[?an]

Again, the reason this glottal stop is surfacing in Vietnamese is claimed to be due to the fact that an empty p-licensed nucleus cannot license a segmentally-empty pointed onset.

The conclusion to be drawn from Vietnamese and Thai is that one finds epenthetic glottals word-initially and word-finally even when there is no phonosyllabic ‘rule’ forcing their appearance. In other words, whatever motivates the epenthetic glottal (in non OCP violation conditions) it isn’t a special rule, rather, as we saw for Tamil, Cupeño, Vietnamese and Thai the conditioning factor for glottal stop epenthesis is a pointed onset with a p-licensed licenser.

2.3. The Constraint and ‘Pointed ON-Pair Repair’ Parameter

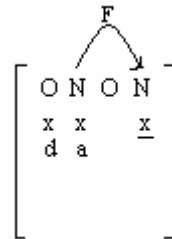
Charette (2006; in press) shows that minimal word-size restrictions in Turkish can be explained by positing a word-final trochaic foot. This would entail that the trochaic foot is necessary for primary stress and every lexeme has this stressed head as its domain head (in line with Kaye et al. 1990; Harris 1997). Symptomatically, this

¹⁰ Most interestingly the nucleus of this ON-pair has to always be empty echoing the word-initial site (Úlfsbjörninn 2006)

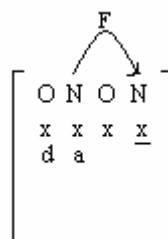
trochaic foot forces the smallest words to be at least comprised of two nuclei which are in a dependency relationship: the first dominating the other, the first phonetically expressed, the latter obligatorily silent. Syllabically speaking, two ON-pairs. What this analysis cannot independently explain, however, is the restriction on *[CV] syllables which should, in our current understanding of the theory, be allowed. Ergo (s13a) is a theoretically licit structure which is ungrammatical in Turkish¹¹.

Turkish 'dag' /da/ [da:] (mountain) (Charette 2006)

*(s13a)



(s13b) UR



(s13a) satisfies Charette's (2006, in press) constraint in violation of the facts. To explain this fact (and something which strict-CV could not do) is posit the structure (s13b) as opposed to (s13a). In (s13b) the second onset is pointed and as such to p-license the word-final nucleus is to deprive this empty category of licensing. Turkish however (opposite to Thai) is a language which faced with this situation chooses to suspend the p-licensing of the word-final nucleus (assigning it the phonetic material of the first nucleus) which would restore its licensing potential. Therefore, [da:] would be able to license the pointed onset rendering the structure licit. The exact same argument can be extended to a great number of languages such as Beijing Mandarin, Cantonese, Mayan languages etc... in which we all see that the second onset is always pointed (Goh 1996; Kaye 2000; Xu 2001).

The above are repair strategies for the universal constraint against two empty categories, where one is the licenser of the other and itself p-licensed. Essentially, this is a specific phono-syllabic environment which violates the ECP. The repair strategies of this particular illegal configuration are parametrically expressed in language as represented by Cupeño-type languages and Turkish-type languages. This parameter can be summarised as follows:

Pointed ON-Pair Repair (to be revised)

- a) spread content of N1 to N2 [yes, no]
- b) insert epenthetic glottal stop [yes, no]

The nature of these parameters is that they are mutually exclusive, so by activating one you dis-activate the other. For economy of theory one could actually claim that the only real parameter is (a) and that the insertion of an epenthetic glottal stop is a

¹¹ There are of course problems with positing that two nuclear points being adjacent as these would in some languages trigger OCP deletion processes like Reduction (Kaye and Gussman 1993; Charette 2006). This however, was never explicitly made a universal and is problematic generally, therefore I don't consider (s13a) to be universally ungrammatical rather, I posit this as an alternating structure for s13b which is disallowed in Turkish but as we will see, seems to work for Dinka.

universal grammar principled effect to satisfy the ECP. That is, languages can do (a) and if so they leave the pointed onset licensed and empty. Or they don't do (a) in which case the word-final nucleus can be p-licensed by the word-final parameter and the pointed onset simply receives phonetic interpretation, this being automatic isn't a parameter at all so the revised 'Pointed ON Pair' parameter is as follows:

Pointed ON-Pair Repair (revised, to be revised at end of section three)

- a) spread the content of N1 to N2 [yes, no]

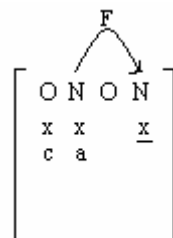
2.3.1. Related side note on Charette (in press)

Charette's trochaic foot languages, the '[CVC]-languages', open up a theoretical possibility for them to have [CV] syllables. Essentially, those without a skeletal point in the second onset (s13a) would be 'CVC languages' *without* a restriction on CV syllables. There do indeed seem to be languages like this. Dinka a Nilo-Saharan language of Southern Sudan has its maximal word [CVC] (Malou 1988). Its compounds are all [CVC-CVC] in which both nuclei are stressed (ibid.). Further supporting the CVC-CVC analysis for compounds is that there is no nasal assimilation on the morpheme boundary ie. [bon-bar], *[bom-bar] (cf. Malou 1988).

Dinka, however, does have a number of lexical nouns and verbs which are [CV]. With the understanding of Charette's foot as a theoretical concept as described in the previous section we can understand Dinka CV syllables to be just like (s13a). Below illustrated with Dinka data: (s14).

Dinka 'ca' /ca/ [ca] (milk) (Malou 1988)

(s14)



In (s14) we can see that the second onset is empty and pointless while its adjacent nucleus is p-licensed by the trochaic foot (cf. Charette in press). Interestingly, in related languages the words which are [CV] in Dinka surface as [CVC] and in other closely related languages this last consonant appears to be an 'epenthetic' 'guttural' consonant (most possibly a variant of our epenthetic glottal stop) (Malou 1988). The argument, therefore, would be that Dinka unlike Thai lost both the segmental content of its word-final ON-pair onset *and* its skeletal point whereas Turkish retained this skeletal point. The effect, as shown, at length in the above sections is that word-finally a p-licensed nucleus has nothing to license and thus doesn't trigger ECP into repair while in Turkish, Thai, Mandarin etc... show ECP repair strategies summarised in the above section.

(s11, s12, s14) all share one thing in common, a *pointed* onset followed by an empty nucleus. Therefore, I would propose that there is a universal constraint on empty nuclei licensing pointed onsets. This constraint when closely analysed can be shown to be nothing but the ECP although we understand, contra to strict-CV, that there are two onset types pointed and non-pointed, one of which, the pointed onset, is a ‘category’ which thus must satisfy ECP while the pointless onset is not subject to ECP (cf. Gussmann and Kaye 1993; Denwood 2006; Úlfsson 2006).

3. Part Two

Part one showed us that the epenthetic glottal stop is nothing more than a symptom of licensing within a particular structural configuration. Part two will go on to show us cases where this pointed onset can be motivated diachronically and shows two language specific consequences of altering the licensing relationship of the segmentally-empty, pointed-ON pair.

3.1. Lhasa Tibetan

All of the Tibetan ‘dialects’ are derived from Classical Tibetan of the Bodic branch of the Sino-Tibetan Language family (Thurgood and LaPolla 2003). Classical Tibetan was written extensively and thus we are able to see diachronic trends which lead to each of the Tibetan ‘dialects’ today. Ladakh is considered to be a ‘conservative’ ‘dialect’ in which all the word-final consonants have been retained while Lhasa on the other hand has undergone extensive diachronic change by which most word-final consonants have been lost (DeLancey 2003).

The cautionary note, regarding this data, extends as far as to say that Lhasa Tibetan is a highly heterogeneous with some speakers retaining word-final velar nasals and some losing them (DeLancey 2003), some speakers retaining pre-nasalisation of consonants while others not (ibid.) and a high increase in tonal contrast (Hari 1979).

In this study therefore I will only concern myself with the more general linguistic variety of Lhasa Tibetan, which is sometimes acknowledged as a lingua franca and called ‘Modern Standard Tibetan’, however, I will use the term Lhasa Tibetan through-out. Importantly the syllabic development of Lhasa Tibetan is that which lost syllabic positions of Classical Tibetan and is reduced to a maximal word of CVX where X stands for either a consonant or a long vowel (Geziben 1995). This linguistic variety will be shown to pattern identically with the language data presented at the end of section one.

In the following table there is demonstrated the way in which diachronic processes have left many Lhasa Tibetan words (DeLancey 2003 for the diachrony and Denwood P. 1999 for data).

Classical Tibetan	Lhasa Tibetan	Examples (Denwood P.1999)	Phonological Structure
Unchanged syllables			
CV	ditto	[di] ‘this’ / [tɕU] ‘water’ [mI] ‘people’ / [tɕa] ‘tea’	$\begin{bmatrix} \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{C} & \text{V} & & \end{bmatrix}$
CV[p/m/k/N]	ditto	[jag] ‘good’ / [maN] ‘many’ [ɕim] ‘good’ / [tEp] ‘book’	$\begin{bmatrix} \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{C} & \text{V} & \text{C} & \end{bmatrix}$
Changed syllables			
CV(+nasal)	C $\tilde{\text{V}}$:	[jĩ:] ‘be’ / [tɕĩ:] ‘went’ [kjÊ:] ‘know’ / [rÊ:] ‘time’	$\begin{bmatrix} \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{C} & \text{V} & [\text{L}] & \end{bmatrix}^{12}$
CV(+liquid)	CV:	[ɕO:] ‘lose’ / [tsa:] ‘finish’ [Ny:] ‘money’ / [Na:] ‘strong’	$\begin{bmatrix} \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{C} & \text{V} & \nearrow & \end{bmatrix}$
CV(+cor+obs) ¹³¹⁴	CV? C $\tilde{\text{V}}$: ¹⁵	[dE?] ‘exist’ / [rE?] ‘to be’ [ta?] ‘tiger’ / [p{?}] ‘tibet’	$\begin{bmatrix} \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{C} & \text{V} & & \end{bmatrix}$

What is of particular interest for this paper is the very last column in which we see that just like Thai and Turkish (seen previously) the words that used to have O2 housing a consonant have now lost the segment of that consonant but retained the skeletal point producing, just as in Thai, an utterance-final glottal stop:

‘stag’ /ta/ [ta?] (DeLancey 2003)

(s15)

$$\begin{bmatrix} \sqrt{\text{x}} \\ \text{O} & \text{N} & \text{O} & \text{N} \\ \text{x} & \text{x} & \text{x} & \underline{\text{x}} \\ \text{t} & \text{a} & \text{?} & \\ & & \text{?} & \\ & & \text{?} & \end{bmatrix}$$

¹² The [L] in O2 makes the vowel sound long (2 skeletal points) and nasalized, when this [L] is followed by a plosive this nasal vowel sounds like an oral vowel followed by a nasal consonant which is homorganic with its following consonant. [jĩ:] --> [jin-da] ‘be-emph’, [jin-gi-...] ‘be-link.’, [jim-b@...] ‘be-link.’ Etc.

¹³ Coronal and variably velars

¹⁴ Some Tibetologists don’t distinguish between [CV?] and [C $\tilde{\text{V}}$]. In this paper it is the object of study I will use DeLancey’s guide to infer the [CV?] from otherwise un-transcribed data.

¹⁵ DeLancey says [CV?] is sometimes transcribed as [C $\tilde{\text{V}}$] (where [$\tilde{\text{V}}$] is a low tone and [V] : high tone).

As we will see in section three this glottal stop alternates with zero in non-utterance final environments.

3.2 Maremmano Italian

Maremmano Italian is a dialect of Italian which shares features with both Northern and Southern dialects while showing aerial similarities with the other Tuscan dialects (Úlfsbjörninn 2006b). There are three phonological processes which characterise Maremmano Italian. Firstly, most of the time, one of the two adjacent vowels, across morpheme boundaries, is subject to deletion:

Maremmano Italian (Úlfsbjörninn 2006b)

11 UR: [la anatra e le anatre sono andate a fa il bagno al mare]

11b SF: [lØ anatra e l:Ø anatre sonØ andatØ a f:a Øl bagnØ al mare]

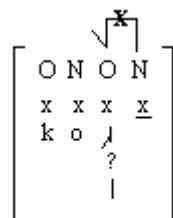
Maremmano Italian is also known for the clipping of its infinitive verbs and prepositions¹⁶. Clipping is the loss of parts of words in historical change (Trask 1997). Importantly, in Maremmano, we see evidence that the ‘lost bits of words’ is exclusively segmental while its syllabic structure is retained. Word-finally these prepositions and infinite verbs evidence a glottal stop where the archaic segmental material would be. Just as with the Lhasa Tibetan this glottal stop alternates with zero once it is taken out of a word-final environment.

Table of Clipped Prepositions (Úlfsbjörninn 2006b):

Preposition	Latin	Italian	Maremmano
in	in	in	i(n)
with	cum	kOn	kO?
on ¹⁷	sub	su	su?
for	per	per	pe?
there ¹⁸		la	la?
to	ad	a	a?

The Maremmano Italian prepositions are therefore seen to be captured by one ‘template’ in which the word-final ON pair is segmentally empty:

(s16)



¹⁶ Actually as we will see later many other words are clipped like ‘sono’ (to be) and ‘non’ (non) but none so systematically as the prepositions and infinitive verbs.

¹⁷ This actually comes from ‘under with motion’ [sub] it became ‘on’ [su]

¹⁸ They do come from demonstratives but in usage they are prepositions [la la gatta si lecca] (there the cat licks herself)

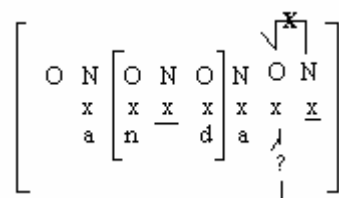
Table of Clipped Infinitive Verbs (Úlfsbjörninn 2006b)

Standard Italian	Maremmano	English
[voFO kantare]	[voFO kanta?]	‘I want to sing’
[voFO sapere]	[voFO sape?]	‘I want to know’
[voFO kapi:re]	[voFO kapi?]	‘I want to understand’

We see that the ‘-re’ at the end of all three infinitival suffixes ‘are/ere/ire’ is segmentally lost, but in its stead there is an utterance-final glottal stop, which we will see in section three alternating with zero. This would motivate the structure (s17) in which we can see the now familiar utterance-final segmentally-empty ON-pair:

Maremmano Italian ‘andare’ /anda/ [anda?] ‘to go’

(s17)



3.3. Cross-Linguistic Comparisons, sections one and two

We have seen many cases now of languages ending their words in word-final segmentally-empty structure comprised of an ON-pair with a pointed onset. We have also repeatedly seen that there seems to be a universal constraint on the ability of a p-licensed empty nucleus to license a segmentally-empty pointed onset and lastly we have seen how some languages choose to deal with this:

Repair Effects	Language
Spread the vowel from N1 to N2	Turkish, Mongolian, Beijing Mandarin
Phonetically realise the onset	Tamil, Cupeño, Thai, Vietnamese, Lhasa Tibetan, Maremmano Italian
?Phonetically realise the nucleus	unknown ¹⁹ ...

As discussed in part three the above can be best described as being a simple violation of the ECP although these languages all demonstrate this ECP effect in a particular environment word-initially for Vietnamese and Tamil and word-finally for all the others. Thus the first row in the above mini-typology is a ‘yes’ setting to our ‘Pointed ON-Pair Repair’ strategy while the second column is the natural effects of not repairing this particular ECP violation.

¹⁹ Not counting spreading as in Turkish etc..., it would be very hard to identify! There would be a lot of words ending in empty vowels in word-final or word-initial environments where they could otherwise be p-licensed. Ps. Also there would be no reason to assume they would be stressed (counts out Turkish).

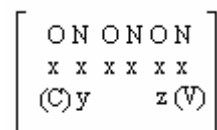
In this section, we are going to examine how this word-final empty structure is satisfied phonologically when it is removed from the word-final context. In particular, I will be exploring languages from the second column of the mini-typology table using Lhasa Tibetan and Maremmano Italian as case studies for they behave in opposite but equally logical ways.

We will see some alternations of glottal stop with zero when words which end in the now familiar word-final empty structure are placed before other elements thus losing their word-final status. In section three I plan to show how Maremmano Italian and Lhasa Tibetan behave exactly oppositely with the former inducing Raddoppiamento Sintattico (RS) and the latter Compensatory Lengthening (CL).

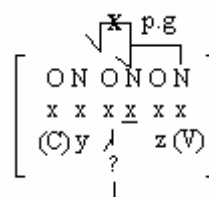
4. Part Three

When the segmentally-empty, pointed, ON-pair occurs *utterance-finally* we can be certain that the empty nucleus flanking the pointed onset is *domain-final* and thus in a position to be p-licensed by the word-final. Kaye's (1995) distinction between analytic and non-analytic morphology and work on Turkish vowel harmony shows that the end of the word is not necessarily the end of the domain and the word-final parameter is better understood to be the granting of p-licensing in domain not word finally. This, however, raises a revealing question. The theory states that the glottal stop is realised because its source of licensing is itself p-licensed. Utterance-finally, therefore, the final nucleus could be p-licensed and, as such, not a good licenser. In a sequence like [koON korina], however, if we assume that the brackets represent the domain and *word-final* ON has lost its *domain* final position and with it its licensing from the word-final parameter. In fact the ON-pair is at the heart of a phonological word. Maremmano Italian and Lhasa Tibetan react by spreading elements into the empty ON just as if it were a landing site for a geminate or a long vowel (s6 repeated). The question is then why does this happen and not (s6b):

(s6 repeated)



*(s6b)



There seem to be two possible and likely alternatives, one is rather conservative and the second is radical.

Firstly, we could assume that the *projection principle* is even more correct than we usually assume and proper government as a licensing force is established in the lexicon and remains constant through the derivation (Kaye et al. 1990). Therefore, seen as when you precede a word with an utterance-final ON pair this is post-lexical and proper government could be argued to be only assigned in the lexicon. The phonological process of RS and CL, however, could be argued to be only relevant post-lexically. This in fact is a problem shared with Optimality Theory which uses the 'one-step derivation' axiom cross-morphemic CL is a problem and has led to either

abandoning the one-step approach or adorning the candidates (post-input, pre-output) (Goldrick 2000).

Secondly, we could assume that proper government doesn't exist. The flavour of sCV in which this essay works is based largely on Cyran (2003) for whom proper government indeed doesn't exist. As I showed in part one, the empty nuclei between consonant clusters are licensed by government licensing and similarly when we look at RS and CL we assume in both cases that the intervening empty nuclei and onset will be p-licensed by being enclosed inside a licensing domain (cf. Cyran 2003). The licensing of the pointed onset by the realised word-final nucleus (as seen in Turkish and Beijing Mandarin) is never explicitly labelled proper government as the original definition of proper government is applicable only between binary terminals in the nuclear projection (Charette 1991). Although this pointed onset licensing is satisfying the ECP in just the same way as proper government does and has the same restrictions, we must remember that two identical effects don't necessarily make for an identical aetiology. In as much as this study is concerned, it hasn't had to use proper government. What we have seen is a licensing of an empty category, the pointed onset. Although this could be proper-government also having the same restriction on it, in terms of p-licensed licensors not being able to license the pointed onset, it is a possibility that this a) isn't proper-government and thus proper government wouldn't be needed at all in phonology or b) this is exactly what proper-government does thus keeping proper government but demoting it to a far smaller role in phonological theory.

It is beyond the scope of this study to ask whether proper government exists, however, either way, we can safely assume that proper government cannot be applied post-lexically (ie. that could be due its non-existence, but that is a separate issue).

What is relevant for this study is that part two has shown the effect of ON when utterance-finally. In certain languages such as Maremmano Italian and Lhasa Tibetan this word-final ON-pair can be followed by words in which it seems that the ON-pair isn't any longer domain-final. That is, elements from left and right can spread into the ON-pair in just the same way as geminates and long vowels are understood by sCV to be formed internal to the lexicon. Essentially, [koON korina] looks identical to (s6 repeated):

$$(s6 \text{ repeated})$$

$$\begin{bmatrix} \text{O N O N O N} \\ \text{x x x x x x} \\ \text{(C)y \quad \quad z(V)} \end{bmatrix}$$

Part three will now show that the two theoretical possibilities regarding this situation are attested. That is, spread y or spread z. This, due to its cross-morphemic nature, is called Raddoppiamento Sintattico and Compensatory Lengthening.

4.1. Raddoppiamento Sintattico vs. Compensatory Lengthening

Although it isn't clear why a language would opt for RS over CL, languages seem to choose one strategy over the other. In Maremmano Italian any clipped infinitive or

preposition followed by a consonant initial word will result in the spreading of the consonant of the following word into its empty structure forming a geminate.

4.1.1. Maremmano Italian RS (Úlfsson 2006b)

Phonological occurrence: $[V?] \rightarrow [V? + C] \rightarrow [VC_x + C_x]$ ²⁰

Data 13-15

RS Triggers ie. words ending in our ON pairs: [pe?] (for) and [so?] (am)

Frame of the sentence: [i fjOri soCV peCV x/y/z]
the flowers are for x/y/z

- 13 [i fjOri so **p:e** **m:arija**] ‘mary’
14 [i fjOri so **p:e** **l:ejj**] ‘her’
15 [i fjOri so **p:e** **l:a Dhon:a**] ‘the lady’

Data 17-19

RS Trigger: [kO?] (with)

- 17 [i fjOri so **k:O** **m:arija**]
18 [i fjOri so **k:O** **l:Ejj**]
19 [i fjOri so **k:O** **l:a Dhon:a**]

Data 20-22

RS Trigger: [no?] (not/ no)

Frame of sentence: [noCV lO soCV x/y/z]
neg it know x/y/z

- 20 [no **l:o** so **p:e** ki s:o?] ‘for who they are’
21 [no **l:o** so **k:i** s:o?] ‘who I am’
22 [no **l:o** so **d:OvEl** kane] ‘where the dog is’

Data 23-25

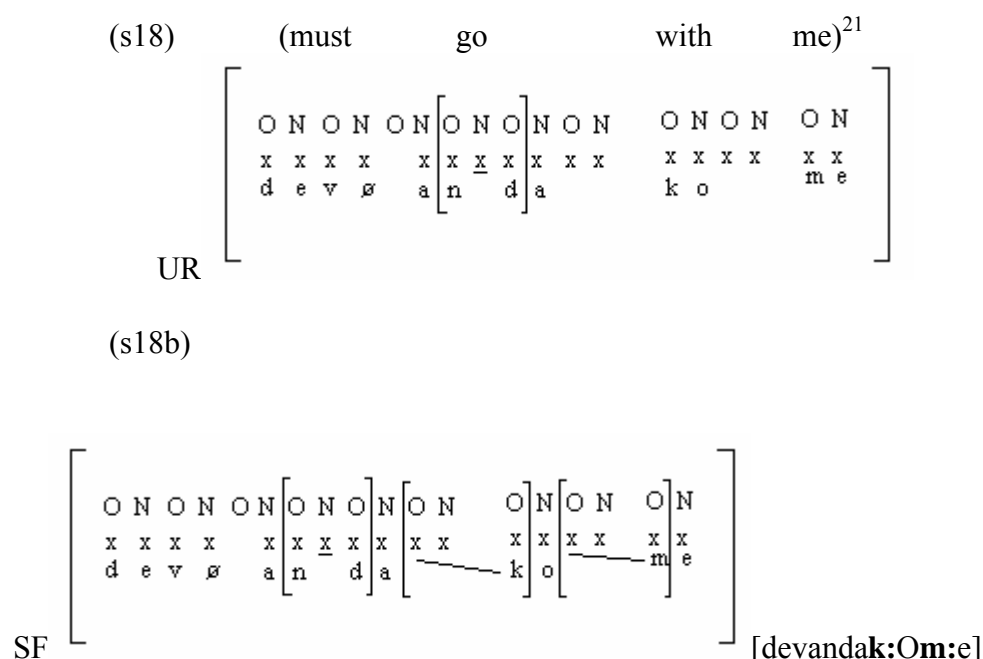
RS Trigger: [kapi?] (understand)

Frame of sentence: [devO capiCV x/y/z]
must understand x/y/z

- 23 [devO kapi **l:a** sitwatzjOne] ‘the situation’
24 [devO kapi **k:ij** era] ‘who it was’
25 [devO kapi **t:i**] ‘you’

²⁰ Where brackets are phonological domain

Essentially the situation above can be shown schematically as follows:



What the gemination does is create an onset-onset relationship between the segmentally-empty onset at the end of the word and the following onset of the next word. This creates a closed [CVC] domain in which the interceding V position would be automatically p-licensed (cf. Cyran 2003). The governed onset is then filled with elements from the governor onset, although this is language specific and in other languages we could predict that only partial spreading will occur (cf. Harris 1997). RS is a good strategy to deal with word-final empty ON-pairs as it simultaneously deals with both empty O and N. The glottal stop that was in the governee onset goes to zero as it never had segmental content and it is then filled with elements from its governor.

4.1.2. Lhasa Tibetan CL

Considering the words of Lhasa Tibetan which end in glottal stop utterance-finally, when we place these words before another item within a phonological domain we see them behaving in exactly the opposite way to Maremmiano Italian. Instead of the following consonant closing the domain thus inducing RS we have the root's own vocalic material spreading into N2.

Phonological occurrence: [V?] → [V? + C] → [V:+C]

CL Triggers: [phö?] (Tibet) and [re?] (Be)

(DeLancey 2003)

kho bod-pa red-pas

²¹ Within (s18) we can see (s16) 'with' and (s17) 'to go' both of which utterance finally end in a glottal stop.

26 [k_ho phö:-pa re:-pä:]

(Is he Tibetan?)

(Denwood P.1999)

khjed-rang bod-la gro...
26b [k_hje-rã: pö:-l@ drO]

(Have you ever been to Tibet?)

(Denwood P. 1999)

khong bod-nas red-pas
26c [k_hoN pö:-ne: ri-be:]

(Is he from Tibet?)

CL Trigger: [k_hye?] (You)

(DeLancey 2003)

khyed-rang chams-pa brgyab dug-gas
27 [k_hye:-raN cham-pa kyap tu-kä:]

(do you have a cold?)

CL Trigger: [ca?] (break) and [re?] again

(DeLancey 2003)

kho-s nga-s bcag-pa-red zer-gyis
28 [k_hö-? Nã-? ca:-pa-re:-s se:-ki]

(he says I broke it)

CL Trigger: [la?-] (work) and [la?-] (office)

(Denwood P. 1999)

khos las-ki bjas chog-gi-red
29 [k_hö le:-g@ tçe tçho:-gy-re:]

(He is/will be allowed to work)

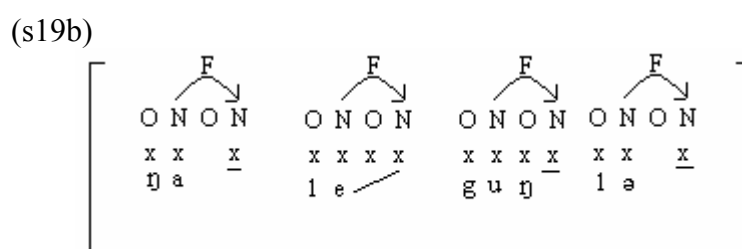
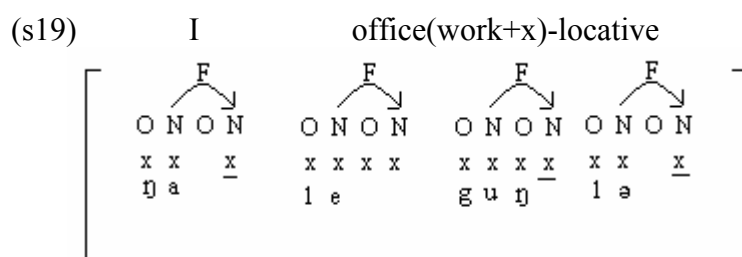
(Denwood P. 1999)

nga las-kung-la phyn
29b [Na le:-gUN-l@ tçhĩ...]

(I had gone to the office)

Where orthographically a coronal plosive was written the segmental content of that onset has been lost to history, the syllabic structure. However, and just like in h-aspiré words the point of the onset has not (Charette 1991). And what we see on the surface is that this empty pointed onset which used to house a glottal stop utterance-finally and in isolation mostly becomes a long vowel before other elements within its phonological domain.

Schematically the situation above can be crystallised schematically as follows:



In (s19) we can see that ‘las’ in (s19) cannot have its word-final nucleus p-licensed by the word-final parameter. As such the repair strategy for this ECP violation, being in Lhasa Tibetan, is to transport the contents from N1 to N2.

Another view, considering Charette’s word-final foot (as illustrated) is that a word-final trochaic foot p-licensing the N2 in [le?] would create the problem of it not being able to license O2 to be empty, again seen as an ECP violation. This scenario triggers in Lhasa Tibetan the effect by which the vowel in N1 spreads into its N2 thus abling this nucleus to license the segmentally-empty onset point.

There doesn’t seem to be any intrinsic reason as to why in both Maremmano and Lhasa Tibetan the glottal stop should only surface utterance finally²². At least in Lhasa Tibetan N2 can be p-licensed and the glottal could satisfy the ECP producing lots of sentences like: [Na le?-guN-l@]. It seems however that the phonetic realisation of the glottal is a ‘last case scenario’ and RS and CL are both favourable. The reason for this is a mystery although it is an observable trend. One possible reason comes from parsing, that is, utterance-finally a particular structural configuration behaves differently to domain internally as an identificational cue of the phonological domain itself (cf. Kaye 1989; Ploch 2004).

This means that our ‘constraint’ from section two has two flavours and an extra clause:

Pointed ON-Pair Repair (revised)

Domain Finally

²² And if you know your Tibetan the glottal is also an epenthetic consonant utterance initially where it coincides with a high-tone vowel beginning the utterance, this however goes to zero when preceded by another element: [dŪV:] + [ʔu:] ---> [dŪV:-u:] ‘breathe’ + honorific

- a) spread the content of N1 to N2 [yes, no]

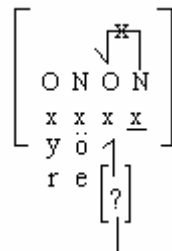
Domain Medially

- a) spread the content of N1 to N2 [yes, no]

Importantly, the above repair strategy uses the same parameter exactly although because this parameters has two different contexts it can be switched to (no) domain-finally but (yes) domain-medially. This would create a language with domain-final glottal stops but domain-medial compensatory lengthening. This setting would be Lhasa Tibetan, while Maremmiano Italian would have both contexts set to (no) and as such utterance-final glottal stops and domain-medial raddoppiamento sintattico.

To show the dynamism and productivity of this principle we can use an illustrative example from Lhasa Tibetan where we can observe the build up of various CL triggers and their behaviour as they loose domain-final status. The affirmative, egophoric, existential mood marker ‘yod’ [yö?] can combine with a generic auxiliary ‘red’ [re?]. Following our earlier discussion we would assume the following structures:

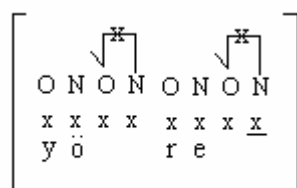
s19c



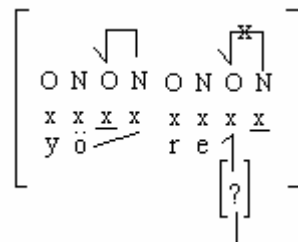
Lhasa Tibetan being what it is sets the word-final context of the ‘Pointed ON-Pair Repair’ to (no). Therefore the p-licensing of the domain-final nucleus is upheld and the pointed onset violating the ECP is filled with phonetic content.

However, when we concatenate the stems we would end up with the following structure and we have two segmentally-empty ON-pairs. These will be dealt with by the ‘Pointed ON-Pair Repair’ although in (s19d) there are two contexts to differentiate a domain-medial and a domain-final context. Lhasa Tibetan therefore maintains that word-finally the parameter is set to (no) although domain-medially the parameter is set to (yes) and as such we get the domain-medial ON-pair have the content of its preceding vowel spread into it. The predicted result [yö:re?] is exactly what is seen in the data (Tournadre and Dorje 2003).

s19d



s19e



Furthermore, if one was to take (s19e) and follow it with a question particle which is also a CL trigger [re?] would lose its domain-final context and be ‘demoted’ to domain-medial context with relevance to the application of the ‘Pointed ON Pair Repair’ the domain-final CL trigger would however, be bound by the domain-final parameter setting of Lhasa Tibetan and so one would predict: [yö:re:pe?] *[yö:re?pe?], this is in fact borne out in the data (Tournadre and Dorje 2003; Miss Ceri May p.c). In this last section I will attempt to the highly related topic of Lowenstamm’s (1996) and (1999) analysis of RS and CL in Biblical Hebrew and show that he was essentially right but that in order to produce the correct data, the analysis must be drawn from a combination of both his hypothesis while remembering the notion that diachronic processes can empty an ON-pair of their segmental content but not necessarily remove either themselves or intermediary skeletal points.

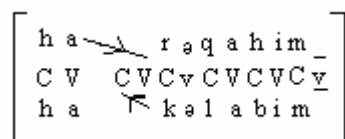
5. Part Four

In 1999 Lowenstamm argues that before every major category in every single language there lies an empty initial CV-site. This is the strong form of the hypothesis and there have been a number of wants/attempts to weaken it. Kula (2006) and Charette (2006) have respectively shown the initial CV may not be truly universal or exclusively initial. Kula argues that ‘clusterless’ languages do not have an initial CV site while Charette shows that in Turkish and in languages like it (essentially the ‘CVC’ languages) the empty CV site is word-final. In this study, I have shown that in Maremmiano Italian and Lhasa Tibetan that word-final, segmentally-empty structure, visible diachronically, can explain all the RS and CL attested in the languages.

There is a discrepancy between Lowenstamm (1996) and (1999). In his (1999) analysis the empty structure to explain RS and CL in Biblical Hebrew is positioned at the beginning of the major categories ie. [cv-klabim]. In (1996) however, the empty structure which was inducing RS and CL was situated at the end of the triggering word such as [ha-cv] ‘the’.

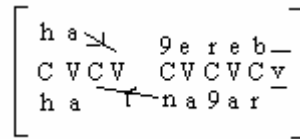
Lowenstamm (1999)

(s20)



Lowenstamm (1996)

(s20b)



5.1. Problems with Lowenstamm’s Analysis of RS and CL

The purpose of the (1999) paper is to explain word-initial clusters. The words in (s20) word initially are [klabim] and [rqaʔim] respectively. In the former the /kl/ are in a closed domain and in the latter /rq/ is not. However, when you add the determiner /ha/ both clusters become broken up by a schwa. The reason for this schwa-zero alternation is never explained in Lowenstamm (1999) however, we could infer that Lowenstamm might argue that the schwa turns up in /ha rqaʔim/ in order to license the CV-site to allow the spreading of the vowel into it, after all, in (1996) he did show that in Semitic a vowel can only spread if the V into which it spreads is licensed *[ka:tpi] vs. [ka:tupi].

If the motivation for spreading the vowel into the empty structure is because it is unlicensed ergo it would require phonetic content, it is illogical to then license the very same structure, satisfying the ECP, in order to allow for another, equally valid, mechanism to satisfy the ECP (ie. phonetic content). In other words, when you insert the schwa between /rɪq/ you automatically license the word-initial CV-site [CV-r@q] thus as the initial CV-site is licensed there is no ECP violation to satisfy (ie. by spreading phonetic content into it).

According to Lowenstamm's own framework rules and restrictions (s21) should output the illicit: (s21a) *[ha k:labim], and (s21b) *[ha:rqaʔim] or (s21c) *[ha r@qaʔim].

Underlying Form

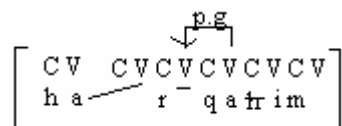
*Surface Form

14/04/2007

The long vowel - without the epenthetic schwa

*Surface Form

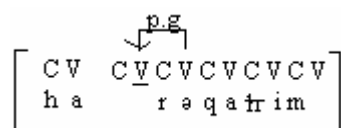
(s21b) *



The epenthetic schwa - without the long vowel

*Surface Form

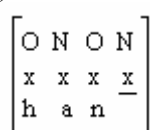
(s21c) *



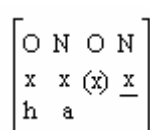
5.2. The Solution to Lowenstamm's (1996; 1999) Problem

The clue to the answer of these problems is found still within Lowenstamm's writings. In (1996) he posited (s20b) in which we see [ha] with empty CV structure attached to it. Most likely, he knew that a good number of Semitists believe the Proto-Semitic form of /ha/ in Biblical Hebrew to be *[han] (Lipinski 2001, Hasselbach 2004). Just as with the Maremmano Italian and Lhasa Tibetan, therefore, we have word-final empty structure which is motivated diachronically:

Proto-Semitic
(s22)



Biblical Hebrew
(s22b)



If though, we posit that the empty CV structure *is* on the end of the determiner and this is solely responsible for RS and CL then we would lose the convincing claim that an initial empty CV at the beginning of the major categories is responsible for the typology of word-initial clusters. The good news is that in the Biblical Hebrew data provided in Lowenstamm (1996), (1999) a word-final empty structure attached to the determiner isn't enough (on its own) to account for the target output. This is for exactly the same reasons as the word-initial empty structure cannot generate the correct output. After all the word-initial empty CV is "midst of a phonological word" Lowenstamm (1999:164) and thus it doesn't matter where the CV was first attached the result will be identical. Alone these two empty structures, both reasonably well motivated, cannot account for the data, whereas a combination of these two empty structures can explain the Biblical Hebrew facts. The diachronically visible word final

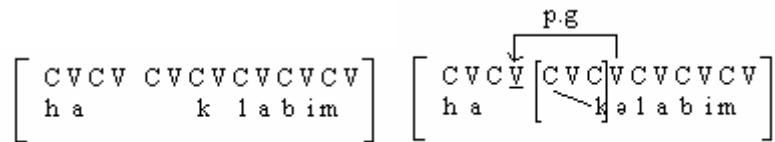
structure of the determiner /haCV/ and the well argued word-initial empty structure combined produce the correct output.
 /haCV CVkʌlabim/ --> [hak:@labim]

Underlying Form

Surface Form

(s23)

(s23b)



In (s23) we see the empty structures and in (s23b) we see that the empty structure is dealt with in the following way. Firstly, just as in Maremmiano Italian, the empty ON-pair attached originally to [klabim] is spread into by the onset forming a geminate and being within a closed domain the intervening V is buried (explained in section one). There is yet another ON-pair to satisfy the ECP of however. The geminate having formed the closed domain can now be jumped by proper government which licenses the CV site to be empty (cf. Scheer 2004). This however, forces the V between the [k] and [l] ‘cluster’ to surface with a schwa (in harmony with sGP assumptions Charette 1991).

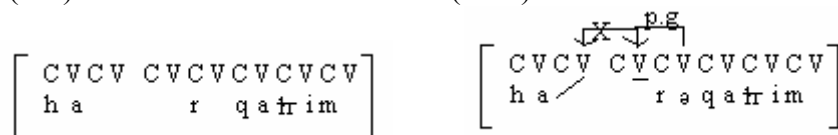
/haCV CVrʌqaʔim/ --> [ha: r@qaʔim]

Underlying Form

Surface Form

(s24)

(s24b)



In (s24) we see the same problems as in (s23) only this time RS isn’t possible as gutturals and [r] cannot geminate in Biblical Hebrew. Another strategy must therefore be adopted. The V between the [r] and the [q] surfaces with schwa and licenses the ‘word-initial’ empty CV. The ‘word-final’ structure at the end of the determiner /ha/, however, is still unable to receive a source of p-licensing and thus the [a] in V1 spreads into V2 producing the lengthening effect known as CL. This structure outputs the correct form: [ha:r@qaʔim].

5.3. The ‘Pointed ON-Pair Repair’ Parameter Refined

One could ask how the ‘Pointed ON-Pair Repair’ could handle the Biblical Hebrew data. I would argue that it does, in fact, explain the Hebrew facts but simply that an extra (and last) clause must be added to our understanding of the domain-medial repair strategy. From section three one could conclude that domain-medial repair was the identical parameter to domain-final repair. Although this may be the case, domain-medial repair comes in two flavours, that is, it contains two parameters (not one) unlike the domain-final repair. This second parameter will have to be x) spread the content of Ox into Oy. That is, spread the onset of the following word into the

preceding empty structure. So what nature of parameter are we talking about? From Biblical Hebrew, a language where both strategies apply domain-medially, we know that the germination is primary and the consonants that phonotactically cannot form geminates are accommodated by spreading N1 to N2. Hence, we can tentatively assume that RS comes first and CL second. So the new domain-medial repair parameter listing look like this:

Domain Medially

- a) spread the content of Ox into Oy [yes, no]
- b) spread the content of N1 to N2 [yes, no]

Therefore the entire ‘Pointed ON-Pair Repair’ looks like the following:

Pointed ON-Pair Repair (final revision)

Domain Finally

- a) spread the content of N1 to N2 [yes, no]

Domain Medially

- a) spread the content of Ox into Oy [yes, no]
- b) spread the content of N1 to N2 [yes, no]

5.4. Why are Biblical Hebrew and Maremmano Italian Different?

The treatment of Biblical Hebrew therefore begs a question; Maremmano Italian is a prime candidate to be classified as an initial-CV language²³ as it has both strong government and initial TR clusters (Kula 2006). However, Maremmano Italian and Biblical Hebrew do not behave similarly in their RS and CL. If Maremmano Italian was just like Biblical Hebrew we would expect: ‘per cremare’ /peCV CVk_vremaCV/ to be [pek:erema?]²⁴ with the cluster [kr] broken up just as in Biblical Hebrew. This isn’t the output however and thus the underlying forms of Biblical Hebrew and Maremmano Italian cannot be the same.

In answer to this query which could be the scope of future research will most probably be aided by the fact that Maremmano might not actually have strong government, TR clusters, or in sGP ‘branching onsets’. If this was true, in Kula’s analysis, one would show Maremmano as not having an initial CV-site at all.

5.4.1. Maremmano Italian with no Branching Onsets

In Standard Government Phonology clusters like [tr], [kr], [gl], [pr] were all seen as possible branching onsets as they obeyed constituent government by being a) strictly local and b) strictly directional (Kaye et al. 1985). The qualification for being a good governor was elemental complexity (Kaye et al. 1985; Harris 1990) the governor having the more complex structure. There was, however, always the possibility that

²³ Lhasa Tibetan looks like, and would be expected to be a word-final foot language (in line with Charette 2006).

²⁴ [e] is the Maremmano epenthetic vowel found in loans to satisfy the word-final parameter /gol/ [gol:e] /pik nik/ [pik:e nik:e] or [piknik:e].

what looked like a branching onset on the surface was actually not a branching onset at all, [aple] vs. [apel] (Charette 1991), these are widely known as bogus clusters in sGP.

Kula (2006) shows that it is quite possible that clusterless languages actually don't have an initial CV-site. A comparison between Biblical Hebrew and Maremamno Italian²⁵ hints that in the latter there may well not be a word-initial site after all. This would (retrogressively) beg the question, are Lowenstamm and Kula both right? Are there consonant cluster-languages without the CV-site? or the other side of the coin; is Maremmanno Italian a clusterless language?

5.4.1.1. Evidence from Maremmanno Cluster Lenition

Evidence from lenition points to the fact that Maremmanno Italian 'clusters' may well just be bogus clusters. In Strict-CV or my 'Strict-CV with the skeleton' a real cluster will be within a closed domain and a bogus cluster will not be (essentially from Cyran 2003). This motivates a prediction by which one would expect the governor not to lenite into a consonant more simplex than its governee. If this were to happen it would violate the maxim of complexity as a condition of government (Kaye et al. 1985; Harris 1990).

In the following table is some data showing the lenition patterns of 'branching onsets' that in fact in the most obvious case ie. the 'governor' of least complexity and the 'governee' of most complexity in Standard Italian and Maremmanno Italian: [g] and [l].

Table 3

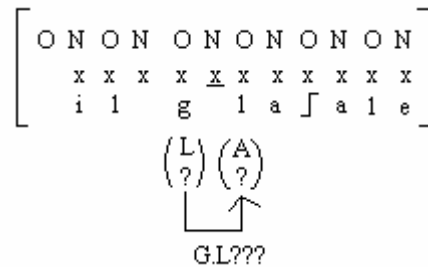
Word-Initial Cluster	Post-Vocal Cluster	Element make up of both 'governors'
[prova-mi] 'try me!'	[la P _h ro:va] 'the test (try)'	(U.H.?) vs. (U.H)
[prE:ti] 'priests'	[i P _h rEti] 'the priests'	ditto
[prEndi-mi] 'catch-me'	[la P _h re:za] 'the catch'	ditto
[pretese] 'pretence'	[la P _h rete:za] 'thepretense'	ditto
[troThē] 'trouts'	[le Th _h roThē] 'the trouts'	(A.H.?) vs. (A.H)
[trip:a] 'tripe'	[la Th _h rip:a] 'the tripe'	ditto
[trE◊wa] 'relief'	[la Th _h rE◊wa] 'the relief'	ditto
[krETHa] 'clay'	[la xrETHa] 'the clay'	(H.?) vs. (H)
[krEPHa] 'die'	[mi xrEPHa] 'me-dies'	ditto
[il kranio] 'the skull'	[il mi xranio] 'my skull'	ditto
[krema] 'cream'	[la xrEma] 'the cream'	ditto
[il glaSa:le] 'the glacier'	[l-Era ◊laSa:le] 'ice age'	(H.L) vs. (L)
[gran:a] 'money'	[la ◊rana] 'wage (the m)'	ditto
[grap:a] 'foul drink'	[la ◊rap:a] 'the grappa'	ditto
[il grande] 'the.m big'	[la ◊rande] 'the.f big'	ditto

²⁵ based on the outputs of the RS of their 'closed domain' 'branching onsets'

We can see below that the [g] does in fact lenite before [l] in a ‘branching onset’ and as such supports the notion that [g] could never have been governing [l] in the first place. This wouldn’t be true of other languages, but in Maremmano Italian lenition says that its word-initial clusters of rising sonority are in fact bogus clusters ie.(s25-26).

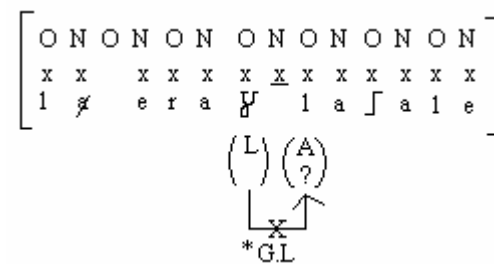
‘il glaciale’ /il glaSa:le/ [il glaSa:le]

(s25)



‘l’era glaciale’ /la era glaSa:le/ [lEra ɔlaSa:le]

(s25b)



The effects this has on this study is that if Maremmano Italian doesn’t have true TR clusters then if we agree with Kula (2006) we know why Maremmano Italian and Biblical Hebrew²⁶ are intrinsically different. The former has only one ON-pair in the midst of the phonological word and as domain-medial the language’s parameter will spread the consonant forming a cross-morphemic geminate and there would be no further ECP violations to handle. Biblical Hebrew on the other hand would have two ON-pairs to satisfy and so in accordance with parameter (a) of the ‘Pointed ON-Pair Repair’ will spread its consonant (phonotactics allowing) licensing one ON pair leaving another ON-pair to be satisfied by proper government. This would have the effect of breaking up the potential cluster as seen in (s23b).

6. Conclusion

Part one outlined the theory of epenthetic glottal stops and motivate the justification of the constraint: a p-licensed empty nucleus cannot license an empty pointed onset. This constraint was then shown to be nothing more than a straightforward ECP violation which due to its specific location within the domain (ie. final) was influenced by other factors, such as the word-final parameter. Part one went on to

²⁶ Biblical Hebrew apparently is a TR language. Despite what one could predict from its clusters.

show that different languages behave in one of two different ways in repairing this ECP violation domain-finally and as such these strategies could be parametised into what I called the ‘Pointed ON-Pair Repair’.

Part two then showed us diachronic evidence from Lhasa Tibetan and Maremmano Italian as regards word-final segmentally-empty syllabic structure and the repair parameter detailed in part one. These languages historically showed us the loss of segmental material and the retention of its syllabic structure.

Part three then showed us how the utterance-final empty structure interacts with elements placed after it. Theoretically speaking this has the effect of taking our ECP violation away from domain-final to the domain-medial position. With our the n understanding of the ‘Pointed ON-Pair Repair’ parameter what we would have expected is that Maremmano Italian and Lhasa Tibetan would behave identically. This hypothesis wasn’t borne out by the data: for Maremmano Italian the ECP violation repair strategy was Raddoppiamento Sintattico, while in Lhasa Tibetan it was Compensatory Lengthening. The data showed us that to properly parametise the repair strategies we must have a differentiation for domain-final and domain-medial. These two locations for the same parameter are independent of each other and so if Maremmano Italian and Lhasa Tibetan select for the same setting domain-finally they can (as we see) select a different setting domain-medially.

Having defined the parameters responsible for the repair strategies which superficially we call RS and CL there is an obvious next step. Lowenstamm (1999) claims that RS and CL in Biblical Hebrew are motivated not by word-final diachronically visible segmental loss and syllabic retention rather by an empty ON-pair (for him CV-site) word-initially. This he argues is universal. What we saw in part four, however, was that Lowenstamm’s rules for generating RS and CL with only one CV-site could not output the correct data as found still within Lowenstamm (1999). The justifications for his structure and outputs was shown to be falsified *a priori*. Part four then showed that actually the CV-site **is** essential to Biblical Hebrew RS and CL but that one CV-site is not enough. Through diachronic segmental loss and syllabic retention however, we can see that the RS and CL trigger ‘ha’ was historically *han. With this essay’s word-final empty structure arguments and Lowenstamm’s (1999) initial CV-site it was then possible to output the correct data (as found within Lowenstamm 1999).

The initial-CV site was evidently important to understand Biblical Hebrew and supposedly universal, at least within languages with real consonant clusters (Kula 2006). So Lhasa Tibetan was immune to the following question but why doesn’t Maremmano Italian, supposedly a branching onset language, behave like Biblical Hebrew with diachronically motivated word-final empty structure and typologically motivated word-initial empty structure. If Kula (2006) is right we could understand Maremmano Italian as not having true branching onsets and thus not requiring a word-initial site, leaving us with the empty word-final empty structure and the correct outputs. Part four concluded by showing lenition data that in deed supported this hypothesis. Lenition in branching onsets in Maremmano Italian showed the supposed governor reducing to a lower elemental complexity than its supposed governee. This would violating the complexity condition of government which is a vital notion in Standard Government Phonology to account for branching onset typology (Kaye et al. 1985; Charette 1990, Harris 1990).

7.

Bibliography

- Bachmaier, R., Kaye, J. and M. Pöchtrager, (2006), “*Stop Glottals! Phonology without ?*”, ms., paper presented at *Conference in Government Phonology 2006*, Universiteit Leiden.
- Charette, M. (1988), “*Some Constraints on Governing Relations in Phonology*”, PhD. Thesis, McGill University.
- Charette, M. (1990), “*License to Govern*”, in *Phonology*, 7:233-253.
- Charette, M. (1991), “*Conditions on Phonological Government*”, Cambridge University Press.
- Charette and Göksel, (1998), “*Licensing Constraints and Vowel Harmony in Turkic Languages*” in Cyran (ed.), *Structure and Interpretation. Studies in Phonology*, 65-88, Folium.
- Charette, M. (2006), “*At the End of the (Turkish) Word*” in *SOAS Working Papers in Linguistics*, 14:23-40.
- Charette, M. (in press), “*The Vital Role of the Trochaic Foot for Understanding Turkish Word Endings*”, *Lingua*.
- Christdas, P. (1988), “*The Phonology and Morphology of Tamil*” PhD Dissertation, Cornell University.
- Christiansen and Christiansen, (1992), “*Kisar Phonology*” in *Phonological Studies in Four Languages of Maluku*, SIL, University of Texas at Arlington and Pattimura University, 35-65.
- Cohn, A. and J. J. McCarthy, (1994), “*Alignment and Parallelism in Indonesian Phonology*”, ms. Cornell University and University of Massachusetts and Amherst.
- Crowhurst, M. (1994), “*Foot Extrametricality and Template Mapping in Cupeño*”, *NLLT*, 12:177-201.
- Cyran, E. (2003), “*Complexity Scales and Licensing Strength in Phonology*”, Wydawnictwo KUL.
- DeLancey, S. (2003), “*Lhasa Tibetan*”, in Thurgood and LaPolla (eds) *The Sino-Tibetan Languages*, 270-285, Routledge.
- Denwood, M. A. (1999), “*A Government Phonology Approach to Thai Tones: Observation and Hypothesis*” in *SOAS Working Papers in Linguistics and Phonetics*, 9:173-190.

- Denwood, M. A. (2006), "A CVCV Template for Turkish", in *The Linguistic Review*, 23:465-504
- Denwood, P. (1999), "*Tibetan*", John Benjamins Publishing Company.
- England, N. (1983), "*A Grammar of Mam, a Mayan Language*", University of Texas Press.
- Geziben, D. S. (1995), "*The Onset Simplification Process in Tibetan*" in *University of Texas at Arlington Working Papers in Linguistics*, 2.
- Goh, Y. S. (1996), "*The Segmental Phonology of Beijing Mandarin*" SOAS PhD Dissertation.
- Goldrick, M. (2000), "*Turbid Output Representations and the Unity of Opacity*" in *Proceedings of the North East Linguistics Society*, 30:231-245
- Gussmann, E. and J. Kaye, (1993), "*Polish Notes from a Dubrovnik Café: the yers*", in *SOAS Working Papers in Linguistics and Phonetics*, 3:427-462
- Hari, A. M. (1979), "*An Investigation of the Tones of Lhasa Tibetan*", Summer Institute of Linguistics.
- Harris, J. (1990), "*Segmental Complexity and Phonological Government*", in *Phonology*, 7:255-300.
- Harris, J. (1994), "*English Sound Structure*", Blackwell.
- Harris, J. and G. Lindsay, (1995), "*The Elements of Phonological Representation*", in Durand and Katamba (eds.) *Frontiers in Phonology*, 34-79.
- Harris, J. (1997), "*Licensing Inheritance: an integrated theory of neutralisation*", in *Phonology*, 14:315-370.
- Hasselbach, R. (2004), "*The Markers of Person, Gender, and Number in the Prefixes of G-Preformative Conjugations in Semitic*" in the *Journal of the American Oriental Society*, 124:23-35.
- Kaye, J., Lowenstamm, J. and J. R. Vergnaud, (1985), "*The Internal Structure of Phonological Representations: A Theory of Charm and Government*" in *Phonology Yearbook*, 2:305-328.
- Kaye, J. (1989), "*Cognitive Phonology*", Lawrence Erlbaum Assoc. Inc.
- Kaye, J. (1990), "'Coda' Licensing", in *Phonology Yearbook*, 7:301-330.
- Kaye, J. (2000), "*A Users Guide to Government Phonology*", ms., University of Ulster.

- Kenstowicz, M. (1976), "Some Rules in Koryak Phonology", in *Studies in the Linguistics Sciences*, 6:22-37.
- Kula, N. (2006), "No Initial CV in Clusterless Languages", *Taalkunde in Nederland*, University of Utrecht, http://www.nancykula.net/TIN_HO-07-02-06.pdf.
- Lipinski, E. (2001), "Semitic Languages: Outline of a Comparative Grammar", 2nd edition, Leuven.
- Lombardi, L. (2002), "Coronal Epenthesis and Markedness" in *Phonology*, 19:219-251.
- Lowenstamm, J. (1996), "CV as the only Syllable Type" in Durand & Laks (eds.) *Current Trends in Phonology, Models and Methods*, 419-441, ESRI.
- Lowenstamm, J. (1999), "The Beginning of the Word" in *Phonologica*, Rennison and Kuhnhammer (eds.), 153-166, Holland Academic Graphics.
- Malou, J. (1988), "Dinka Phonology", University of Texas Press.
- McWhorter, J. (2001), "The Power of Babel", Harper Perennial.
- Mithun, M. and H. Basri, (1986), "The Phonology of Selayarese" in *Oceanic Linguistics*, 5:210-254.
- Pan, N. and W. Snyder, (2003), "Setting the Parameters of Syllable Structure in Early Dutch", in *Proceedings of the 27th Annual Boston University Conference on Language Development*, 2:614-625.
- Ploch, S. (2004), "What is Wrong with Linguistics?" in *SOAS Working Papers in Linguistics and Phonetics*, 13.
- Pöchtrager, M. A. (2006), "The Structure of Length", PhD Thesis, University of Vienna.
- Scheer, T. (1998), "Governing Domains are Head Final" in Cyran (ed.) *Structure and Interpretation. Studies in Phonology*, 261-285, Folium.
- Scheer, T. (1998b), "A Unified Model of Proper Government", in *the Linguistic Review*, 15:41-67.
- Scheer, T. (2004), "A Lateral Theory of Phonology", Mouton de Gruyter.
- Thompson, L. (1959), "Saigon Phonemics", in *Language*, 43:362-371.
- Thurgood, G. and R. LaPolla (eds.), (2003), "The Sino-Tibetan Languages", Routledge.
- Trask, L. (1997), "Historical Linguistics", Arnold.

Tournadre, N. and S. Dorje, “*Manual of Standard Tibetan*”, Snow Lion

Úlfssbjörninn, S. (2005), “*A Principles and Parameters Approach to Consonant Cluster Acquisition*”, ms., the School of Oriental and African Studies.

Úlfssbjörninn, S. (2006), “*The Vietnamese Syllabic Template; The difference between pointed and pointless onsets*”, ms., The School of Oriental and African Studies.

Úlfssbjörninn, S. (2006b), “*Some Differences between Italiano and Maremmano*” ms., the School of Oriental and African Studies.

Xu, Z. (2001), “*Phonological Analysis of Honkeui To*”, PhD dissertation, Guangdong University.