GARO GEMINATION: A CASE OF FAITHFULNESS PRODUCING NON-MORAIC GEMINATES

Cheman Baira A'gitok
University of Calgary

1. Introduction

Garo, an understudied Tibeto-Burman language of North-East India (Burling, 2003), has predictable geminates: when a morpheme that ends with a consonant combines with a morpheme that begins with a vowel to form a word, the final consonant of the first morpheme geminates. Cross-linguistically, geminate consonants are primarily of two types: some languages have underlying geminate consonants that contrast with singletons (Davis, 2011). These types of geminates are explained by positing a mora in the underlying form that is associated to the consonant (Hayes, 1989). Languages that have these geminates have words that contrast based solely on the length of the consonants. Another type of geminates is the result of a complete assimilation. These geminates arise when one consonant assimilates to all the features of an adjacent consonant such that the two consonants become featurally identical giving rise to surface geminates (Kotzor et al., 2017; Ridouane, 2010).

Neither the moraic or the assimilation analysis fit with the Garo gemination data, however. The present paper looks to answer the question of what the Garo gemination pattern is and how it can be formally accounted for in Optimality Theory (Prince & Smolensky, 2008). I conclude that due to the monosyllabic and agglutinating nature of the language, Garo geminates arise out of a conflict between restriction placed on the syllabification of the morpheme final consonant by an alignment constraint on the one hand, and the preference for an unmarked CV(C) syllable shape on the other hand. The findings also have important implications for the typology of gemination.

2. Background about Garo

Garo is a Tibeto-Burman language of the Boro-Garo group (Bradley, 1997) spoken primarily in North-East India, mainly in Meghalaya, but also in Assam. There is also a sizable number of speakers in neighbouring Bangladesh (Burling, 2003). There are roughly one million speakers across eight mutually intelligible Garo dialects[1] associated with different regions of Garo Hills. The A'we dialect, associated with the town of Tura, is the

[1] Traditionally, Garos include three additional linguistic systems in their dialects. Linguistically speaking however, these dialects are not Garo languages although they belong to the Bodo-Garo group. These dialects may best be called "cultural dialects," since the speakers of these languages are very much Garo in terms of their culture, and the larger Garo society also counts them an integral part of its culture.

Actes du congrès annuel de l'Association canadienne de linguistique 2021.
Proceedings of the 2021 annual conference of the Canadian Linguistic Association.

standard dialect used in teaching and writing, in addition to being the common dialect among Garos. The present paper is concerned with the standard dialect as spoken in Tura.

Garo has 17 phonemic consonants (Table 1). Not all of these consonants can occur in all syllable positions, however. The sounds /b, d, g, ɾ, s, h, w, t͡s, d͡z/ can only occur in onsets, while /ʔ, ŋ, l/ are limited to codas. The rest of the sounds, i.e., /p, t, k, m, n/ can occur in both onset and coda positions. The voiceless stops and affricate /p, t, k, t͡s/ are aspirated in onset position and the consonants that can occur in the coda position can geminate.[2]

Table 1. Garo phonemic consonants.

	Bilabial	Alveolar	Velar	Glottal
Plosives	p b	t d	k g	ʔ
Nasals	m	n	ŋ	
Taps		ɾ		
Fricatives		s		h
Approximants	w			
Lateral Approximants		l		
Affricates		t͡s d͡z		

Garo does allow onset consonant clusters, but the number of possible combinations is relatively few. The two classes of consonant clusters that are possible in Garo are of the configurations: /C_1ɾ/, where C_1 represents the consonants that can combine with /ɾ/ to form a cluster. The sounds that can go into the C_1 positions include /p, b, t, d, k, g, m, s, t͡s, d͡z/. The second configuration, i.e., /sC_2/ is a lot more restricted. The consonants that can go into the C_2 position are /p, t, k, ɾ/. Coda clusters are also possible in the language but there is only one possible configuration in this position, /Cʔ/. The consonants that can go into the C position are limited to the sonorants /m, n, ŋ, l/.

When it comes to the vowels, Garo has a fairly common six vowel inventory. These vowels are /i, e, a, ə, o, u/. There are almost no distributional restrictions on the vowels except for the fact that /i/ only occurs in open syllables and /ə/ only occurs in closed syllables[3]. Diphthongs are possible in the language, but the combination possibilities are few. The possible diphthongs include the /a/ vowel, with the configuration being /aV_2/ and the only vowels that can go into the V_2 position are the high vowels /i, u/.

The syllable is a central element in the phonology of Garo. As has been alluded to above, a lot of distributional restrictions and phonological processes are defined with reference to the syllable. When it comes to the possible structure of the syllable, the vowel

[2] It is unclear whether or not /ʔ/ and /ŋ/ can geminate however.

[3] It is possible to analyse the distribution of these vowels as being allophonic or being in a complementary distribution. But a consideration of the restriction on syllabification in the language (dealt with in the coming sections) makes the phonemic analysis of these vowels more appropriate from learnability viewpoint. The same thing can be said about the distribution of /ɾ/ and /l/.

is the only mandatory element, thus, vowel-only syllables are possible in the language. The syllable is free to have either an onset or a coda or both. The presence of one in the syllable does not preclude the presence of the other, nor does the absence of one preclude the presence of the other. Combining these with the fact that Garo allows consonant clusters in both the onset and coda positions, the possible syllable shape in the language can be schematized as (C)(C)V(C)(C). It has to be said at this point that the statements about the possible syllable shape is more accurately said about the underlying segmental make-up of the morphemes since the language is monosyllabic. It makes more sense to say the morphemes contain at most two consonants on either side of the vowel in Garo due to restrictions placed on the syllabification of these sounds whereby the final consonant of a morpheme does not resyllabify as the onset of the following syllable. For example, /wat/ + /-a/ surfaces as [wat.ta] "release-Pre," and /dok/ + /-a/ surfaces as [dok.ka] "hit-Pre," always displaying gemination, while the resyllabified forms *[wa.tʰa] and *[do.kʰa] are unattested in the language. These will be further elaborated on in the coming sections.

Garo is a quantity-insensitive language with fixed stress on the final syllable (Burling, 2003). Due to this, the stress always falls on the final syllable regardless of whether the final syllable is open or closed. The position of the stress is also unaffected by whether or not the preceding syllable is open or closed. This can be seen in the forms such as [dok.ˈna] "hit-Inf," [t͡sʰa.ˈna] "grow-Inf," and [naʔ.ˈtʰok] "fish," where the stress is always on the final syllable. Garo also lacks any length contrast on vowels.

With respect to morphology, Garo is a monosyllabic language with agglutinating morphology. In verbs in particular, Garo can have long and morphologically complex words. Verbs in their simplest form have a root morpheme and obligatory tense and often, aspect markers. This can be seen in (1), where morpheme boundaries are marked with "+".

(1)	a.	/t͡se/	+ /-a/	➔	[t͡sʰe.a]	
		win	Pre		"win-Pre"	
	b.	/ɾu/	+ /-aha/	➔	[ɾu.a.ha]	
		pour	Perf		"pour-Perf"	

In addition to the obligatory tense and aspect markers that follow the root in verbs, Garo morphology allows additional affixes to attach to the root before the tense and aspect affixes. These affixes encode additional grammatical information such as deixis and causation. The presence of these additional affixes in the verb form can be seen in (2) where they occur before the tense and aspect markers.

(2)	a.	/ɾeʔ/	+ /-aŋ/	+ /-na/	➔	[ɾeʔ.aŋ.na]	
		walk	movement away from speaker	infinitive		"to go"	

b.	/ɾaʔ/	+	/-baʔ/	+	/-na/	➔	[ɾaʔ.ba.na]
	take		movement towards the speaker		infinitive		"to bring"

3. Gemination in Garo

As I discussed above, Garo is an agglutinating language that allows grammatical morphemes to attach between a verb root and the obligatory tense/aspect suffixes. When this is combined with the fact that the language allows for the morphemes to have a variety of syllable shapes within the (C)(C)V(C)(C) template, it opens up the possibility for morphemes of various shapes to combine sequentially in verb forms. Out of these possible combinations, there are two that are of interest for the present paper. The first of these combinations are those where morphemes that end in a consonant precede a morpheme that begins with a vowel (i.e., a sequence of (C)VC + V(C) morphemes), as in (3), where some words have additional grammatical suffixes before the tense and aspect suffixes.

(3)	a.	/t͡san/	+	/-a/			➔	[t͡sʰan.na]
		count		Pre				"count-Pre"
	b.	/kat/	+	/-a/			➔	[kʰat.ta]
		run		Pre				"run-Pre"
	c.	/t͡sam/	+	/-a/			➔	[t͡sʰam.ma]
		portion		Pre				"portion-Pre"
	d.	/wat/	+	/-at/	+	/-a/	➔	[wat.at.ta]
		release		causative		Pre		"make (someone) go-Pre"
	e.	/t͡san/	+	/-at/	+	/-a/	➔	[t͡sʰan.nat.ta]
		count		causative		Pre		"make (someone) count-Pre"

As these examples demonstrate, whenever a morpheme that ends in a consonant is followed by a morpheme that begins with a vowel, the final consonant of the first morpheme always geminates rather than resyllabifying as an onset of the following syllable. This process applies without exception, throughout the word and it is not sensitive to stress or foot constituency. If gemination was sensitive to prosodic prominence, a difference would be expected in the forms (3d) and (3e), but the consonant geminates in both the second and the third syllables. This helps to rule out any conditioning of the process by prosodic prominence.

Gemination is observed under similar circumstances even in forms that could be called compound verbs. Consonant gemination is seen when two root morphemes of appropriate shape combine in the right order, i.e., VC + V, to form a compound verb. This

can be seen in (4c, 5c) where the first member of the compound ends in a consonant and the following member begins with a vowel. (In (4) and (5), the constituent roots are shown in their simplex forms first, after which their combination into compound forms are shown.) The data in (4-5) show that gemination occurs in compound forms too, indicating that it is a reflex of morpheme concatenation in general and it is not limited to suffixation. A combination of VC and V morphemes always produce a geminate consonant. Also, as I discussed in the previous section, resyllabification of the type VC+V → V.CV is not attested in the language. The combinations of this type always result in a consonant gemination rather than a resyllabification.

(4)	a.	/sal/ pull	+	/-a/ Pre			→	[sal.la] "pull-Pre"
	b.	/on/ low	+	/-a/ Pre			→	[on.na] "low-Pre"
	c.	/sal/ pull	+	/on/ low	+	/-a/ Pre	→	[sal.on.na] (compound) "pull down-Pre"
(5)	a.	/dok/ hit	+	/-a/ Pre			→	[dok.ka] "hit-Pre"
	b.	/ok/ remove	+	/-a/ Pre			→	[ok.ka] "remove-Pre"
	c.	/dok/ hit	+	/ok/ remove	+	/-a/ Pre	→	[dok.kok.ka] (compound) "remove by hitting (something)-Pre"

3.1 Forms where gemination does not occur

While gemination is seen to occur without exception in VC + V sequences, the mirror image of it is unattested in the language. Specifically, gemination never occurs in V + CV sequences, where a morpheme that ends in a vowel combines with a following morpheme that begins with a consonant, as in (6). In these cases, no gemination is observed and the initial consonant of the following morpheme syllabifies as an onset.

(6)	a.	/t͡sa/ emerge	+	/-baʔ/ movement towards the speaker	+	/-na/ infinitive	→	[t͡sʰa.ba.na] "to start to emerge"

b. /ka/ + /-do/ + /-bo/ ➔ [kʰa.do.bo]
tie climb imperative "tie up-imperative"

Similarly, gemination is also not attested in VC + CV sequences, where a morpheme that ends in a consonant combines with another morpheme that begins with a consonant. This can be seen in (7). Thus, gemination is blocked when the following morpheme begins with a consonant.

(7) a. /t͡sak/ + /-na/ ➔ [t͡sʰak.na]
support infinitive "to support"

b. /tap/ + /-ɾoŋ/ + /-na/ ➔ [tʰap.ɾoŋ.na]
paste habitual infinitive "to always paste"

It is also interesting to note what happens when a morpheme that ends in a vowel combines with another morpheme that begins with a vowel, i.e., in (C)V + V(C) sequences. As we see in (8), in V + V sequences, there is no resolution to the hiatus through consonant epenthesis and a vowel hiatus appears in the surface form. In fact, Garo phonology lacks epenthesis as a phonological process more generally. The lack of epenthesis is not only seen word internally, but also word initially (8b).

(8) a. /t͡sa/ + / -a/ ➔ [t͡sʰa.a]
emerge Pre "emerge-Pre"

b. /o/ + /-a/ ➔ [o.a]
open Pre "open-Pre"

Another environment where VC + V combinations (which result in gemination) are possible are at word junctures. A question naturally emerges therefore as to whether or not there is consonant gemination in this environment. The data reveals, however, that there is no gemination across word boundaries[4] even though the syllable configuration that could trigger the process is present. Gemination across word boundaries are unattested in the language. This can be seen in (9).

(9) a. /aʔ + pal/ /oʔ + gəmmən/ ➔ [aʔ.pʰal] [oʔ.gəm.mən]
agricultural field cleared "a cleared field"

[4] Since gemination occurs in compounds, but not across word boundaries, it is likely that the phonological domain in which gemination applies is the phonological word. The prosodic structure of compound constructions in Garo is not known at the moment, however. So, it is possible that the relevant prosodic boundary is a higher-level boundary. This is beyond the scope of the current paper however, and will need to be investigated in a future paper.

b.	/bi + gəl/	/anʔ + gəl/	➔ [bi.gəl] [anʔ.gəl]
	skin	hue	"skin tone"

4. Existing models and their incompatibility with Garo data

While consonant gemination remains unanalysed in the language, Duanmu (1994), which is based on the data from Burling (1992), puts forward a proposal of a fixed syllable structure [CVX] in Garo to account for vowel related phonological processes in the language. These vowel processes include the distributional restrictions on /i/ and /ə/, and also the mid ~ high vowel alternation in the southern Garo dialect (see Duanmu (1994) and Burling (1992) for more information). What is relevant about Duanmu's proposal for this paper is that the proposed [CVX] template could be translated into the moraic model, which would assign a mora to all coda consonants.

Taking Duanmu's proposal into account, the moraic account needs to be explored as an explanation for the geminate consonants in Garo. Cross-linguistically, there are languages that have geminate consonants arising from an underlying mora. Languages like Japanese contrast long and short consonants, e.g., [sa.ka] "hill" and [sak.ka] "author." Long consonants in forms such as [sak.ka] are explained by positing an underlying mora associated to the consonant (Hayes, 1989). In [sak.ka], the [k] syllabifies as the onset of the second syllable first, but also need to associate to the coda position due to an underlying mora on it as onsets cannot have a mora. This leads to its double association and thus phonetic lengthening. Syllabification of consonants are essentially unpredictable in languages with underlying geminates. The difference in the underlying specification of the [k] in [sa.ka] and [sak.ka] are represented in Figure (1).

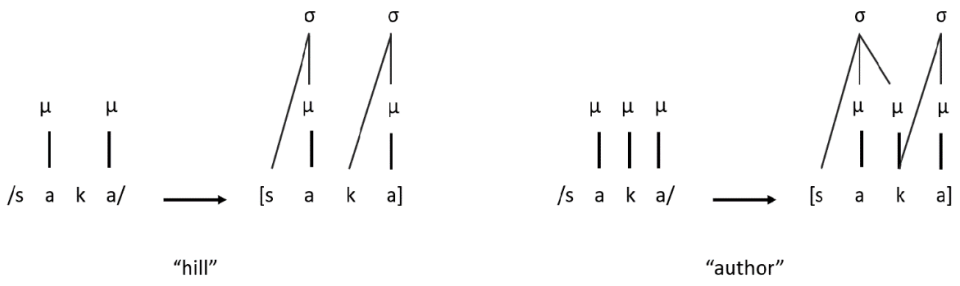

Figure 1. Syllabification of Japanese words with and without geminates in moraic theory. Data from Davis (2011).

Garo pattern is very different from the Japanese, however, as the morpheme final consonants in VC + V sequences always syllabify as the coda of the first syllable and then geminate, e.g., in combinations such as /pak/ + /-a/ → [pʰak.ka] "pour-Pre," and /nam/ + /-a/ → [nam.ma] "good-Pre," the morpheme final consonants always geminate. Morpheme final consonants that do not geminate and simply syllabify as the onset of the following syllable are not attested in the language, e.g., syllabifications such as /pak/ + /-a/ →

*[pʰa.kʰa] and /nam/ + /-a/ → *[na.ma] are unattested in the language in VC + V sequences (Figure 2).

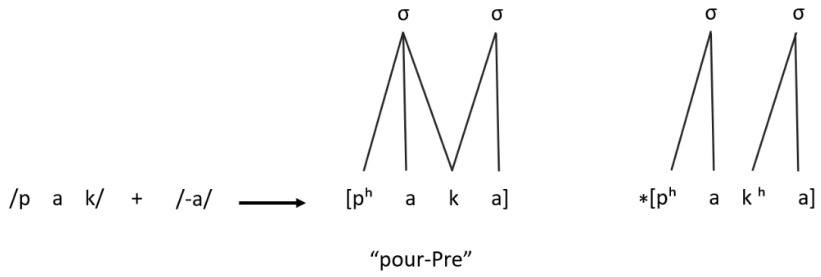

Figure 2. Syllabification of morpheme final consonants in Garo.

Due to morpheme final consonants always geminating in VC + V sequences, the moraic account would need to specify a mora for all morpheme final consonants. Syllabification of morpheme final consonants are completely predictable in Garo, however. The morpheme final consonants do not syllabify differently, that is they do not sometimes syllabify as the onset of the following syllable and then other times geminate. They consistently geminate without any exception. Thus, the mora becomes redundant. This is as opposed to languages like Japanese, where the mora has an actual function in the analysis as it explains the unpredictable syllabification of consonants in what would be identical words as [sa.ka] vs. [sak.ka]. In Garo however, one ends up with a mora in the underlying form for a predictable phonological process, which goes against the very idea of the underlying forms, since underlying representations are supposed to capture what is unpredictable. When the redundancy of the mora is coupled with the fact that Garo is a quantity-insensitive system with fixed stress, one ends up with a mora that is not only redundant, but also ad hoc to gemination. These factors rule out underlying mora as a desirable explanation for gemination in the language. By extension, Duanmu's syllable model for Garo also does not work.

Another existing explanation for geminate consonants that needs to be considered for the Garo data is that of assimilated consonants. Languages like Bengali have geminate consonants that arise due to assimilation. In Bengali, the /r/ consonant assimilates to the manner of articulation of the following coronal consonant in words like /mar + t + o/ → [mat:o] "beat-Past-Habitual-3P" (Kotzor, Wetterlin, & Lahiri, 2017). Due to this assimilation, the coronal consonant lengthens to give surface geminates. These geminates are a product of phonological processes and do not have an underlying specification for a mora. The assimilation account does not fit with the Garo data, however, since there is no gemination seen in cases of VC + CV sequences. For example, there is no gemination in the word such as /t͡sak/ + /-na/ → [t͡sʰak.na] "support-Inf," with forms such as *[t͡sʰan.na] and *[t͡sʰak.ka] being unattested. It is in these sequences that geminates arise in languages like Bengali. Additionally, if it was the case that the following morpheme had a glottal stop onset that was assimilating to produce a geminate, it should show up in V + V sequences, but these forms surface with a hiatus, e.g., /o/ + /-a/ → [o.a] "open-Pre".

It is thus clear that Garo data is not explained by existing accounts of consonant gemination. The following section puts forward an analysis for the Garo data based on syllabification.

5. Current analysis

As the previous section makes clear, the existing accounts of geminate consonants cannot explain the Garo data. The moraic analysis does not work because the morpheme final consonants in VC + V sequences have a predictable syllabification. They do not sometimes syllabify as the onset of the following syllable and then other times geminate. They always syllabify as the coda of the preceding syllable and then geminate, making the mora redundant. The assimilation account also does not work because of the fact that there is no gemination seen in VC + CV sequences, the exact environment where assimilated geminates arise. A close look at the data reveals crucial patterns about where gemination is occurring however. It is always in the case of VC + V morpheme sequences that geminates arise. This combinatorial configuration is a crucial piece of the puzzle as Garo is a monosyllabic language.[5] What can be seen in these sequences is that the final consonant of the preceding morpheme is not syllabifying as the onset of the following syllable. It is thus clear that the ends of syllables need to align with the end with of morphemes in Garo.

The fact that gemination occurs however, shows two things: one, it shows that although resyllabification and epenthesis through which a syllable could get an onset is not allowed, the language still tends towards having an unmarked syllable shape CV(C) by having an onset. This is where gemination occurs. Since outright resyllabification is not possible, the language doubly associates the morpheme final consonant to give an onset to a following vowel initial morpheme. Second, it shows that although it is imperative that the end of syllables align with the end of morphemes, it is not necessary that the beginning of syllables align with the beginning of morphemes. Having an unmarked syllable shape through gemination essentially takes precedence over alignment of the beginnings of syllables and morphemes.

The picture that emerges is that of a conflict between various forces in Garo phonology, as well as an interaction between phonology and morphology. There are restrictions on epenthesis and syllabification, but the language still tends towards an unmarked syllable shape CV(C). This conflict is very straightforwardly modelled formally in OT (Prince & Smolensky, 2008) through constraint ranking, as I will discuss in the next sections. The restriction on syllabification due to phonology-morphology interaction is handled by an extant set of constraints known as alignment constraints (McCarthy & Prince, 1993). The reason for why gemination arises at all can be found in the active markedness constraint that demands unmarked syllable shapes. It is the interaction of these constraints

[5] Interestingly, children are taught in Garo primary schools to break down words into constituent morphemes. For example, words like *paka* (/pak + a/ → [pʰak.ka]) "pour-Pre" are taught as ***pak-a***, where the two morphemes are separated with a dash.

in addition to the high ranking of additional faithfulness and alignment constraints that give rise to Garo geminates.

5.1 Lack of resyllabification

The datasets in the previous sections show that a morpheme final consonant does not resyllabify as an onset of a following vowel-initial morpheme when they combine. As can be seen in the /t͡san/ + /-a/ → [t͡sʰan.na] "count-Pre," the final consonant of the first morpheme syllabifies as the coda of the first syllable. This shows that the ends of syllables and morphemes need to align in Garo. This part of the pattern is captured by the alignment constraints proposed by McCarthy & Prince (1993):

ALIGN(MORPH-R, σ-R)
The right edge of a morpheme coincides with the right edge of a syllable.

This constraint requires the alignment of the right-edges of morpheme and syllable. This constraint needs to be ranked relatively high in Garo in order to rule out resyllabification of the final consonant. This constraint needs to outrank a well-formedness constraint that requires a resyllabification of the morpheme final consonant to the onset of the following syllable in other languages, i.e., forms such as /t͡san/ + /-a/ → *[t͡sʰa.na]. The constraint that demands onsets for syllables has been defined in the literature as:

ONSET
*[$_{\sigma}$ V *('Syllables must have onsets.')*

The relative ranking of these constraints can be expressed as:

ALIGN(MORPH-R, σ-R) >> ONSET

This constraint ranking rules out resyllabification. The presence of gemination in the data reveals however, that although ONSET is ranked lower than ALIGN(MORPH-R, σ-R), it is still active in that it outranks another constraint.

5.2 Onset over strict alignment

The presence of gemination reveals that although onsets cannot be assigned by resyllabification, the phonological system of the language still prefers to have syllable onsets. This reveals that although the language prioritizes the alignment of right-edges of morphemes and syllables, having onsets is still prioritized over aligning the left-edges. The constraint that requires the alignment of the left-edges of morphemes and syllables is defined as:

ALIGN(MORPH-L, σ-L)
The left edge of a morpheme coincides with the left edge of a syllable.

This constraint needs to be outranked by the ONSET constraint in order for consonant gemination to be possible. If this constraint outranked the ONSET constraint, then the forms such as /t͡san/ + /-a/ would surface with syllables that maximally align with morpheme structure, i.e., *[ts͡ʰan.a]. These forms are ungrammatical in the language however, hence the need for the ONSET constraint to outrank the ALIGN(MORPH-L, σ-L) constraint. The relative ranking of these constraints can be expressed as:

ONSET >> ALIGN(MORPH-L, σ-L)

5.3 Lack of epenthesis

The issue that needs to be addressed at this point however, is the absence of epenthesis. As was seen in the previous section, the emergence of geminates is due to the language preferring syllable onsets over the maximal preservation of morpheme structure. The question then arises as to why epenthesis is not observed if syllable onsets are preferred. Since the ONSET constraint dominates the ALIGN(MORPH-L, σ-L) constraint, forms such as: /o/ + /-a/ → [o.a] "open-Pre" should have an epenthetic consonant in the onset positions. The fact that there is no epenthesis in these forms reveal that ONSET itself is outranked by a faithfulness constraint that disallows epenthetic segments. This constraint is defined in the literature as:

DEP-IO
Output segments must have input correspondents ('No epenthesis').

This constraint rules out epenthesis, but it need to outrank ONSET in the constraint hierarchy in order to produce the Garo patterns. The relative ranking of DEP-IO with regards to ONSET can be expressed as:

DEP-IO >> ONSET

This constraint ranking rules out the forms: /o/ + /-a/ → [Co.Ca], where [C] is an epenthetic consonant. The failure to resolve vowel hiatus is thus captured by this relative ranking of the constraints.

5.4 Prosodic-word junctures

The prosodic-word junctures do not display any gemination even though the syllable structure that could trigger the process exists. This indicates that there are high ranking word-level alignment constraints that outrank the ONSET constraint. These alignment constraints are defined in the literature as:

ALIGN-R
The right edge of the grammatical word coincides with the right edge of the PrWd

ALIGN-L

The left edge of the grammatical word coincides with the left edge of the PrWd

In order to completely rule out gemination across prosodic-word boundaries, both of these alignment constraints need to outrank ONSET. The relative ranking of these constraints with regards to one another cannot be deduced from the data at hand however. The relative ranking of these alignment constraints with regard to ONSET can be expressed as:

ALIGN-R, ALIGN-L >> ONSET

This relative ranking rules out forms such as /bi.gəl/ /anʔ.gəl/ → *[bi.gəl] [lanʔ.gəl] "skin tone" for (9b) where the final consonant of the first word geminates to the onset position of the initial syllable of the following word.

5.5 Final ranking and tableaux

Considering the individual relative rankings of various constraints in the previous section together, the following constraint hierarchy can be arrived at:

ALIGN-R, ALIGN-L, DEP-IO, ALIGN(MORPH-R, σ-R) >> ONSET >> ALIGN(MORPH-L, σ-L)

This constraint ranking should be able to generate the Garo patterns if the analysis has been correct. This can be tested by creating tableaux (1-3) with the final ranking and inputting representative examples into them. Since the ALIGN-R and ALIGN-L constraints are undoubtedly highly ranked, and also since they prevent, rather than cause gemination, they will not be included in the tableaux.

Tableau 1.

/t͡san + -a/	DEP-IO	ALIGN(MORPH-R, σ-R)	ONSET	ALIGN(MORPH-L, σ-L)
☞ [t͡sʰan.na]				*
[t͡sʰan.a]			*!	
[t͡sʰa.na]		*!		*
[t͡sʰa.a]		*!	*	

Tableau 2.

/t͡sa + -baʔ + -na/	DEP-IO	ALIGN(MORPH-R, σ-R)	ONSET	ALIGN(MORPH-L, σ-L)
☞ [t͡sʰa.ba.na]				
[t͡sʰab.ban.na]		*!*		
[t͡sʰab.an.a]		*!*	**	**

Tableau 3.

/t͡sa + a/	DEP-IO	ALIGN(MORPH-R, σ-R)	ONSET	ALIGN(MORPH-L, σ-L)
☞ [t͡sʰa.a]			*	
[t͡sʰa.Ca]	*!			*
[t͡sʰaC$_1$.C$_1$a]	*!	*		*
[t͡sʰaC.aC]	**!	**	*	

The above tableaux show that the constraint hierarchy selects the correct optimal candidate. Tableau 1 shows how geminate consonants arise. Other candidates violate the higher-ranking constraints, so [t͡sʰan.na] is chosen as the optimal candidate even though it violates the ALIGN(MORPH-L, σ-L) constraint. Tableau 2 shows why regressive gemination is not possible in the language. The constraint hierarchy choses the maximally optimal candidate [t͡sʰa.ba.na] since the other candidates violate the higher ranking ALIGN(MORPH-R, σ-R) constraint. Likewise, Tableau 3 shows why [t͡sʰa.a] is the optimal candidate although it violates the ONSET constraint. Other candidates violate the higher-ranking constraint DEP-IO which rules them out. This tableau shows why epenthesis is not an option in the language.

6. Discussion

The present paper investigates Garo geminate consonants with the goal of discovering the gemination pattern and to formally account for the pattern in the Optimality Theory framework. The current analysis has revealed that the gemination process in Garo is due to a restriction placed on the syllabification of morpheme final consonants by the alignment constraint ALIGN(MORPH-R, σ-R) which is ranked higher than the ONSET constraint. When VC + V morpheme sequences combine to form a word, the morpheme final consonants cannot resyllabify as the onset of the following syllable since the alignment of right-edges of morphemes and syllables are of great priority in the language. This is unusual cross-linguistically, as intervocalic consonants typically syllabify as the onset of the following syllable, and this is what is also predicted by syllabification theories (Blevins, 1996). The phonological system of the language however, still prefers unmarked syllable shapes by having onsets whenever possible. The conflict between alignment of right edges of morphemes and syllables and the demand for onsets is essentially what gives rise to gemination in VC + V sequences. The very fact that gemination occurs in VC + V sequences comes at a cost of misalignment of the left-edges of morphemes and syllables, however.

Interestingly, although the language tends towards unmarked syllable shapes where possible, epenthesis is not attested, even when there is no possibility for gemination to occur. Due to this, vowel hiatuses are not resolved in the language and epenthetic consonants are not found. This is very easily explained by the DEP-IO however, which rules

out epenthesis. DEP-IO needs to be ranked pretty high in the constraint ranking of Garo anyway due to lack of epenthesis as a phonological process. Lack of resolution of hiatus is odd only if the relative ranking ONSET >> ALIGN(MORPH-L, σ-L) is considered, as in this case the language could simply insert an epenthetic consonant to resolve the hiatus. The lack of epenthesis is resolved by the ranking DEP-IO >> ONSET >> ALIGN(MORPH-L, σ-L). It is not only in case of vowel hiatus that epenthesis does not occur, but it is absent in the language altogether.

Having explained the gemination process as arising out of complex constraint interaction, a look at the typology of geminate consonants presents an interesting picture. Typologically, geminate consonants are of two types. The first type is the underlying geminates, which are underlyingly specified for a mora (e.g., in Japanese). The second type of geminates seen in languages is a product of assimilation (e.g., in Bengali). Assimilated geminates arise due to one of two adjacent consonants assimilating to the features of the other consonant. The case of Garo geminates does not fit within either type in the existing typology, however. Garo geminates are not a product of an underlying mora (type one) since the morpheme final consonants in VC + V always geminate nor are they a product of assimilation (type two) since no gemination occurs in VC + CV sequences. It will be worthwhile to investigate how common Garo-type geminates are, crosslinguistically. Monosyllabic languages that are also agglutinating, like Garo, are the natural candidates for such a study to shed light on the relation between gemination and the morphological characteristics of such a language.

There are still a lot of unaddressed issues regarding gemination in Garo. Although this paper has put forward a formal analysis to account for the pattern, there is still a need for a phonetic study to test the degree and quality of the phonetic difference between geminate consonants and other consonant types such as singletons and consonant sequences. There is particular interest with regards to the nasals since geminate nasals are ambiguous with sequences of nasals. It is also unclear if the velar nasal geminates at all since it can only occur in coda positions in the language. Additionally, while voiceless plosives and affricates aspirate in onset positions, geminate plosives do not aspirate. Gemination also interacts with glottal stop deletion, which also needs to be both formally and phonetically studied. These issues need to be investigated further in future studies.

7. Conclusion

Geminate consonants are typically explained as either being a consequence of being underlyingly specified for a mora, or being a product of assimilation processes. A look at the typological data supports this view as geminate consonants in languages can be analysed as either an underlying or an assimilated geminate. Data from Garo, an understudied Tibeto-Burman language, however, presents a third kind of geminates. Garo geminate consonants are not underlying geminates, and although they are surface geminates, they are not a product of assimilation, either. My proposal is that gemination is the result of the interaction between ALIGN(MORPH-R, σ-R) and ONSET constraints. Garo thus presents a strong case for the expansion of the typology of geminate consonants and

the need to look closely at monosyllabic agglutinating languages with respect to gemination pattern and explore the potential effect of their morphological characteristics in phonological processes such as consonant gemination.

References

Bradley, D. (1997). Tibeto-Burman languages and classification. In *Papers in Southeast Asian Linguistics* (Vol. 14). Pacific Linguistics. https://doi.org/10.15144/PL-A86.cover

Burling, R. (1992). Garo as a minimal tone language. *Linguistics of Tibeto Burman Area*, *15:2*(Fall 1992), 33–51.

Burling, R. (2003). *The Language of the Modhupur Mandi (Garo), Volume 1*. Michigan Publishing, University of Michigan Library.

Davis, S. (2011). Geminates. In *The Blackwell Companion to Phonology*. John Wiley & Sons.

Duanmu, S. (1994). The Phonology of the Glottal Stop in Garo. *Linguistics of the Tibeto-Burman Area*.

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry*, *20*(2), 253–306.

Kotzor, S., Wetterlin, A., & Lahiri, A. (2017). Bengali geminates: Processing and representation. In H. Kubozono (Ed.), *The phonetics and phonology of geminate consonants*. Oxford University Press.

McCarthy, J. J., & Prince, A. (1993). Generalized Alignment. *Yearbook of Morphology*. https://doi.org/10.1002/9780470756171.ch7

Prince, A., & Smolensky, P. (2008). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University. https://doi.org/10.1002/9780470756171.ch1

Ridouane, R. (2010). Gemination at the junction of phonetics and phonology. *Laboratory Phonology*, *10*.