

3.0. Introduction

In this chapter, I present extensive descriptive data from the phonotactics of Georgian. This is meant as background to the assumptions that I make in the analysis of Georgian phonotactics, specifically the analysis of Georgian consonant sequences in Chapter 6.

The study is restricted to Modern Literary Georgian. A more comprehensive study would be needed to account for richer dialect data, e.g. incorporating vowel systems with length distinctions and umlaut. Nevertheless, dialect data are sometimes discussed when relevant for the justification of the plausibility of particular arguments.

This chapter is organised as follows: in 3.1 a general introduction to the Georgian language is given. Morphological patterns are discussed with respect to the phonotactic structure of a word; in 3.2 the phonemic inventory and general syntagmatic regularities of vowels are considered; in 3.3 special attention is paid to the consonantal system, represented by three sets: stops, fricatives and sonorants. Two important claims are made: the sonorant /r/ is syllabic in a consonantal environment, and combinations of C + /v/ are complex segments; in 3.4 specific characteristics of the Georgian minimal word are discussed. The disyllabic minimal word constraint is proposed; in 3.5 a general introduction to consonant syntagmatics is offered. All possible combinations of consonants of different lengths and constituency are discussed. Harmonic clusters are argued to be complex segments, on the basis of distributional, comparative and historical evidence. It is proposed that the longest true consonant sequences are biconsonantal; in 3.6 the hypothesis concerning the derived nature of consonant sequences is tested. The restrictions on consonants at a distance (across a vowel) and in adjacency (in sequences) are compared; in 3.7 the generalisations are summarised and general conclusions on Georgian phonotactics are drawn.

3.1. The Georgian language

Georgian belongs to the Kartvelian (South Caucasian) language family. The name of the language group is related to the ethnonym *kartveli* 'Georgian'. Georgian is the official state language of Georgia, with more than 5 million speakers. Outside Geor-

gia, there are Georgian-speaking populations in Azerbaidjan (Saingilo),¹ Turkey (Shavsheti, Imerkhevi), Iran (Fereidan) and the North Caucasus (Sochi, Kizlar-Mozdok, Orjonikidze). Besides Georgian, the Kartvelian language family consists of three other languages, Megrelian (Mingrelian), Laz (Chan) and Svan. The Kartvelian languages are spoken within the territory of Georgia, the Central and West Caucasus and the vast territory of the South Caucasus. Megrelian is spoken in the western part of Georgia, namely, in the northern part of the Kolkheti Lowlands and in the neighbouring mountainous region. Two dialects of this language can be distinguished: Senakuri and Zugdidur-Samurzakanuli. Laz is spoken near the Black Sea, from the village of Sarpi (on the southern border of Georgia) to Kerem (Turkey). It consists of three dialects: Atinuri, Vicur-Arkabuli and Khopuri. Svan is spoken in the northwest of Georgia. Four dialects of this language can be distinguished: Balszemouri, Balskvemouri, Lashkhuri and Lentekhuri.

The kinship and common origin of the Modern Kartvelian languages has been confirmed by numerous studies on phonetic and morphological correspondences (Güldenstädt 1787, Deeters 1926, 1930, Bopp 1846, 1847, Brosset 1849, Gamkrelidze & Machavariani 1965, Sardschweladse & Fähnrich 1990). Much research has been done to establish the genetic affiliation of Kartvelian languages with other language families or individual languages, e.g. Basque (Bouda 1949, 1950), Indo-European (Bopp 1846, 1847), Semitic (Marr 1908) and North Caucasian (Bouda 1954–56, Chikobava 1965). Although a number of typological similarities have been found, to my knowledge no one has demonstrated beyond doubt the genetic affiliation of the Kartvelian language group to any other language group in the world.

According to Shanidze's (1973) classification, which is mainly based on geographical factors, there are six groups of Georgian dialects: 1. Pkhouri (Khevsuruli, Mokheuri, Tushuri); 2. Mtiulur-Pshauri (Mtiulur-Gudamakruli, Pshauri); 3. Kartlur-Kakhuri (Kartluri, Kakhuri, Javakhuri, Meskhuri, Kizikuri); 4. Dasavluri (Imeruli, Guruli, Rachuli, Lechkhumuri); 5. Samkhret-Dasavluri (Acharuli, Imerkheuli); and 6. Ingilouri.

Georgian is the only written language of the Kartvelian language family. The oldest written documents that are still preserved can be dated back to the fourth century. According to some specialists, the writing system is much older than the texts, and its origin is related to the Georgian king Parnavaz, who reigned in the third century B.C. Gamkrelidze (1989) has argued for a connection between the origin of the Georgian alphabet and the conversion of Georgia to Christianity. Taking into account the paradigmatic and syntagmatic characteristics of the Georgian alphabet, he demonstrates the influence of the Greek alphabet on the Georgian one. The same idea had been expressed by Gardthousen (1879), Bakradze (1889) and K'ek'elidze (1929), but had not been systematically demonstrated. The alloglottographic method² of writing was widely used in Georgia and in the Caucasus region

¹ Names of the geographical areas are given in brackets.

² The alloglottographic method involves the transcription of a language with symbols of the writing system of another language. The method is commonly used when a language does not have its own writing system.

in general, before the spread of Christianity. Prior to the introduction of the alphabetic system, the Georgian language was encoded through the Aramaic writing system.

The modern Georgian writing system is based on the round-form cursive, which was developed from the angular book script of the ninth century; the latter was a direct descendant of the Old Georgian script. The Georgian writing system includes a symbol, represented by a single grapheme, for each of the distinctive sounds (phonemes) of the Georgian language. The phonemic inventory of Georgian consists of 33 phonemes: 28 consonants and 5 vowels. In the table in (1) all phonemes are presented in alphabetical order. Phonemes are presented in their original Georgian graphic form (i), followed by the name of the symbol (ii), the notation used throughout this thesis (iii) and IPA equivalents (iv). Note that the symbol ' denotes glottalisation in obstruents.

(1) *The Georgian Alphabet*

(i)	(ii)	(iii)	(iv)	რ	rae	r	[r]
ს	an	a	[a]	ს	san	s	[s]
ბ	ban	b	[b]	ბ	t'as	t'	[t']
გ	gan	g	[g]	უ	un	u	[u]
დ	don	d	[d]	პ	par	p	[p]
ე	en	e	[e]	კ	kar	k	[k]
ვ	vin	v	[v]	ყ	yan	y	[y]
ზ	zen	z	[z]	ჯ	χ'ar	χ'	[χ']
თ	tan	t	[t]	შ	šin	š	[ʃ]
ი	in	i	[i]	ჩ	čin	č	[tʃ]
ქ	k'an	k'	[k']	ც	can	c	[ts]
ლ	las	l	[l]	ძ	jil	j	[dz]
მ	man	m	[m]	წ	c'il	c'	[t's]
ნ	nar	n	[n]	ჭ	č'ar	č'	[t'ʃ]
ო	on	o	[o]	ხ	xan	x	[χ]
პ	p'ar	p'	[p']	ჯ	jan	j	[dʒ]
ჟ	žan	ž	[ʒ]	ჰ	hae	h	[h]

Before considering the phonemic inventory, a short introduction to Georgian morphology is in order. Georgian morphology is rich, with a very productive inflectional and derivational morphology. Since the technique of morpheme combination is both agglutinative and inflectional, Georgian can be characterised as a Grammatical language (see Chapter 2 for the terminology). Grammatical affixes attach before or after a root, and may sometimes even violate the integrity of the root if a metathesis process takes place. Both the noun and the verb are characterised by prefixing and

suffixing morphology, but verbal morphology is much more complex. Consider the morphologically complex nominal (2a) and verbal (2b) forms in (2).

- (2) a) kud-ian-eb-is-tvis ‘for people wearing hats’
 kud- N ROOT
 -ian- POSS
 -eb- PL
 -is- GEN
 -tvis POSS
- b) ga-u-k’et-eb-in-eb-i-a ‘somebody made somebody else do something’
 ga- PREV
 -u- PRE-RADICAL VOWEL (OBJECTIVE VERSION MARKER)
 -k’et- V ROOT
 -eb- THEM SUFF
 -in- CAUS
 -eb- THEM SUFF
 -i- PERF
 -a 3rd SUBJECT MARKER

When affixes attach to a root, several phonological processes can take place, e.g. vowel deletion, metathesis and epenthesis. The deletion of the root vowel is the most widespread process in both nominal and verbal morphology. As a result of vowel deletion, consonant sequences emerge, as illustrated in (3).

- (3) a) NOM GEN
 xmal-i xml-is ‘sword’
 cxvar-i cxvr-is ‘sheep’
- b) da-č’er-i ‘cut IMP’ da-č’r-a ‘he had cut’
 PREV - V ROOT PREV - V ROOT - 3rd MARKER

Surface consonant sequences also emerge as a result of simply adding consonantal prefixes or suffixes to a root, e. g.

- (4) a) gv-c’er ‘you write us’
 2nd PL MARKER - V ROOT
- b) gv-xat’-av ‘you paint us’
 2nd PL MARKER - V ROOT - THEM SUFF
- c) m-c’er-al-i ‘writer’
 NOMINALISER - V STEM - PRT MARKER - NOM MARKER

Thus, one of the ways to account for the surface consonant sequences of Georgian is to study the morphological structure of such words. Sequences may be the result either of the deletion of a root vowel, which generally happens when a vowel-initial affix is added to a root, as illustrated in (3), or of the addition of a consonantal affix to a root, as illustrated in (4).

I have already mentioned that, in general, lexical and grammatical morphemes differ in their use of the phonological inventory. Morphological languages demonstrate the discrepancy between the two: grammatical morphemes use a subset of the phonological inventory of a language, whereas the full set is employed for lexical morphemes. Georgian, as a Grammatical language, also demonstrates this pattern. I will briefly summarise my findings concerning this issue.

Of the total of 28 consonants of Georgian, 13 can occur in both lexical and grammatical (derivational and inflectional) morphemes; while the other 15 occur exclusively in lexical morphemes.

- (5) a) Consonants appearing only in lexical morphemes:
/p p' t' k č j c' č' j z š ž χ' x h/
b) Consonants appearing in lexical and grammatical morphemes:
/b t d g k' c s γ v m l r n/

It is interesting to note that the set in (5a) includes almost all affricates (i.e. consonants with a complex stop + fricative constituency), e.g. /č c' č' j j'/, and other marked consonants, e.g. /h/ and /ž/, the distribution of which is quite restricted, e.g. /h/ occurs only in word-initial position, while the fricative /ž/ occurs only in a few words of Georgian origin.

Consonants in grammatical affixes given in the set in (5b) have different distribution patterns depending on their position in a derived word.

- (6) a) Word-initial consonants: /v m n s³ d g/
b) Word-medial consonants: /v m l r n t d g b k' γ/
c) Word-final consonants: /v m n s t d t b⁴ c/

It is interesting to observe that the velars /k'/ and /γ/ appear only in word-medial position, the affricate /c/, which is the only affricate consonant that appears in grammatical morphemes, occurs only in word-final position and the liquids /r/ and /l/ occur only in word-medial position. Coronals, nasals and the sonorant /v/ appear to be quite free in their distribution. All of this has direct consequences for the surface complexity of consonant sequences in Georgian, and I return to this issue later. One more thing to notice is that of the 13 consonants given in (5b), only eight, /v m n s t d g b/, occur in inflectional affixes, e.g.

³ It should be noted that depending on the root-initial sound, the person marker /s-/ has several allomorphs, for instance, /h-/ and /š-/. In Old Georgian, the morpheme also had an allomorph /x-/.

⁴ Word-final /d/ and /b/ devoice and are pronounced as [t] and [p], respectively, e.g. /k'arg-ad/ 'well' is pronounced as [k'argat] and /v-a-k'eteb/ 'I do' is pronounced as [vak'etep].

- (7) /v-/ 1st SUBJECT MARKER
 /m-/ 1st OBJECT MARKER
 /s-/ 3rd OBJECT MARKER
 /-t/ PL MARKER
 /-d-/ POSS MARKER
 /g-/ 2nd OBJECT MARKER
 /-b/ is attested in the thematic suffix /-eb/

The constituency of the set is reminiscent of the ‘primary’ consonants discussed in Chapter 2, which appear to be the most unmarked consonants not only in Georgian, but also cross-linguistically. In previous chapters I discussed the correlation between two asymmetries, consonant vs. vowel and lexicon vs. grammar, and proposed that consonants and the lexicon are on the same plane, as opposed to vowels, which are on a par with the grammar. Georgian data also substantiate this cross-linguistic observation. According to Ertelishvili (1970), there are no Georgian lexical roots containing only one vowel; however lexical, monoconsonantal and biconsonantal roots are quite common. For example, Ertelishvili (1970) gives lists of 14 such monoconsonantal and 45 biconsonantal roots. There are also longer roots containing three consonants (42 such roots are attested) or maximally four consonants (14 such roots are attested). I return to the patterns of consonantal roots by discussing their respective histories in Chapter 6. I will propose that such roots are derived; they emerge as the result of vowel deletion, and are sometimes the result of a lexicalisation process or a conflation of two roots/stems. Note that grammatical morphemes containing a vowel only are quite common in Georgian. Monoconsonantal affixes are also quite common. On an even larger scale, the association of consonants with the lexical part of the morpheme and vowels with the grammatical part seems quite obvious (see the forms given in (2)).

3.2. The phonemic inventory

During the history of the Georgian language, the phonemic inventory has not undergone many changes. During the last ten centuries, the phoneme set was reduced by three elements. The semivowel /y/ was lost, the bilabial spirant /w/ merged with one of the variants of the labio-dental /v/ and the velar stop /q/ converged with the spirant /x/, although this sound is still retained in mountainous dialects of Georgian (Vogt 1961). The latter process caused the emergence of some homonymous forms.

- | | | | |
|-----|------------------------|---------------------|----------|
| (8) | <i>Modern Georgian</i> | <i>Old Georgian</i> | |
| | xeli | qeli | ‘hand’ |
| | xeli | xeli | ‘stupid’ |
| | xerxi | qerxi | ‘skill’ |
| | xerxi | xerxi | ‘saw’ |

Since the Georgian vowel system is much simpler than the consonant system, the vowel system will be considered first.

3.2.1. The vowel system

The five vowels of Georgian can be characterised by two parameters: height and roundness.

- (9) *Vowels*
- | | | |
|----------------|---|--|
| Low | a | low, open, slightly fronted [a] (French <i>patte</i>) |
| High unrounded | i | front, spread lips, high, between close and half close [i] (English <i>bit</i>) |
| High rounded | u | same height as /i/ [u] (English <i>book</i> with marked lip rounding) |
| Mid unrounded | e | front, spread lips, between half close and half open [ɛ] (English <i>get</i>) |
| Mid rounded | o | same height as /e/ [ɔ] (German <i>Glocken</i>) |
- (Akhvlediani 1949)

Schematically, the Georgian vowel system can be presented as a triangle.

- (10) *The vowel inventory*

i	[i]	u	[u]
e	[ɛ]	o	[ɔ]
	a	[a]	

Length is not a relevant feature of the vowel system and hence does not play a role either in the phonological processes or in accent assignment.⁵ There are no diphthongs. There are some alternations involving vowels. Some of the alternations are phonologically conditioned, but most of them are morphologised, participating in ablaut in a verbal paradigm, e.g. the /e/ ~ /i/ alternation in pairs such as /da-grex-s/ ‘he will twist’ ~ /da-grix-a/ ‘he twisted’. In this case, the opposition between future and past tenses, in addition to inflectional suffixes /-s/ vs. /-a/, is expressed by the vowel alternation /e/ ~ /i/. Among the vowel alternations, the following can be listed:

⁵ Phonetic lengthening of vowels is observed before voiced fricatives and voiced stops, and especially before nasal sonorants (Akhvlediani 1949:282).

(11)		<i>NOM</i>	<i>GEN</i>		
a)	o ~ v or ø	mindor-i	mindvr-is	'field'	
		xoxob-i	xoxb-is	'pheasant'	
b)	a ~ ø	bal-i	bl-is	'cherry'	
c)	e ~ ø or i	švel-i	švl-is	'deer'	
		c'avedi	'I went'	c'avida	'he went'

The high vowels /i/ and /u/ are considered as more marked than the other vowels since they do not undergo any changes or deletion. Two distributional generalisations about the vowels /i/ and /u/ also substantiate this claim. They are given in (12).⁶

- (12) a) Although the vowel /u/ is commonly found in the final position of monosyllabic lexical words, e.g. /bu/ 'owl', /ru/ 'brook', /χ'ru/ 'deaf', its occurrence in the final position of polysyllabic lexical words is very restricted.
- b) There are no monosyllabic lexical words with final /i/. Although there are two words, /k'i/ 'yes' and /vin/ 'who', which seem to satisfy this condition, they are in fact not lexical words: the first is a confirmation particle and the second is a pronominal element. In monosyllabic words, the vowel /i/ is always a grammatical morpheme, e.g. /d-is/ (GEN) < /da/ 'sister', /č'r-i-s/ 'somebody cuts' < /č'ra/ 'to cut', etc.

These two observations suggest that the distribution of the high vowels is sensitive not only to the syllabic structure of a word (monosyllabic vs. disyllabic), but also to the asymmetry between lexical and grammatical morphemes. The vowel /u/ is preferred in lexical morphemes, while the vowel /i/ in grammatical morphemes. This morphologically motivated asymmetry between /i/ and /u/ is quite evident in grammatical affixes. The vowel /i/ is often found in inflectional affixes, while /u/ occurs only in derivational affixes, e.g. /u-/ is the particle of negation in forms such as /u-kud-o/ 'without hat', /u-namus-o/ 'shameless', and the objective version marker in forms such as /u-k'eteb-s/ 'somebody does something to somebody else'.

There is a small set of monosyllabic, monomorphemic words in Georgian. All of them are vowel-final.⁷ These words have a preference for final /u/, not /i/. This regularity is expected for acoustic and perceptual reasons. The vowel /u/ is more sonorous, and therefore more salient than /i/. This was confirmed by acoustic and perceptual studies on the Georgian vowels (Žgent'i 1956). In addition, phonetic studies on the long-distance influence of vowels on consonants reveal that the rounded vowels /o/, and especially /u/, have more influence on consonants than others; Georgian listeners could identify the trace of the deleted vowels in the cases when such vowels were labial, especially the vowel /u/ (K'iziria 1985). Thus, the

⁶ The other three vowels, /a e o/, do undergo deletion and are characterised by quite free distribution throughout the word.

⁷ In general, neither monosyllabic nor other types of non-derived words are consonant-final in Georgian.

occurrence of the vowel /u/ in prominent position, that is in final position of monomorphemic words, is more plausible than the occurrence of the vowel /i/ in the same position.

3.2.2. Hiatus

One basic phonotactic restriction on monomorphemic words is that two adjacent vowels are disallowed.⁸ Two vowels can appear across morpheme boundaries. It is interesting that the prefix vs. suffix asymmetry shows up in such cases. Hiatus is tolerated at prefix # stem boundary, while at the stem # suffix boundary it is not. The most common way of resolving hiatus across morpheme boundaries is epenthesis. In most cases, the epenthetic segment is the sonorant /v/, e.g.

(13)	<i>NOM</i>		<i>Derived word</i>	
	rje	‘milk’	me-rje-v-e	‘milkman’
	t’χ’e	‘forest’	me-t’χ’e-v-e	‘forester’
	uto	‘iron’	a-uto-v-eb-s	‘somebody irons’

The epenthetic element, i.e. the sonorant /v/, is an interesting element in its own right, since its distribution, both diachronically and synchronically, can explain several facts of Georgian phonotactics. I will return to the behaviour of the sonorant /v/ when considering the consonant system, and especially the group of sonorants.

In a few cases, an epenthetic bilabial stop /b/ is found, e.g.

(14)	ezo	‘yard’	m-ezo-b-el-i	‘neighbour’
------	-----	--------	--------------	-------------

The insertion of /v/ is prohibited, however, when the following morpheme starts with either of the rounded vowels /o/ or /u/, e.g.

(15)	šina	‘inside’	sa-šina-o	‘home, domestic’
------	------	----------	-----------	------------------

In general, all vowels except for /o/ and /u/ may be preceded or followed by any consonant: in monomorphemic words no */v/ + /o/ or */v/ + /u/ combinations are accepted. This restriction can be related to a general phonotactic restriction on adjacent bilabials in Georgian. To avoid combinations of */v/ + /o/ or */v/ + /u/ across morpheme boundaries, the sonorant /v/ is deleted.

(16)	tagv-i	‘mouse’	sa-tag-ur-i	‘mouse trap’
	tav-i	‘head’	sa-ta-ur-i	‘title’

Another strategy to avoid the hiatus across morpheme boundaries is vowel deletion.

⁸ There are a few monomorphemic loan words with vowel complexes in Georgian, e.g. /musaipi/ ‘talk’ (Arabic), /paipuri/ ‘porcelain’ (Greek), /p’aik’i/ ‘pawn’ (Persian), /daira/ ‘tambourine’ (Arabic), /maudi/ ‘cloth’ (Arabic), etc. There are also cases with vowel complexes which are the result of compounding, e.g. /č’aobi/ < /č’a-obi/ ‘swamp’, /daisi/ ‘sunset’ < /da-isi/, etc. (see Uturgaidze 1976:40).

- | | | | |
|------|------------|------------|-----------|
| (17) | <i>NOM</i> | <i>GEN</i> | |
| | jma | jm-is | ‘brother’ |
| | xe | x-is | ‘tree’ |

In general, there are no assimilation processes between vowels and vowel-consonant combinations. Literary Georgian does not allow assimilation processes, due to the strong influence of the orthographic normativity (Akhvlediani 1949). However, some assimilation processes are found in dialectal forms, e.g.

- | | | | |
|------|--------------------------|------------------------|------------------------|
| (18) | <i>Literary Georgian</i> | <i>Imeruli dialect</i> | |
| | c’a-i-y-o | c’eiyo | ‘he took something’ |
| | mo-i-t’an-a | meit’ana | ‘he brought something’ |

3.2.3. Labials

There are some phonetic processes suggesting that the opposition involving labiality is an important one in Georgian phonotactics. Sequences of rounded vowels are disallowed both in monomorphemic environments and across morpheme boundaries. A dissimilation process occurs when two rounded vowels appear adjacently in derived words.

- | | | | | |
|------|----------|---|---------|-------------|
| (19) | indo-uri | > | indauri | ‘turkey’ |
| | sa-uto-o | > | sautao | ‘iron desk’ |

The restriction on the occurrence of adjacent labials holds not only for vowel combinations, but also for vowel-consonant (see (41)) and consonant-consonant (see section 3.5) combinations. This observation suggests that there is a formal identity of labiality in consonants and roundedness in vowels.⁹

With regard to the observation on the labials, the Georgian data fit the cross-linguistic generalisation that is formulated as the Obligatory Contour Principle. That is, segments that are ‘identical’ are not permitted in a sequence (see Chapter 2).

3.2.4. Conclusions

The following generalisations summarise the paradigmatic and syntagmatic characteristics of the Georgian vowel system. There is very little assimilation between vowels and there are no diphthongs. Monomorphemically, adjacent vowels are also disallowed. The vowel alternations are either the result of morphological affixation or themselves have a morphological function.

The high vowels /i/ and /u/ have a defective distribution that is sensitive both to the number of syllables and the morphological composition of a word. They do not undergo the deletion process of Modern Georgian.

⁹ In recent years a number of researchers have argued for a (partial) identification of Place features for consonants and vowels (see Reighard 1972, Campbell 1974, Anderson & Ewen 1987, Clements 1989, Levelt 1994, van de Weijer 1996).

The bilabial correlation is one of the most important, conditioning several phonological processes across morpheme boundaries.

3.3. The consonant system

As has already been pointed out, there are 28 consonants in the phonemic inventory of Georgian. For classification, it is convenient to divide them into three major sets: (i) stops and affricates, (ii) fricatives and (iii) sonorants.

Plosives: stops and affricates. The Georgian plosives can be characterised by place of articulation (according to which bilabial, dental, alveolar, palato-alveolar and velar plosives are distinguished) and laryngeal feature (distinguishing voiced, voiceless and glottalised sounds). The Georgian plosives are presented in (20).

(20) *Stops and affricates*

	Bilabial	Dental	Alveolar	Palato-alveolar	Velar
Voiced	b	d	j	ǰ	g
Voiceless	p	t	c	č	k
Glottalised	p'	t'	c'	č'	k'

From an acoustic point of view, the Georgian voiced obstruents are characterised by a low degree of voicing. This can be confirmed by transcriptions of Georgian speech by foreign listeners. The Georgian voiced obstruents are almost always perceived as either glottalised or voiceless (Žgent'i 1956). Some phonetic studies suggest that the voiceless consonants are aspirated (Žgent'i 1956). Thus, the classification according to laryngeal feature can be as follows: non-aspirated, aspirated and glottalised. More phonetic studies need to be carried out to establish which feature is more adequate for the description of Georgian consonants, [voice] or [aspiration].

Georgian has word-final devoicing. Since no minimal word ends in a consonant in Georgian, devoicing occurs in grammatical affixes, e.g. /v-a-k'et-eb/ is pronounced as [vak'etep] 'I do something', where /b/ belongs to the thematic suffix /-eb/ and /k'ac-ad/ is pronounced as [kac-at] 'man ABL', where /-ad/ is the Ablative case marker. Devoicing does not occur when the suffixes are followed by another suffix, for instance in forms such as /v-a-k'et-eb-di/ [v-a-k'et-eb-d-i] 'I was doing', and /k'ac-ad-a/ [k'ac-ad-a] 'like a man'.

A few words must be said about the affricates, of which there are six in Georgian: /j c c'/ (alveolars) and /ǰ č č'/ (palato-alveolars). There are some cases when affricates appeared historically as a result of the merger of stops and fricatives, e.g.:

(21)	<i>Old Georgian</i>	<i>Modern Georgian</i>	
	at-švidmet'i	čvidmet'i	'seventeen'
	at-samet'i	camet'i	'thirteen'

In addition, there are cases of de-affrication.¹⁰ Evidence for this comes from dialectal forms.

(22)	<i>Literary Georgian</i>	<i>Imeruli dialect</i>	
	sc'ori	st'ori	'right'
	sc'avla	st'avla	'to study'
	dajdoma	daždoma	'to sit'
	ocdaati	ozdaati	'thirty' (Žgent'i 1956:237)

This is not to say that affricates represent mere combinations of stops and fricatives. Phonotactic restrictions discussed below indicate that affricates are single segments (see (23) and (24)). Experimental studies show that the phonetic duration of affricates is no longer than that of simple stops (Žgent'i 1956). Thus, Georgian affricates, like affricates of other languages, can be represented as segments linked to one timing slot or root node (see e.g. van de Weijer 1996).

Recently, Kehrein (1999) has proposed the elimination of the class of affricates from the phonological description and their inclusion in the natural group of stops. Even though, for the sake of convenience, the affricates are grouped together with stops in the above description, there are some phonotactic restrictions that suggest that affricates form a natural class of their own. For instance, the combination of a coronal stop followed by an affricate is not permitted in Georgian, e.g. the clusters *dc and *jd are ill-formed, while the reverse order is accepted, e.g. the clusters cd and jd are attested. There is another restriction in combination with fricatives. Affricates can be preceded by fricatives but never followed by them. For instance, the clusters sc' and šj are attested, while the clusters *c's and *jš are not. Both regularities also apply across a vowel. All these restrictions hold only within a lexical morpheme and can be violated across morpheme boundaries. The phonotactic restrictions concerning the combination of coronal and affricate consonants are depicted in (23) and those concerning the combination of fricative and affricate consonants are illustrated in (24).

- (23) a) *coronal stop + (V) + [coronal stop + fricative]_{affricate}
 b) [coronal stop + fricative]_{affricate} + (V) + coronal stop

- (24) a) *[coronal stop + fricative]_{affricate} + (V) + fricative
 b) fricative + (V) + [coronal stop + fricative]_{affricate}

In both (23) and (24), the OCP seems to be at work. In (23a), the restriction is imposed on adjacent coronal stops, and in (24a) the restriction holds for adjacent fricatives. Thus, optimal combinations are consonants with non-identical manner specification; in the case of stops, the restriction is even more specific, because it applies only to coronal stops. More such constraints, referring to place of articula-

¹⁰ The claim that de-affrication in Georgian dialects is a later development is substantiated by Old Georgian data, where corresponding segments are affricates, as in Modern Georgian.

tion of stops, are considered in the following discussion of the syntagmatic regularities of consonants. These two observations (i.e. (23) and (24)) demonstrate the relevance of distinguishing affricates as a separate class (see also Butskhrikidze & van de Weijer 2001b).

Fricatives. Fricatives can be characterised by place of articulation (by distinguishing between alveolar, palato-alveolar and velar sounds) and laryngeal feature (by distinguishing between voiced and voiceless sounds). The fricatives are presented in (25).

(25) *Fricatives*

	Alveolar	Palato-alveolar	Velar
Voiced	z	ʒ	ɣ
Voiceless	s	š	x

There are two phonemes which have not yet been considered. One is the glottalised uvular fricative /χʷ/, alternatively transcribed as a glottalised uvular stop (i.e. /qʷ/). This sound can be grouped together with the velar fricatives. It behaves like the fricative /x/ and its voiced counterpart, in that it can form a particular type of complex segments known as harmonic groups (see also below) when combined with anterior consonants. Another is the voiceless laryngeal /h/. It is the only laryngeal phoneme in Georgian. The distribution of /h/ is restricted to word-initial position, but even in this position it tends to delete in Modern Georgian. The following correspondences exemplify this:

(26)	<i>Old Georgian</i>	<i>Modern Georgian</i>	
	hazri	azri	‘mind’
	hambavi	ambavi	‘story’
	hasak’i	asak’i	‘age’
	hambor	ambori	‘hug’

Most of the words beginning with the consonant /h/ are loan words, e.g. from Greek: /hipoteza/ ‘hypothesis’, /harmonia/ ‘harmony’, etc. This sound never occurs in consonant sequences. Because of its defective distribution, /h/ will not be considered in the following discussion of consonant phonotactics.

Sonorants. There are 5 sonorants: the liquids /r l/, the nasals /n m/ and the labio-velar /v/. The sonorant /v/ has the allophones [v ɸ w], depending on the phonetic context:

- (27) a) /v/ has the allophone [v] in word-initial position and between vowels, e.g. /viri/ [viri] ‘donkey’, /venaxi/ [venaxi] ‘wine-yard’, /tavi/ [tavi] ‘head’, /telavi/ [telavi] ‘the name of a town’.
- b) /v/ has the allophone [ϕ] when followed by voiceless (aspirated or glottalised) consonants, e.g. /v-pikr-ob/ [ϕpikrop] ‘I think’, /v-c’er/ [ϕc’er] ‘I write’.
- c) /v/ has the allophone [w] when preceded by an obstruent, especially by dorsal consonants, e.g. /χ’vavili/ [χ’wavili] ‘flower’, /kva/ [kwa] ‘stone’. The allophone [w] in such words is actually pronounced as a secondary articulation on the preceding consonants. Thus, more precise phonetic transcriptions of the words are [χ^wavili] and [k^wa], respectively.

The characterisation of /v/ as a sonorant can be established on the basis of its distribution and behaviour in phonological processes. They are discussed in section 3.3.1.

In languages with long consonant sequences sonorants are often observed to be syllabic. Investigating the syllabic status of Georgian sonorants is an important issue for an adequate analysis of long consonant sequences. Consequently, the following section addresses the status of sonorants. The behaviour of the sonorant /v/ is considered extensively in a separate section. Two claims are made:

- (28) a) Georgian sonorants are phonetically syllabic in consonant sequences. The claim is based on phonological processes, distributional patterns, and historical and comparative evidence. This especially concerns the most sonorous consonant /t/.
- b) Combinations of C + /v/ can be treated as complex labialised segments, i.e. C^w. The claim is based on phonological, distributional and historical evidence. Reduplication data also substantiate this claim (see Chapter 5).

3.3.1. Syllabicity of sonorants

Akhvlediani (1949) assumes the phonetic syllabicity of the Modern Georgian sonorants. He argues that the syllabicity of the sonorants is context-dependent. It is realised only in clusters when the sonorant is surrounded by less sonorant consonants, e.g.

(29)	NOM	DAT		
	naym-i	naym-s	[nayms]	‘mine’
	ipn-i	ipn-s	[ipɲs]	‘ash tree’
	saxl-i	saxl-s	[saxɭs]	‘house’
	tetr-i	tetr-s	[tetɾs]	‘white’

The forms which demonstrate the syllabicity of the sonorants contain clusters in derived word-final position (forms in the Dative case). I argue that the syllabicity of

sonorants can also be assumed in consonant sequences occurring in initial position in monomorphemic words. Thus, the sonorant /l/ in /k'ɫde/ [k'ɫde] 'rock' is phonetically syllabic, as is the sonorant /r/ in /trtvili/ [trtvili] 'hoar-frost'. To strengthen this claim, historical and comparative evidence as well as synchronic phonological processes and distributional regularities will be considered.

In their study of the correspondences between four languages of the Kartvelian language group and their investigation of the historical development of these languages, Gamkrelidze & Machavariani (1965) propose that Proto-Kartvelian sonorants are syllabic in certain phonetic contexts.¹¹ An explanation of the regular correspondences between the Kartvelian languages, ablaut in the verbal system and present-day syncope in the nominal paradigms in Modern Georgian is related to the syllabicity of sonorant consonants. Below I give some examples of Proto-Kartvelian syllabic sonorants, with correspondences in Kartvelian languages. The examples are taken from Gamkrelidze & Machavariani (1965).

(30)	<i>Proto-Kartv.</i>	<i>Georgian</i>	<i>Megrelian</i>	<i>Laz</i>	<i>Svan</i>	
	tɾt	trt-ol-a	tirt-ol-i	tirt-ini		'trembling'
	k'ɫde	k'ɫde	k'irda/k'irde		k'oǰ	'rock'
	m̥ze	mze			miž/mež	'sun'

I will demonstrate that sonorants of Modern Georgian have retained their syllabicity in words like *trtola*, *k'ɫde* and *mze* (see e.g. (39) for *trtola* and (31) for *k'ɫde*).

Phonetically syllabic [ɾ] and [ɫ] are assumed to be present in the mountain dialects and some lowland dialects of Georgian (Gamkrelidze & Machavariani 1965). The following correspondences are considered to be illustrative:

(31)	<i>Old Georgian</i>	<i>Modern Georgian</i>	<i>Khevsuruli</i>	<i>Ingilouri</i>	
	rgoli	rgoli	girgoli		'ring'
	grk'ali	rk'ali	girk'ali	k'irk'ali	'arc'
		ɣrma	ɣirma		'deep'
		trtvili	tvirtvili	tirtul	'hoar-frost'
		črdili	čirdili		'shadow'
		k'ɫde	k'ilde		'rock'

The sonorants /r/ and /l/ seem to be syllabic in both Old Georgian and Modern Georgian. The corresponding Khevsuruli and Ingilouri forms with the vowel /i/ before the sonorants /r/ and /l/ substantiate this claim.

There are some phonological processes and distributional patterns which substantiate the claim about the syllabicity of Georgian sonorants, especially of the sonorant /r/. Consider in this respect the following: "The syllabicity of syllabic consonants never arises spontaneously from a marginal consonant, as far as I can ascertain. The source of the syllabicity is always a vowel. By far the commonest process

¹¹ Asatiani (1983) suggests that stops and fricatives are also syllabic in Proto-Kartvelian.

of origin is syllabic syncope, that is, loss of a vowel and shift of syllabicity to one of its margins” (Bell 1978:105).

Georgian has some phonological processes in which vowel loss is related to immediately following sonorants. One of these processes is syncope.¹² It occurs in the Modern Georgian nominal paradigm. Words with the stem-final syllables *-al*, *-ar*, *-el*, *-er*, *-an*, *-en*, *-am*, *-em*, *-ol*, *-or*, *-on*¹³ undergo vowel deletion when followed by vowel-initial suffixes (case markers, the plural suffix, etc.). One of the requirements for syncope to occur is that a word must contain at least two syllables. Some examples of syncope in the nominal paradigm are given in (32).

(32)	NOM	mercxal-i ¹⁴	‘swallow’
	ERG	mercxal-ma	
	DAT	mercxal-s	
	GEN	mercxl-is	
	INST	mercxl-it	
	ABL	mercxl-ad	

As shown in (32), syncope occurs in three cases: the genitive, instrumental and ablative. In all of these cases suffixes are of the -VC type: */-is/*, */-it/* and */-ad/*, respectively.

Syncope does not occur when the stem does not end in a sonorant. Consider the case in ((33)).

(33)	NOM	k’amat-i	‘debate’
	ERG	k’amat-ma	
	DAT	k’amat-s	
	GEN	k’amat-is	
	INST	k’amat-it	
	ABL	k’amat-ad	

Another phonological process is truncation, which also occurs in the nominal paradigm. Vowels are lost in words ending with vowels other than */i/* (the nominative case marker) and */u/*. In case words have two final syllables containing sonorants, both processes (truncation and syncope) apply simultaneously, e.g. */t’omara/* ‘sack’ in GEN */t’omr-is/*, INST */t’omr-it/*, ABL */t’omr-ad/*. In these cases, two vowels are deleted, the stem-final */a/* and the stem-medial */a/*.

¹² Syncope not only occurs in nominal paradigms (e.g. in the GEN, ABL and INST cases), but also with other types of suffixes with a -VC type structure, for example when the plural suffix */-eb/* is added to a stem, e.g. */bal-i/* ‘cherry’ ~ */bl-eb-i/* ‘cherries’.

¹³ There are some exceptional cases when syncope occurs in words with a stem-final obstruent */b/*, e.g. */k’ak’ab-i/* ‘partridge’ ~ */k’ak’b-is/* (GEN) and */xoxob-i/* ‘pheasant’ ~ */xoxb-is/* (GEN).

¹⁴ The nominative case marker */-i/* always follows consonant-final stems and never causes the reduction of stem-final vowels. In this respect, the nominative case marker is different from other vowel-initial suffixes.

In both the processes described above, a vowel deletes only when it is followed by a sonorant. The syncope process is also characteristic of verbal forms.

- (34) *še-i-p'χ'ar-i* 'arrest IMP' *še-i-p'χ'r-ob* 'you will arrest'
 ga-č'er-i 'cut IMP' *ga-č'r-i* 'you will cut'

The vowel-deletion process, which takes place under adjacency to sonorants, indicates the syllabicity of Georgian sonorants.

There are interesting distributional characteristics of the sonorants observed in clusters. The co-occurrence of the sonorants seems sensitive to the number of obstruents in a cluster. Two observations can be made about the co-occurrence of the sonorant /v/ with other sonorants, depending on the length and constituency of a cluster:

- (35) a) In a cluster CS_1S_2 (obstruent + sonorant + sonorant), S_1 is always the sonorant /v/, while S_2 is either /n/, /l/ or /r/, e.g. /k'vn/ in /k'vnesa/ 'to moan', /k'vn/ in /kvnet'a/ 'to bite', /gvr/ in /gvrit'i/ 'turtle-dove', /c'vr/ in /c'vrili/ 'petty, small', /c'vl/ in /c'vlili/ 'mite', /xvr/ in /xvreli/ 'hole'. Note that in most of the cases obstruents are dorsals, which, in general, are the best landing sites for labialised consonants (Maddieson 1984). The generalisation is related to the claim that the sonorant /v/ is a secondary articulation on the preceding consonant. I return to this issue by considering extensive data in the following section.
- b) In clusters with two or more obstruents (e.g. $C_1S_1C_2S_2$), the order of sonorants is different. The most sonorous sonorant /r/ takes the first position and the sonorant /v/ can only be the second sonorant of the cluster, e.g. /drt'v/ in /drt'vinva/ 'to grumble', /grgv/ in /grgvinva/ 'thunder' and /brč'χ'v/ in /brč'χ'viali/ 'sparkling'.

With regard to these generalisations, it is interesting to examine the combinations where C_1 is filled by a harmonic group or an /s/ + obstruent cluster. Both combinations count as one obstruent for these generalisations, e.g. /pxvn/ in /pxvnili/ 'powder', /sxvl/ in /sxvla/ 'chop off'. These patterns are in accordance with the generalisation formulated in (35a). Thus, these clusters could be formalised as CS_1S_2 . I will discuss both types of clusters in more detail later.

From the case of sonorant distribution generalised in (35b), one could argue that a cluster of the type $C_1S_1C_2S_2$ cannot be analysed as one unit, but instead is divisible into sub-parts, as follows: C_1S_1/C_2S_2 . The integrity of long sequences of the type $C_1S_1C_2S_2$ is challenged since the sonorant /r/ occupies the S_1 position and seems to be syllabic in this context. I will attempt a better formalisation of this assumption in Chapter 6, and for the time being, merely mention that for the examples given in (35b), the division would be as follows: *dr/t'vinva*, *gr/gvinva*, *br/c'χ'viali*.

Apart from the syllabicity of /r/, both generalisations (35a) and (35b) suggest the treatment of harmonic clusters and /s/ + obstruent clusters as one element C

(consonant). I return to these consonant sequences in the following sections and in Chapter 5. These generalisations also clearly illustrate that the sonorant /v/ occurs in consonant sequences predominantly after a dorsal consonant and can be treated as a secondary articulation on the dorsal consonant. The behaviour of the sonorant /v/ is discussed in the following section.

In addition, the sonorant /r/ displays interesting phonotactic behaviour in consonant sequences. For instance, sequences of the stop + fricative type are never separated by a sonorant, e.g. sequences such as **brz*, **pls*, **glz* are not attested, whereas sequences of the type fricative + stop are always separated by the sonorant /r/. Thus, sequences such *zrd* in /zrda/ ‘to grow’ and in /zrdiloba/ ‘politeness’ and *xrt* in /xrt’ili/ ‘gristle’ are commonly found in Georgian (Nebieridze 1975). It seems that sequences violating the SSP, i.e. clusters of the fricative + stop type, are separated by syllabic /r/, e.g. /r/ in /xrt’ili/ is syllabic (see also the examples in (37)).

The behaviour of the sonorant /r/ is quite complex in Georgian and I will consider it in detail.

The sonorant /r/ can precede and follow any consonant or vowel. It is described as a ‘fleeting’ sonorant (Vogt 1961). In other words, it can appear optionally when it is surrounded by consonants with identical laryngeal specifications.

- | | | | | |
|------|----------|---|---------|---------|
| (36) | prta | ~ | pta | ‘wing’ |
| | grdemli | ~ | gdemli | ‘anvil’ |
| | brjeni | ~ | bjeni | ‘wise’ |
| | brjaneba | ~ | bjaneba | ‘order’ |

As shown in the examples in (36), the sonorant /r/ can disappear when preceded and followed by consonants with identical laryngeal specifications.

‘/r/-loss’ does not take place when /r/ is preceded and followed by consonants which do not have the same laryngeal specification. Thus, the presence of the sonorant /r/ is obligatory in the following forms.

- | | | |
|------|-------------------------|-------------|
| (37) | brč’χ’ali ¹⁵ | ‘claw’ |
| | k’rjalva | ‘reverence’ |

In addition, the presence of the sonorant /r/ is obligatory when it is surrounded by identical consonants, e.g. in forms such as:

- | | | |
|------|----------|--------------|
| (38) | trtvili | ‘hoar-frost’ |
| | grgvinva | ‘thunder’ |

According to Ertelishvili (1970), stems with the sonorant /r/ between identical consonants are always derived from reduplicated forms.

¹⁵ ‘/r/-loss’ is attested in dialectal forms, but only after assimilation has taken place: /brč’χ’ali/ > /p’rč’χ’ali/ > /p’č’χ’ali/ ‘claw’.

(39)	<i>Stems with /r/</i>			
		<i>Literary Georgian</i>		<i>Dialectal form</i>
	Stems	C_1rC_1		CVrCVC
	trt	trt-ol-a 'to tremble'		tar-tal-eb-s 'tremble'
		<i>Literary Georgian</i>		<i>Paradigmatically related form</i>
	prp	prp-en-a 'to adore'		par-pat'i 'to fly around'
				pr-en-a 'to fly'

The examples given in (39) suggest that the sonorant /r/ in (38) is syllabic.

The phonological and phonotactic properties of the Georgian sonorants, especially those of the sonorant /r/, suggest that they should be considered as syllabic consonants. However, the syllabicity of Georgian sonorants has to be substantiated by phonetic studies.

3.3.2. The sonorant /v/ as a secondary articulation

Nepveu (1994) and Bush (1997), after observing that many of the clusters in Georgian contain /v/, and that it varies in pronunciation between [v], [ϕ] and [w], suggest that it is a 'defective segment'. Nepveu argues that it is specified only for labial place of articulation, and acquires other features from the preceding consonant. It can be treated as dependent segment, i.e. a secondary articulation on the preceding consonant. I accept the analysis of the sonorant /v/ as a secondary articulation in consonant sequences and substantiate the claim with additional data from phonological processes, distributional regularities and historical evidence, and with reduplication data presented in Chapter 5.

Phonological processes involving the sonorant /v/ include metathesis, which is extensively discussed in Butskhrikidze & van de Weijer (2001a). Here I merely give a short introduction to the metathesis process relevant for the present discussion of the status of /v/. /v/, which is part of the thematic suffix in verb forms, violates integrity of the root when followed by the infinitival suffix /-a/. For example, the third person singular of the root /xar/ 'to gnaw' is /xr-av-s/ (with deletion of the root vowel). The infinitival form, however, is /xvr-a/ (with deletion of both the root vowel and the vowel in the thematic suffix), where the thematic suffix consonant /v/ occurs between the two root consonants.

(40) *Regular metathesis in verb forms*

ROOT	PRES 3 rd SG (/-av/ THEM SUFF)	INF (/-a/ INF SUFF)	
xar	xr-av-s	xvr-a	‘to gnaw’
k’ar	k’r-av-s	k’vr-a	‘to tie’
xan	xn-av-s	xvn-a	‘to plough’
k’al	k’l-av-s	k’vl-a	‘to kill’
sxal	sxl-av-s	sxvl-a	‘to chop off’
jer	jr-av-s	jvr-a	‘to move’

In addition, there are some isolated forms displaying metathesis. There are cases when metathesis has occurred as a result of a diachronic change, e.g. /rva/ ‘eight’ > /at-rvamet’i/ ‘eighteen’ (Old Georgian) > /tvramet’i/ (Modern Georgian). There are cases of metathesis in dialectal forms too, e.g. /gač’irveba/ ‘difficulty’ (Literary Georgian) ~ /gač’ivreba/ (Imeruli dialect).

There are a few striking conditions on /v/-metathesis: the root must end in a sonorant other than /m/ (e.g. metathesis does not occur with a root like /xed/ ‘to see’) and the root should not start with a labial (metathesis does not occur with a root like /ber/ ‘to blow up’). The condition can be formalised as follows: C /r l n/ /v/ > C /v/ /r l n/. The condition is related to the constraint which bans two bilabials monomorphemically. I will return to this constraint later. It is interesting to observe that consonants preceding /v/ are predominantly dorsal (see (40)). Dorsal consonants are the best landing sites for labialisation cross-linguistically (Maddieson 1984).

Note that the metathesis process just described is closely related to the stem-vowel deletion process. One of the factors preventing vowel deletion in general in Georgian is avoidance of homonymous forms, e.g. in the form /k’erv-a/ ‘to sew’. Vowel deletion and consequently metathesis do not occur, because homonymous forms * /k’vra/ ‘to sew’ and /k’vra/ ‘to bind’ would emerge. Thus, even though all conditions for the metathesis process are met in the form /k’erv-a/, it does not take place, because the language avoids creating homonymous forms.¹⁶

There are a number of other processes affecting the sound /v/, e.g. /v/-loss; /v/-alternation with /o/; /v/-epenthesis, etc. They are discussed briefly below.

/v/-loss: stem-final /v/ is lost when followed by affix beginning with the round vowels /o/ or /u/. Consider the paradigmatically related forms in (41).

(41)	NOM		Derived word		
	tav-i	‘head’	meta-ur-i	(*me-tav-ur-i)	‘leader’
	p’at’iv-i	‘honour’	p’at’i-osan-i	(*p’at’ivosan-i)	‘honest’

¹⁶ The same constraint holds for the syncope process discussed in section 3.3.1. For instance, the genitive form of the noun /kari/ ‘wind’ is /kar-is/. Stem-vowel deletion is expected, but does not occur, because there already exists the form /kris/, which means ‘blows’. Thus, the stem vowel is retained in the form /kar-is/ to avoid homonymy.

‘/v/-loss’ under adjacency to labialised vowels can be related to ‘/v/-loss’ when followed by the bilabial nasal /m/ (e.g. /gamo-tkv-am-s/ ‘somebody is pronouncing’, but /gamo-tkma/ ‘pronunciation’), and to the distributional restriction on this sonorant with regard to the bilabial stops. The following discussion will demonstrate that combinations of bilabial consonants with /v/ are disallowed. It is interesting to note that, similarly to the case of hiatus, there is asymmetry between prefix # stem and stem # suffix contexts. While monomorphemic and stem # suffix environments do not tolerate two adjacent bilabials, the prefix # stem context allows such combinations, e.g. in form such as /v-mušaob/ ‘I am working’ and /v-mecadineob/ ‘I am studying’. Thus, the combination of *vm* is allowed across prefix # stem boundary; while in other contexts it is unattested.

/v/-alternation with /o/: the substitution of the vowel /o/ with /v/ occurs in derived words of the following type:

(42)	<i>NOM</i>	<i>GEN</i>	
	mindor-i	mindvr-is	‘field’
	p’amidor-i	p’amidvr-is	‘tomato’
	nior-i	nivr-is	‘garlic’
	nigoz-i	nigvz-is	‘nut’

The substitution of the /o/ by /v/ does not occur when the /o/ is preceded by labial consonant (recall that labial + /v/ is an ill-formed cluster). Hence one can find /sap’oni/ ‘soap’ ~ /sap’nis/, but not /*sap’vnis/. For the same reason, the substitution of the vowel /o/ by the sonorant /v/ is blocked in case it is followed by the bilabial sonorant /m/, e.g. in /diyomi/ ‘toponym’ ~ /diyomis/ (GEN).

/v/-epenthesis: to resolve hiatus, the sonorant /v/ is inserted between successive vowels, e.g. /jado/ ‘witchcraft’ ~ /a-jado-v-eb-s/ ‘somebody does witchcraft’.

There is some historical evidence for /v/ serving as a secondary articulation on consonants, especially on dorsals. According to some studies (e.g. Gamkrelidze & Machavariani 1965), there was a distinction between syllabic /u/ and non-syllabic /u̥/ in Proto-Kartvelian. The merger of the non-syllabic /u̥/ with the sonorant /v/ happened in a later period of the development of Georgian. The Modern Georgian sonorant /v/ corresponds to Old Georgian non-syllabic /u̥/. The correspondences are illustrated in the examples given in (43).

(43)	<i>Old Georgian</i>	<i>Modern Georgian</i>		
	sik’u̥dili	sik’vdili	[sik’ ^w dili]	‘death’
	eku̥si	ekvsi	[ek ^w si]	‘six’
	varsk’u̥lavi	varsk’vlavi	[varsk’ ^w lavi]	‘star’
	čūen	čven	[č ^w en]	‘us’
	k’u̥ali	k’vali	[k’ ^w ali]	‘trace’

As shown in the examples in (43), just as in those in (40), [w] appears after dorsal consonants in almost all cases, and it seems, is a cognate to the Old Georgian /u̥/.

The important consequence of establishing the status of /v/ as a secondary articulation in consonant sequences is that the combination of C + /v/ can be described not as a true cluster, but as a complex segment, i.e. a labialised consonant, the sequence of a consonant followed by [^v].

3.4. The minimal word

The definition of the minimal word is language-specific. An interesting property of the Georgian minimal word is its morphological constituency. To use the term *monomorphemic word* for Georgian is not especially insightful, because only vowel-final stems are monomorphemic, whereas consonant-final stems require a vowel (the nominative case marker *i*) to give a well-formed minimal word. Thus, in Georgian, words have two types of morphological structure: (i) word = stem, e.g. /da/ ‘sister’, /k'alata/ ‘basket’ and (ii) word = stem + the nominative case marker, e.g. /saxl-i/ ‘house’, /xel-i/ ‘hand’. Morphological constituency plays an important role in the phonotactic organisation of a word. In the following sections I consider the word-initial and word-final positions in relation to the morphological and phonotactic structure of the word.

In many languages, lexical words of only one mora or syllable are avoided: a minimal bimoraic/disyllabic requirement is imposed (McCarthy & Prince 1986b, Kenstowicz 1994, among others). There are several types of evidence for such a disyllabic minimality constraint in Georgian.

Georgian has only vowel-final monosyllabic words of the CV, CCV, CCCV type, e.g. /xe/ ‘tree’, /rk'o/ ‘acorn’, /brge/ ‘tall’, etc.); there are no lexical words of the CVC type. Consonant-final stems always receive a suffixal vowel to form a well-formed minimal word, as in /k'ac-i/ ‘man’, /xel-i/ ‘hand’, etc. Vowel-final stems do not have a nominative case marker. Thus, they surface as bare stems, e.g. /magida/ ‘table’, /deda/ ‘mother’, /taro/ ‘shelf’, /sok'o/ ‘mushroom’. The generalisation that words of the CVC type do not exist in Georgian indicates that the minimal Georgian word is of the CVCV type, i.e. disyllabic. This observation is interesting with regard to the maximum number of consonants in consonant sequences. As I argue later in this chapter, ‘true clusters’ in Georgian are maximally biconsonantal.

Since the minimal word is by definition (McCarthy & Prince 1986b) coextensive with a foot, another argument supporting the proposal of disyllabic minimality in Georgian could be stress assignment. According to Tevdoradze (1978), primary stress always falls on the first syllable of a word. In polysyllabic words, secondary stress occurs. In four-syllable words, secondary stress falls on the second syllable from the end, in five-syllable words it occurs on the third syllable from the end, and in six-syllable words on the fourth syllable from the end. According to this observation, a trochaic foot can be formed in polysyllabic words. That is to say, the constituent responsible for stress assignment in Georgian is a disyllabic trochaic foot. This could be formalised as follows:

(44)	a)	($\acute{\sigma}$ σ)	déda	‘mother’
	b)	($\acute{\sigma}$ σ) σ	déda-li	‘hen’
	c)	($\acute{\sigma}$ σ)($\grave{\sigma}$ σ)	déda-èna	‘mother tongue’
	d)	($\acute{\sigma}$ σ)($\grave{\sigma}$ σ) σ	déda-švil-oba	‘motherly’
	e)	($\acute{\sigma}$ σ)($\grave{\sigma}$ σ)(σ σ)	déda-búd-ian-ad	‘with all the family’

The stress patterns in (44) illustrate that words containing more than two syllables are morphologically complex. Monomorphemic trisyllabic and longer words of Georgian origin are very rare. These are mostly loan words, e.g. /ko'nst'it'ucia/ ‘constitution’, /p'arlamet'i/ ‘parliament’, etc. This observation was confirmed by studies on the phonotactic structure of Georgian nominal and verbal stems (Ertelishvili 1970, 1980).

Additional evidence for the size of the minimal word comes from morphophonological processes. I will consider the formation of inalienable constructions in Georgian. A brief introduction to this matter is in order. “The prototypical morphological operation is affixation to a base. In most cases affixation occurs without regard to the phonological nature of the base. Of course, once affixation takes place, phonological rules can come into play. But in general, affixation occurs earlier and the phonology is left with the task of assigning a phonetic representation consistent with the rules and constraints of a language to the result” (Kenstowicz 1994:622).

However, there are cases in which the affixation process itself must take account of the phonology of the base. If the required structure does not obtain, affixation fails to occur. This is the case with the inalienable construction in Georgian.

Only kinship terms participate in inalienable possessive constructions. To form the construction, the terms should meet the following two conditions: they should be at least disyllabic, and they should have the vowel /a/ stem-finally. The latter is a general characteristic feature of Georgian kinship terms. The lexical items able to form the inalienable construction are as follows: /mama/ ‘father’, /deda/ ‘mother’, /bija/ ‘uncle’, /deida/ ‘aunt’, /mamida/ ‘sister of father’, /bicola/ ‘uncle’s wife’, /bebia/ ‘grandmother’, /babua/ ‘grandfather’. A sample paradigm of inalienable possessive constructions is given below.

(45)	<i>SG</i>			<i>PL</i>	
	mama-čemi	‘my father’		mama-čveni	‘our father’
	mama-šeni	‘your father’		mama-tkveni	‘your father’
	mama-misi	‘her/his father’		mati mama	‘their father’

The inalienable constructions are forms with the first person and second person singular and plural cases. An exception occurs in the third person plural, where /*mama-mati/ is ill-formed and the correct form is /mati mama/, with the pronoun preceding the noun and occurring separately. The same happens with other kinship terms that allow the inalienable possessive construction, such as /mati deda/ ‘their mother’, /mati bebia/ ‘their grandmother’, etc.

Monosyllabic kinship terms such as /da/ ‘sister’ and /jma/ ‘brother’ fail to allow the inalienable possessive construction. They take possessive pronouns independently and precede them, e.g. /čemi da/ ‘my sister’, /čemi jma/ ‘my brother’. Forms such as /*da-čemi/ and /*jma-čemi/ are ill-formed.

In one of the dialects of Georgian, Imeruli, where the same terms are disyllabic (/daia/ ‘sister’; /jamia/ ‘brother’), affixation applies and the constructions /daia-čemi/ and /jamia-čemi/ are well-formed (Butskhrikidze 1995). The blocking of the affixation process therefore seems to be conditioned by the disyllabic requirement of the base.

The process of monosyllabic lengthening provides further evidence for the disyllabicity of minimal words. In yes-no questions, whenever the last word of the question is monosyllabic, it is lengthened to two syllables by repeating the vowel. The process occurs if and only if the word in question is monosyllabic.

(46) *Monosyllabic lengthening*

- a) k’art’opils pckvnis? ‘Is he peeling potatoes?’
[prckniis]
- b) k’art’opili dač’ra? ‘Did he cut potatoes?’
[dač’ra]

In the example in (46a), lengthening of the monosyllabic word /pckvnis/ occurs, and it is pronounced with two *i*’s [prckniis], while in the example in (46b), lengthening does not occur, because the word /dač’ra/ is disyllabic. For more about the process, see Bush (1997). Finally, the reduplication process discussed in Chapter 5 also substantiates the claim that the minimal word is disyllabic in Georgian.

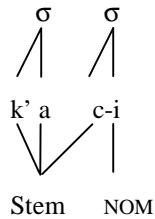
To conclude, the evidence discussed above suggests the existence of a disyllabic minimal word constraint in Georgian.

3.4.1. Word-final position

As has already been mentioned, a well-formed minimal word in Georgian must be vowel-final. The final vowel is either the nominative case marker or part of the stem. The occurrence of a single consonant or a consonant sequence is disallowed in word-final position. There are some exceptions in adverbs, most of which are sonorant-final (/c’in/ ‘in front’, /xval/ ‘tomorrow’, /gušin/ ‘yesterday’), or have voiceless obstruents (stops or fricatives) in final position (e.g. /zevit/ ‘up’, /kvevit/ ‘below’, /k’argad/ [k’argat] ‘well’). In the latter case, /-ad/ [at] is the ablative case marker and since we are not dealing with derived forms and their phonotactic patterns these cases will not be considered in the thesis.

As for the structural characteristics of the Georgian stem, the most important observation is that stems are predominantly consonant-final (Appendix 1 provides extensive data on the structural patterns of the nominal and verbal forms of Georgian). Syllable boundaries are not aligned with morphological boundaries in Georgian, e.g. /k’ac-i/ ‘man’ is syllabified as /k’a.ci/. The syllabic and morphological constituency of this word is depicted in (47):

(47) /k'aci/ 'man'



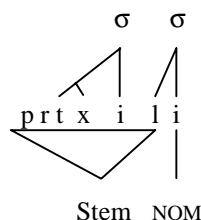
The schema in (47) is a representation of the mismatch between syllable and morphological boundaries. It also illustrates that consonantal material is within the stem domain. Notice that the relation that might hold between the consonants cannot be expressed in terms of syllables because the syllable boundary locates these consonants (i.e. the consonants /k'/ and /c/) in different constituents. In that case, the formulation of additional constraints operating on the adjacent constituents would be needed. The mismatch depicted in (47) has direct consequences for the possible types of analyses of the consonantal patterns of a language. Two options can be explored. One is an analysis which accounts for the study of consonantal patterns within the stem domain. The other is an analysis which accounts for the study of consonantal patterns in terms of a syllable constituent. The former is the one proposed and advocated in this thesis. The analysis is called the Gradual Consonant Analysis, and is extensively discussed in Chapter 6. The latter option has been considered in previous analyses of Georgian consonant clusters, and is summarised and discussed in Chapter 7. As for the stem, it can end in any of the 27 consonants (i.e. except /h/) and maximally form a five-member sequence. The patterns of stem-final consonant sequences are summarised in Appendix 3.

3.4.2. Word-initial position

All 33 phonemes of Georgian can appear in word-initial position. Words of Georgian origin are almost always consonant-initial. There are some words with a vowel in initial position, but most of them are loan words (mostly from Persian or Latin). Unlike word-final position, word-initial position is characterised by long consonant clustering, with sequences of up to six members.

Word-initial position corresponds to stem-initial position in Georgian (see section 3.5 for exceptional cases), while word-initial position does not correspond to syllable-initial position. The mismatch is illustrated by way of the word /prtʰxili/ 'careful' and is formalised in the schema in (48).

(48) /prtxili/ ‘careful’



The schema in (48) shows that in /prtxili/ the first two consonants, /p/ and /r/, cannot be syllabified together with /t/ and /x/, since the syllable-initial position (onset) is already filled by two consonants (/t/ and /x/).¹⁷ Thus, the word cannot be exhaustively syllabified. We are left with two word-initial consonants (/p/ and /r/) which are not part of the syllable. Nevertheless, the whole *prtxil* sequence is part of a structural constituent stem that encompasses the four-member cluster as a whole. The following questions arise with respect to consonant sequences of this type:

- (49) a) Are there constraints on consonant combinations in such long consonant sequences, or is their constituency random?
 b) If the co-occurrence is not random, then what are the principles governing them?

These are questions I will address in the remainder of the thesis. However, for the present, it is important to keep in mind the domain mismatches depicted in (47) and (48).

3.4.3. Conclusions

To conclude, two important observations have been made with respect to the characteristics of the Georgian minimal word.

The disyllabicity of the minimal word is substantiated by phonological (accent assignment), phonetic (monosyllabic lengthening) and morphological (the alienable/inalienable construction) evidence. Additional evidence will come from the reduplication process discussed in Chapter 5. The Georgian minimal word can be formally presented as a general disyllabic template of the $C_1V_1C_2V_2$ type.

Syllable-final position coincides with word-final position, while syllable-initial position does not coincide with word-initial position. The mismatch can be related to the fact that the morphological boundary is not aligned with the syllable boundary in Georgian. Consequently, the phonotactics of Georgian can be better interpreted not only by looking at syllable constituency, but by also taking into account the mor-

¹⁷ At this point, the optional occurrence of the sonorant /r/ and the treatment of harmonic clusters (here *cx*) as complex segments are disregarded. These will be considered later, in the analysis of consonant sequences (see Chapter 6).

phological constituency of a word. Stem-initial position in Georgian corresponds with word-initial position, while stem-final position does not correspond with word-final position (except vowel-final stems). This is exactly the opposite situation to the one described above, i.e. the syllable/word correlation. In the following section, I discuss word-initial, i.e. coinciding with stem-initial, consonant combinations.

3.5. General introduction to consonant syntagmatics

Apart from the study of consonant grouping that was carried out in section 3.3, another classification referring to the place of articulation is needed to account for the syntagmatic patterns of consonant combinations. Thus, we arrive at the table in (50), in which primary focus is on place and manner of articulation.

(50) *Consonants grouped according to place of articulation and manner*

Manner	Stop			Nasal	Fricative						Liquid
Place											
Bilabial	b	p	p'	m							
Alveolar	d	t	t'	n	j	ʃ	c'	z	s	l	r
Palato-alveolar					c	č	č'	ʒ	š		
Velar	g	k	k'					ɣ	x		
Uvular									χ'		

There are many co-occurrence restrictions of the OCP type in Georgian, referring primarily to place of articulation.

Geminates are disallowed in monomorphemic contexts.¹⁸ The classification discussed below is based on Uturgaidze (1976). It accounts for the majority of the Georgian data.

The first restriction concerns the labial sounds (including the labio-dental sonorant /v/). This class is denoted as C1. It contains the sounds /b p p' m v/.

It is not permissible to have two adjacent labial consonants in a monomorphemic environment. Any combination between the members of this class is disallowed. Thus, *mb-, *mp-, *mp'-, *mv-, *bm-, *bp-, *bp'-, *bv-, *pm-, *pb-, *pv-, *vm- (recall that /v/ is lost when followed by /m/), etc. are not found. Recall that combinations of rounded vowels, as well as rounded vowels and the sonorant /v/, are also disallowed (see sections 3.2.2 and 3.2.3).

The second class, C2, contains the coronals, /d t t' j c c' z s ʃ č č' ʒ š/, in which three places of articulation are distinguished:

¹⁸ There are no geminates in Georgian. However, identical consonants do occur across morpheme boundaries, e.g. /xaz-ze/ 'on the line', /v-varjišob/ 'I am training', /mat-tan/ 'with them'. In order to avoid gemination, one of the consonants undergoes deletion, e.g. /t'ani-samosi/ /*t'anis-samosi/ 'cloth', etc. Foreign words with geminates always enter the Georgian lexicon with only one consonant, e.g. /alegoria/ 'allegory', /k'lasi/ 'class', etc.

Dentals: /d t t' /
 Alveolars: /j c c' z s /
 Palato-alveolars: /j̣ č č' ž š /

C2 has two restrictions on the co-occurrence of its members:

a) Combinations of homorganic phonemes are not accepted (i.e. the same restriction as for labials);

b) A posterior coronal may precede an anterior coronal, but never follow it. Thus, the combinations *št*, *jd*, *cd* are attested, while **tš* **dj*, **dc* never occur.

The third class, C3, contains the sonorants, /r l n/. It has the following restriction: the members of the class never combine (i.e. the same restriction as for the first and the second classes).

The velar and uvular consonants are grouped together. This class is referred to as C4, and contains /g k k' ɣ x χ' h/. It has the following restriction: the members of the class never combine with each other (i.e. the same restriction as for the first, the second and the third classes).

To conclude, the generalisation is as follows:

(51) Obstruents with identical place of articulation never combine.¹⁹

Two additional factors are of importance when consonant sequences are discussed.

(52) a) The places of articulation of the members of the consonant sequence.
 b) The laryngeal specification.

With regard to the first factor, regressive and non-regressive consonant sequences can be distinguished. 'Regressive' refers to anterior + posterior consonant sequences (i.e. labial-coronal, coronal-dorsal, labial-dorsal). 'Non-regressive' refers to posterior + anterior consonant sequences (i.e. dorsal-coronal, coronal-labial, dorsal-labial). With regard to the second factor (i.e. (52b)), clusters are defined as 'homogeneous' or 'heterogeneous'. In 'homogeneous' clusters, members share the laryngeal feature. Thus, they are voiced, voiceless or glottalised. Most scholars (Akhvlediani 1949, Vogt 1961, Melikishvili 1997) consider regressivity and homogeneity to be the preferred patterns for Georgian consonant sequences.

Both conditions, regressivity and homogeneity, are met in the consonant combinations referred to as 'harmonic clusters'. Because of the long tradition of the study of the harmonic clusters in Georgian, and because of the importance of their representation in my analysis, the following section is devoted to these.

¹⁹ In other words, sequences of homorganic consonants are disallowed in Georgian. The claim does not concern the class of sonorants since combinations of e.g. coronal obstruents and coronal sonorants are well-formed. For instance, sequences such as *dn*, *dr*, *t'l*, etc. are attested.

3.5.1. Harmonic clusters

3.5.1.0. Introduction

The history of the study of the harmonic clusters in the Georgian linguistic literature goes back to the nineteenth century. Over this period of time, harmonic clusters have been defined in many different ways, but the two requirements, regressivity and homogeneity, are recognised in almost all proposals (Khundadze 1901, Marr 1925, Akhvlediani 1949, Machavariani 1965 and Uturgaidze 1971, among others). Clusters of [-dorsal] [+dorsal] obstruents are called harmonic in Georgian because they share a laryngeal specification. There are two types of harmonic clusters, referred to as types A and B. They differ in their constituency. Type A refers to combinations of stops, affricates and fricatives with the velar stops /g k k'/ and type B refers to combinations of stops, affricates and fricatives with the fricatives /ɣ x χ'/. In some studies (Marr 1925, Machavariani 1965), combinations of fricatives with dorsal obstruents are also treated as harmonic clusters. In these clusters, the regressive order is respected and, taking into account the fact that laryngeal feature is less distinctive for fricative consonants than for stops, homogeneity is maintained. I incorporate these additions and present the clusters as follows:

(53)	<i>Type A</i>			<i>Type B</i>		
	bg	pk	p'k'	bɣ	px	p'χ'
	dg	tk	t'k'	dɣ	tx	t'χ'
	jg	ck	c'k'	jɣ	cx	c'χ'
	ǰg	čk	č'k'	ǰɣ	čx	č'χ'
	zg ²⁰	sk		zɣ	sx	
	žg	šk		žɣ	šx	

In order to define the status of the harmonic clusters, two types of evidence could be considered: phonological and phonetic. Phonological evidence is considered in this chapter. Phonetic evidence, which consists of a perceptual experiment on harmonic clusters, is discussed extensively in Chapter 4.

3.5.1.1. Phonological behaviour

There are several types of evidence: phonological processes, distributional facts, historical considerations, etc. that suggest an analysis of harmonic clusters as complex segments. The observations are considered below.

Harmonic clusters never occur across morpheme boundaries. They always belong to one morpheme, specifically to the lexical morpheme, the stem.

²⁰ I. Melikishvili has pointed out to me that harmonic clusters such as *zg* and *žg* are almost unattested. They were probably included in the set in order to maintain the symmetry with the Type B set.

- (54) a) *Stem-initial position*
 pkvil-i 'flour'
 t'k'bil-i 'sweet'
 dye 'day'
 čxir-i 'stick'
 cxvir-i 'nose'
- b) *Stem-final position*
 ortkl-i 'steam'
 marc'χ'v-i 'strawberry'
 čončx-i 'skeleton'
 otx-i 'four'
 vepxv-i 'tiger'

Harmonic clusters are the only obstruent sequences to appear in stem-final position. No other type of obstruent sequence is attested in this position (Vogt 1961, Deprez 1988).

Phonological processes affect both members of a harmonic cluster, i.e. in case one consonant of a harmonic cluster change, the other member also changes, such that the processes shown below, for instance, are not attested (Dzidzishvili 1966).

- (55) *dg ↗ tg *tk ↗ dk
 *dg ↘ dk *tk ↘ tg

Thus, both members of a cluster undergo a change, e.g. the alternation *dg ~ tk* is attested in Old Georgian form such as *c'ardgra ~ c'artkra* 'introduced' (Dzidzishvili 1966:232).

Some forms containing harmonic clusters have parallel forms with another type of harmonic cluster, e.g. both *burdγuni ~ bur't'χ'uni* 'muttering' and *bdγvriali ~ bč'χ'vriali* 'glitter, sparkle' are attested in Modern Georgian.

Consonant sequences other than harmonic clusters are characterised by optional /r/-insertion (Vogt 1958, Chikobava 1971, Deprez 1988).

Harmonic clusters always syllabify together intervocally, e.g. /si.t'χ'va/ 'word', /ce.cxli/ 'fire', while other obstruent clusters are never tautosyllabic, e.g. /mar.t'i/ 'March', /bev.ri/ 'much' (Akhvlediani 1949 and Žgent'i 1956, among others). Evidence for these syllabification patterns mostly comes from native speaker intuitions.

In reduplicated forms, harmonic clusters retain their complexity, e.g. /čkar-čkara/ 'quickly', /cxel-cxeli/ 'hot'. Other types of clusters do not usually participate in reduplication. Reduplication is discussed in more detail in Chapter 5.

Harmonic clusters are found in all Kartvelian languages (i.e. Svan, Megrelian and Laz). There are many examples of correspondences of harmonic clusters between these languages.

(56)	<i>Georgian</i>	<i>Megrelian</i>	<i>Laz</i>	<i>Svan</i>	
	mat'χ'l-i	mont'χ'or	mont'k'or		'wool'
	txra	txorua	ontxoru	lištxri	'to dig'

As shown in the examples in (56), the harmonic clusters do not undergo any simplification processes, and have direct correspondences in all Kartvelian languages.

3.5.1.2. Conclusions

To conclude, several types of evidence: phonological processes, distributional facts and historical considerations suggest that harmonic clusters can be analysed as complex segments. Phonetic evidence is discussed in Chapter 4. The harmonic clusters are considered as complex segments in numerous studies, e.g. Machavariani & Gamkrelidze (1965), Deprez (1988), Bush (1997), Nepveu (1994), Cho & King (1997), van Lit (1988), among others. The harmonic clusters are treated as being parallel to the C + /v/ combinations by Gamkrelidze & Machavariani (1965). In the following sections, I strengthen this position with additional arguments from the patterns of long consonant sequences and phonotactic characteristics of the stems of the CVC type.

3.5.2. Consonant combinations in adjacency

It has already been mentioned that consonant sequences appear word-initially. There are no word-final consonant sequences. Since word-medial clusters show the sub-patterns of word-initial clusters, I do not consider them in my analysis. For the patterns of stem-final word-medial consonant sequences, see Appendix 3.

In this section I would like to demonstrate that the complexity of word-initial consonant sequences is quite illusory in Georgian. The two-consonant clusters that seem to be the building blocks of much longer consonant sequences are of the type obstruent + sonorant. Thus they largely obey the SSP. The most frequently realised clusters are combinations of obstruent + sonorant, harmonic clusters and consonants that share a laryngeal specification and have [front] + [back] place of articulation. The clusters that do not obey the SSP are largely morphologically complex.

Since the domain of my analysis is the minimal word, two types of consonant sequences, /v/ + C and /m/ + C, will not be considered. In both cases, /v/ and /m/ are of morphological origin.

In monomorphemic words, /v/ can never be the first member of a cluster. It appears in word-initial position in clusters, but always as a prefix, e.g. /v-rčeb-i/ 'I am staying'. Two words, /vseba/ 'to fill' and /vrce/i 'wide' are exceptions to this generalisation. Unlike /m/, the sonorant /v/ can appear in other positions in clusters, but in almost all cases /v/ has a morphological origin and is subject to the metathesis process. For instance, in /k'vra/ 'to bind', /v/ is derived from the thematic /-av/ morpheme, /k'r-av-s/ 'somebody binds somebody else' (see (40)).

The sonorant /m/ is a nominaliser and can precede any consonant. The morphological origin of /m-/ is obvious in forms such as /m-c'eral-i/ 'writer', /m-k'v'leli/ 'killer' and /m-c'vrtnel-i/ 'trainer'. However there are some words where it is difficult to trace the morphological origin of /m-/, e.g. in forms such as /mz-e/ 'sun', /mgel-i/ 'wolf' and /mt'red-i/ 'pigeon'.

Comparative data could shed some light on the constituency of the forms with /m/ + C clusters. Consider, for example, correspondences between forms of Modern Georgian and another Kartvelian language, Megrelian.

(57)	<i>Modern Georgian</i>	<i>Megrelian</i>	
	mxari	xuji	'shoulder'
	msxali	sxuli	'pear'
	mxali	xuli	'name of a meal'
	mč'adi	č'k'idi	'maize-bread'

The Megrelian forms lack the word-initial /m-/, which may suggest that the /m-/ in Modern Georgian is a grammatical morpheme with no direct correspondences in other Kartvelian languages.

In addition, there are correspondences between Literary Georgian and other Georgian dialects (e.g. the West Georgian dialects Acharuli, Guruli and Imeruli) involving parallel forms with and without word-initial /m-/. Consider the correspondences in (58). The data come from Gudava (1979).

(58)	a)	<i>Literary Georgian</i>	<i>Acharuli</i>	
		mze	ze	'sun'
		mc'are	c're	'hot, bitter'
		mk'lavi	k'lavi	'arm'
		mc'χ'emsi	c'χ'esi	'shepherd'
	b)	<i>Literary Georgian</i>	<i>Imeruli</i>	
		mta	ta	'mountain'
		mgeli	geli	'wolf'
		mdidari	d'idari	'rich'
		mšobeli	šobeli	'parent'

The dialectal forms systematically lack /m-/. These correspondences once again substantiate the claim that /m-/ should be a prefixal morpheme rather than a part of the lexical morpheme in Literary Georgian. Thus, sequences of the /m/ + C and /v/ + C type are not considered to be genuine clusters in Georgian.

Typological studies on consonant sequences suggest the following generalisation: "... all languages exhibit the following property: if clusters of n Cs are possible syllable-initially, then clusters of $n-1$ Cs are also possible syllable-initially, and if clusters of n Cs are possible syllable-finally, then clusters of $n-1$ Cs are also possible finally" (Blevins 1995:217; see also Greenberg 1978). Maximally, six-consonant

clusters can appear word-initially in Georgian. This implies that there are also two-, three-, four- and five-member clusters.

Studying the constituency of two-member to six-member consonant sequences reveals a quite transparent structure (see (59)).²¹ Here I merely give a sample of the sequences beginning with the consonant /b/. I will try to generalise on the basis of these examples and point out interesting patterns.

(59)	<i>Consonant sequences beginning with /b/</i>	CC	CCC	CCCC	CCCCC
a)	Stop + stop	bg –	– –	bdyv –	bdyvn bdyvr
b)	Stop + affricate	bj bč'	bjγ	bc'k'	
c)	Stop + fricative	bγ – bz bž	bγv bγl		
d)	Stop + sonorant	bn br	brg brk' brj brt' brm	brč'χ'	brč'χ'v
		–	–		

The consonant sequences given in bold are special in the sense that they do not have a shorter counterpart. There are three such sequences: **bc'k'**, **bdyv**, **brč'χ'**. For example, the *bc'k'* sequence occurs in word-initial position, while **bc'* is unattested. The same can be said about the other two cases: sequences such as **bd* and **brč'* are not attested word-initially. All these consonant sequences contain harmonic clusters (i.e. *c'k'*, *dγ* and *č'χ'*, respectively). This indicates that *c'k'*, *dγ* and *c'χ'* appear as single units. This observation is an additional argument supporting the treatment of the harmonic clusters as complex segments. Hence, the representation of the sequences should be as follows: $c\overline{k'}$, $d\overline{\gamma}$, $\overline{c'}\overline{\chi'}$. After taking into account what we have said about harmonic clusters, we can analyse the $bc\overline{k'}$ sequence as a two-member consonant sequence. In the case of *bdyv*, another assumption about C + /v/ clusters is needed in addition to the assumption about harmonic clusters. I have already argued that the C + /v/ combinations are complex segments. Thus, even in the case of

²¹ For extensive data see Appendix 2.

an apparent four-member sequence, we end up with a two-member one, i.e. $bd\bar{y}^w$. Finally, consider the four-member $br\check{c}'\chi'$ sequence. The sequence is attested in only one word, $br\check{c}'\chi'ali$ 'claw', which has a parallel form $p'r\check{c}'\chi'ali$, with optional 'fleeting' /r/ (recall the description of the sonorant /r/ as 'fleeting' in section 3.3.1). The sequence $p'c^{\check{c}}\chi'$ exists in parallel with the sequence $p'r\check{c}'\chi'$. Thus, in these cases, as in the previous cases (i.e. as in cases of the $bc\bar{k}$ and $bd\bar{y}^w$ clusters), we are left with the two-member sequence $p'\check{c}'\chi'$.

The same transparency of long consonant sequences is observed in the following generalisation, suggested by Gvinadze (1970).

- (60) I /b p p' m/
 II /r/
 III /d t t' j c c' j' č č' z s ž š/
 IV /g k k' γ x χ'/
 V /v/
 VI /r l m n/

One consonant from each set can combine in the strict order given and form maximally a six-member cluster, e.g. /brdγvna/ 'to fight', /prckvna/ 'to peel'. Any set can be skipped, but the order between the sets should be respected.

Consonants in the sequence must have the same laryngeal specification and must be regressive. For instance, the five-member sequences are: $bd\gamma vr-$, $p'r\check{c}'\chi'v-$, etc.

As shown in (60), the sonorant /m/ is included in the first set. It seems that Gvinadze (1970) did not exclude word-initial /m/ + C combinations from her analysis. As I have mentioned, I do not consider /m/ + C combinations here, because of the morphological origin of /m/. The double occurrence of /r/, in both the second and sixth sets, is explainable by the optionality of the first /r/ in long consonant sequences.

It can also be seen from (60) that the first set consists of bilabial stops. This is remarkable, since we have already seen a number of distributional restrictions on bilabial consonants and rounded vowels in Georgian. The appearance of bilabials in the initial position of long consonant sequences can be related to the phonotactic restrictions that hold in general for the bilabials (including rounded vowels) in Georgian phonotactics. The initial position in the sequence is taken by bilabials because they have less combinatory power than other segments, such as coronals. The behaviour of the Georgian bilabials shows their marked character, and it is in accordance with the cross-linguistic observation that bilabials are marked obstruents (Melikishvili 1976).

The second consonant in a consonant sequence is the sonorant /r/, which is optional. For instance, it is optional in all six-member sequences. In general, in long consonant sequences it is optional when preceded and followed by consonants with an identical laryngeal specification. It is obligatory when preceded and followed by

non-homogenous obstruents. As was argued earlier, in such contexts the /t/ is syllabic.

The consonants from the third and fourth sets usually form harmonic clusters. They can be followed by the sonorant /v/ from set 5, which, as I have argued earlier, can be analysed as a secondary articulation on the preceding consonant. The longest consonant sequences are closed by a sonorant consonant from set 6.

Since the longest clusters of Georgian, as suggested in the discussion of the examples in (59), can be analysed as two-member clusters, the main topic of further discussion will concern two-member clusters. The following generalisation justifies the focus on two-member clusters:

(61) There are no three-member clusters which do not contain a sonorant.

Since Georgian sonorants are assumed to be syllabic in my analysis, apparent three-member sequences including a sonorant will be analysed as spurious sequences with the syllabic sonorant. For instance, the sequences *kld* in *klde* 'rock' and *pvt* in *prta* 'wing' include the syllabic sonorants *l̥* and *ɾ̥*, and in the analysis will be broken up by an unspecified vowel (see Chapter 6). Three-member obstruent sequences, e.g. *t'k'b*, *p'c'k'*, *bjg*, are considered as two-member sequences because they contain harmonic clusters *t'k'*, *c'k'*, *bg*, respectively.

Surface stem/word-initial two-member sequences are given in (62). The table depicts two characteristics of biconsonantal sequences discussed above, regressivity and homogeneity. The table does not include /m/ + C and /v/ + C sequences.

(62) Possible (+) and impossible (-) initial biconsonantal clusters

C ₁ /C ₂	b	p	p'	d	t	t'	g	k	k'	j	c	c'	ǰ	č	č'	z	s	ž	š	ɣ	x	χ'	v	l	r	m	n	
b	-	-	-	-	-	-	+	-	-	-	-	-	+	-	+	+	-	+	-	+	-	-	-	+	+	-	+	
p	-	-	-	-	+	-	-	+	-	-	-	-	-	+	-	-	+	-	+	-	+	-	-	-	+	+	-	+
p'	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+
d	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+
t	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-
t'	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-
g	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	+	+	+	+
k	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	+	-	-	-	-	+	+	+	+	+
k'	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
j	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	+	+
c	+	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	+	+
c'	+	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
ǰ	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	-	+	-
č	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+
č'	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	+	+
s	-	+	+	-	-	+	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+
ž	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	+	-
š	-	+	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+
ɣ	-	-	-	+	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
x	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	+	+
χ'	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
l	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+
r	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-	-	-
n	-	-	-	+	+	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

The shaded areas in some parts of the table denote high concentrations of pluses, which are the most frequently realised clusters. There are three such areas in the table: (i) the last five columns, (ii) two rows, (iii) the rest of the shaded areas. I will discuss each of these in turn.

(i) Most '+'s in (62) are concentrated in the last five columns of the table. These are combinations of consonants with the sonorants /v l r m n/, i.e. sequences of the C + sonorant type. These show that stem/word-initial clusters have a preference for rising sonority. As argued in section 3.3.2, sequences of C + /v/ can be analysed as complex, labialised segments.

(ii) Two shaded rows present clusters with *s* and *r* as their first member. These are clusters with falling sonority. The two are different in that the sequences of the /s/ + C type can be analysed as complex segments, while in /r/ + C sequences the sonorant /r/ appears to be syllabic and its occurrence as a first member of two-member sequences is related to a number of factors to be discussed below.

The frequency of the /s/ + C clusters can be related to the large number of old and recent borrowings from the Indo-European languages, e.g. Greek, Latin, English, etc. For instance, *sp* is attested in /spero/ ‘sphere’; *sp*’ is attested in /sp’ort’i/ ‘sport’ and /sp’irali/ ‘spiral’; etc. Interestingly, the sequences of the /s/ + obstruent type behave as complex segments in Modern Georgian. Some of them are treated as harmonic clusters, e.g. *sk* and *sx* (see (53)). Reduplication patterns also provide additional evidence for treating these sequences as complex segments (see Chapter 5).

The sequences of the /r/ + C type are very common in Georgian. Melikishvili (1997) proposes an explanation for the appearance of the sonorant /r/ in stem/word-initial position. The arguments are briefly summarised below.

a) The first factor causing the appearance of /r/ in word-initial position is the transition from closed syllable to open syllable in the development of Georgian.

b) The second factor conditioning the appearance of /r/ in word-initial position is related to the process whereby the fricative /s/ in word-initial position in Old Georgian²² changed into the sonorant /r/ in Modern Georgian.

(63)	<i>Old Georgian</i>	<i>Modern Georgian</i>	
	sje	rje	‘milk’
	sjali	rjali	‘groom’
	sjuli	rjuli	‘religion’
	stveli	rtveli	‘vintage’

In all of these examples the /s/ seems to have a morphological origin but the change to /r/ still needs to be explained.

c) Melikishvili (1997) relates the occurrence of /r/ cluster-initially to a cluster-simplification process. She says that the Proto-Kartvelian syllabic /r/ became non-syllabic, which was followed by deletion of a preceding consonant, i.e. *C₁C > CrC > rC:

(64)	<i>Proto-Kartvelian</i>	<i>Old Georgian</i>	<i>Modern Georgian</i>	
	grk’ali >	grk’ali >	rk’ali	‘arc’
	grc’χ’ili >	grc’χ’ili >	rc’χ’ili	‘flea’
	k’rk’o >	k’rk’o >	rk’o	‘acorn’
	grgoli >	rgoli >	rgoli	‘ring’

(Melikishvili 1997:33)

In section 3.3.1 I argue that in both cases ((63) and (64)) /r/ in Modern Georgian is phonetically syllabic.

(iii) The rest of the shaded areas refer to harmonic clusters.

Let us now consider some of the biconsonantal sequences which are not within the shaded areas of the table but are attested in Modern Georgian, such as *gz*, *cd*, *ct*, *tb*, etc. Ertelishvili (1970) and Uturgaidze (1976) convincingly argue that such sequences are derived from CVC stems. Consider the following forms:

²² /s-/ in Old Georgian could have a morphological origin, as a prefix.

(65)	gz	gza	‘way’	gezi	‘direction, way’
	ct/cd	cda	‘to try’	e-cad-e	‘try you’
	tb/t’p	tbili	‘warm’	gan-t’ep-i-t	‘get warm (Old Georgian)’

These types of consonant sequences seem marginal to the Georgian language. Sometimes, as shown in the examples in (65), it is possible to trace the vowel between the consonants by looking at parallel forms of Modern or Old Georgian, but sometimes it is difficult to find such evidence, for instance in case of clusters such as $\chi'b$ and $k'b$. Comparative data shed some light on the origin of such sequences, e.g. the sequence $\chi'b$ attested in the word / $\chi'ba$ / ‘jaw’ is reconstructed in Proto-Kartvelian as / $\chi'ab$ /; the corresponding form in Svan is / $\chi'ab$ / (Sardschweladse & Fähnrich 1990). The consonant sequence $k'b$, which is attested in the word / $k'bili$ / ‘tooth’, can also be reconstructed as / $k'Vb$ /, because the corresponding form in Megrelian and Laz is / $k'ib-ir-i$ /. Thus, the form with a vowel between the consonants / k' /and / b / occurs in Kartvelian languages.

To conclude, three types of biconsonantal sequences are commonly attested.

(66) *Preferred biconsonantal sequences*

- a) Obstruent + sonorant
- b) Harmonic clusters
- c) / s / + obstruent
- d) / r / + obstruent

Of these four types of consonant sequences, only the sequence of obstruent + sonorant (i.e. (66a)) is a genuine biconsonantal cluster, since, as advocated in this thesis, harmonic clusters are treated as complex segments. The sequence of / s / + obstruent type is also considered as complex segment (see Chapter 5 for additional evidence) and / r / + obstruent sequences are not considered as genuine biconsonantal clusters because of their derived nature.

3.5.3. Conclusions

The study of word-initial consonant sequences has shown that the complexity of consonant sequences is quite illusory in Georgian. The two-consonant clusters which seem to be the building blocks of much longer consonant sequences are of the type obstruent + sonorant. Thus they largely obey the SSP.

The most frequently realised clusters are combinations of obstruent + sonorant, harmonic clusters and consonants that share a laryngeal specification and have [front] + [back] place of articulation. Clusters that do not obey the SSP are largely of morphological origin.

3.6. Testing the hypothesis

In Chapter 2, I proposed that all two-member consonant clusters are derived by vowel deletion; their co-occurrence restrictions are derivable from stems of the CVC type.²³ The proposal was formulated as a hypothesis, which is repeated as (67).

(67) *Hypothesis*

If a language has clusters of the C_iC_j type, then the language will have stems of the C_iVC_j type.

The merits of this hypothesis are that it is easily refutable, easily testable and has predictive power. The following discussion is devoted to the verification of the hypothesis.

I have already formulated the basic characteristics of biconsonantal sequences in Georgian. Now I turn to consonant combinations across a vowel.

Firstly, I summarise the studies by Kobalava (1967) on Modern Georgian stems of the CVC type and Melikishvili (1997) on the typological study of the Kartvelian root structure. Secondly, I present the results of my findings. Finally, the restrictions on combinations of consonants at a distance and in adjacency will be compared.

3.6.1. Studies by Kobalava (1967)

The investigation of Kobalava (1967) was based on consonant combinations in stems of the CVC type. Using an explanatory dictionary of Georgian she studied all nominal and verbal stems of the CVC type. Particular attention was paid to consonants.²⁴ Of a potential 729 possible consonant combinations only 457 are attested (i.e. 63 %). The unattested 272 possibilities reflect the structural properties of Georgian phonotactics. These findings can be summarised as follows:²⁵

a) Consonants from the same set do not co-occur in CVC stems (excluding combinations of identical consonants):

(68) *Stops and Affricates*

/b p p' /

/d t t' /

/j c c' /

/j č č' /

/g k k' /

²³ Consonant clusters here refer to genuine clusters, and not ones which are the result of complex segment formation or morphological merger, i.e. the result of a lexicalisation process.

²⁴ Kobalava (1965) mentions a possible correlation between constraints on adjacent and distant consonant combinations, but does not investigate this further.

²⁵ There are some exceptions (mostly loan words) to the generalisations in (68) and (69), but they certainly reflect the basic phonotactic properties of Georgian stems.

Fricatives

/z s/

/ž š/

/ɣ x ɣ' /

b) Affricates and fricatives do not combine with each other: /j c c' z s ʃ č č' ž š/.

c) Some consonant combinations in a certain order are allowed, while the reverse order is unacceptable, e.g. /b p p' / + V + /m v/ is attested, while */m v/ + V + /b p p' / is unattested.

d) If the first consonant is a glottalised stop or an affricate, the second voiced consonant must have a place of articulation that is further to the front than that of the first consonant (restrictions for the second voiceless or glottalised consonants are not attested). The generalisation is expressed in (69).

(69)	p'	V	*b	*d	*j	*j̣	*g
	t'	V	b	*d	*j	*j̣	*g
	c'	V	b	d	*j	*j̣	*g
	č'	V	b	d	j	*j̣	*g
	k'	V	b	d	j	j̣	*g

Thus, the combinations *t'b*, *c'b*, *k'b* are attested, while the combinations **t'p'*, **c'j*, **c'g*, etc. are disallowed.

e) The occurrence of two /t/'s or two /v/'s is disallowed within a CVC stem.

All possible (denoted by the symbol +) and impossible (denoted by the symbol –) consonant combinations within the stems of the CVC type are summarised in (70). The shaded areas denote the high concentrations of pluses. Note that most pluses appear in columns where the second consonant is a sonorant. Thus, in Georgian the realisation of the CVS (sonorant) structure is almost 100%. Recall that adjacent consonants also predominantly have a CS structure.

(70) *Consonant combinations within the stem of the CVC type (Kobalava 1967)*

C ₂ /C ₁	b	g	d	v	z	t	k'	l	m	n	p'	ž	r	s	t'	p	k	γ	χ'	š	č	j	c'	č'	x	ǰ
b	+	+	+	+	+		+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+
g	+	+	+	+	+	+	-	+	+	+		+	+	-		+	-		-	+		+			+	+
d	+	+	+	+	+	-	+	+	+	+		+	+	+	-	+	+	+	+	+		+				+
v	-	+	+	-	+			+	-	+	-	+	+			-				+		+				+
z		+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+			-	-	-	-	-	-	+
t	+	+	-	+		+	+	+	+	+			+	-	+		+			+						+
k'	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+		+	+	+	+	
l	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
m	-	+		+		+		+	+	+	-	+	+	+	+	-	+			+					+	+
n	+		+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+				+	+	+
p'	-	-	-	+			+	+	+	+	+	+	+	+	-					+			-	+	+	+
ž				+	-		+	+	+	+	+	+	+	-	+					-	-	-	-	-	-	-
r	+	+	+	+	+		+	+	+	+		+	+		+	+	+		+	+		+	+	+	+	+
s	+			+	-	+	+	+	+	+	+	-	+		+	+	+	+		-	-	-	-	-	-	+
t'	+	-	-	+	+	-	+	+	+	+	+		+	+	+	+				+	+	+	-			+
p	-	+		+		+	+	+	+	+	-	+	+	+	+		+	+	+	+	+	+	+	+	+	+
k	+	-	+	+	+	+	-	+	+	+			+	+	+	+	+		-	+	+	+			+	+
γ	+		+	+	+		+	+	+	+	+	+	+		+	+		+	-	+		+			+	-
χ'	+	-	+	+	+	+	-	+	+	+	+	+	+		+	+	-	-	+	+	+				+	-
š	+		+	+	-	+	+	+	+	+	+	-	+	-	+		+			+	-	-	-	-	-	+
č		+	+	+	-	+	+	+	+	+		-	+	-	+	+	+			+	-	+	-	-	-	+
c			+	+	-	+	+	+	+	+		-	+	-	+	+				-	-	+	-	-	-	+
j	+	+		+	-		+	+	+	+		-	+	-		+		+		-	-	-	-	-	-	+
c'		-	+	+	-	+	+	+	+	+	+	-	+	-	+		+			-	-	-	-	+	+	-
č'		-	+	+	-	+	+	+	+	+	+	-	+	-	+		+	+		-	-	-	-	-	+	+
x			+	+	+	+	+	+	+	+	+	+	+	+	+				-	-	+	+	+	+	+	+
ǰ	+	+	+	-	-	+	+	+	+	+	+	-	+	-	+	+	+	+	+	-	-	-	-	-	-	+

3.6.2. Studies by Melikishvili (1997)

Melikishvili’s studies on Kartvelian root patterns are of great interest for my analysis, not only because of the data and generalisations (in addition to those already observed by Gamkrelidze & Machavariani 1965, Fänrich 1978, etc.), but most of all because of her theoretical standpoint. When looking at the root/stem and the syllable structures of different types of genetically distinct languages, she establishes the following correlation: the basic phonotactic constraints for the stem/root and syllable structures are the same cross-linguistically.²⁶ While accepting the general idea, I will extend it and offer some additional proposals to strengthen it.

²⁶ The same type of claim is made on the basis of studies of the Semitic languages (see, e.g. Belova 1991, Diakonoff 1970, 1988).

The consequence of the correlation between the root/stem and syllable structures is to expect the Sonority Sequencing Principle – originally formulated for the syllable domain – also to function on the stem domain. Melikishvili demonstrates this for the Kartvelian CVC stems. Firstly, she discusses the distributional characteristics of sonorants in CVC stems. These occur in initial and final position in accordance with the general sonority schema: /m/ → /n/ → /l/ → /r/.²⁷ The sonorant /r/ is the most dispreferred in the initial position,²⁸ although it is often found in final position. The reverse situation holds with the least sonorous consonant /m/, which is more likely to occur stem-initially than stem-finally. These observations for Kartvelian stems are supported by similar studies on stem patterns of other languages (Indo-European (Magnusson 1967, 1979), Semitic (Greenberg 1970) and German (Twaddell 1939, 1941)).

More original are Melikishvili's generalisations on the distributional characteristics of stops in CVC sequences. Assuming the importance of defining the sonority scale not only for manner, but also for place of articulation features, she proposes the following constraint for accounting for the phonotactic restrictions within the stems of the CVC type.

(71) *The Compensatory Principle (CP)*

The co-occurrence patterns of heterorganic consonants across a vowel are such that if the first consonant is more sonorous than the second according to the laryngeal specification, the second consonant will be more sonorous according to the place of articulation. If the first consonant is more sonorous than the second according to the place of articulation, then the second consonant will be more sonorous according to the laryngeal specification (Melikishvili 1997:57, my translation).

The implications of the principle are illustrated with respect to the Kartvelian roots of the CVC type in (72):

(72)

Laryngeal Specification		Place of Articulation	
		Rising Sonority Regressive	Falling Sonority Non-regressive
Rising Sonority	Glottalised-voiced	p'-g	k'-b
	Voiceless-voiced	p-g	k-b
	Glottalised-voiceless	p'-k	k'-p
Falling Sonority	Voiced-glottalised	b-k'	*g-p'
	Voiced-voiceless	b-k	*g-p
	Voiceless-glottalised	p-k'	*k-p'

²⁷ Nebieridze (1974) proposes this sonority hierarchy.

²⁸ The absence of the sonorant /r/ in stem-initial position (in terms of syllable constituency, onset position) is attested in a number of languages, e.g. Basque, Armenian, Greek, Khunzic and Andic.

The combinations of k'-b, k-b and k'-p, as well as their reverse order b-k', b-k and p-k' are attested, while combinations violating the requirement do not occur. Combinations of *g-p', *g-p and *k-p' are not permitted. In the latter combinations, the sonority falls, according to both place of articulation and the laryngeal specifications, while in the former cases sonority increases in one of the two dimensions: it either increases in terms of place of articulation or in terms of laryngeal specification.

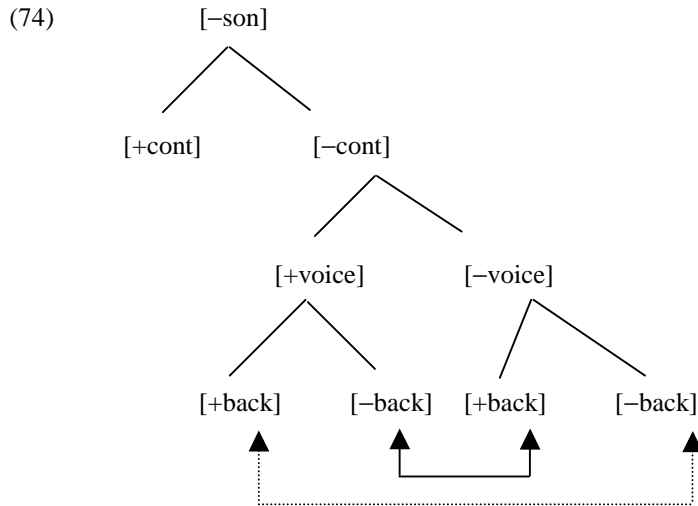
The CP holds for Modern Georgian consonant combinations across a vowel. As Melikishvili points out, the generalisation made by Kobalava follows from this principle. For ease of exposition I repeat the generalisation.

- (73)
- | | | | | | | |
|----|---|----|----|----|-----|----|
| p' | V | *b | *d | *j | *j̃ | *g |
| t' | V | b | *d | *j | *j̃ | *g |
| c' | V | b | d | *j | *j̃ | *g |
| č' | V | b | d | j | *j̃ | *g |
| k' | V | b | d | j | j̃ | *g |

Thus, *p'b, *p'd, *p'j, *p'j̃, *p'g, *t'd, *dj, etc. are not attested clusters; while t'b, c'b, c'd, etc. are well-formed clusters.

Melikishvili's proposal is justified not only by language data, but also from an articulatory-perceptual perspective. Two requirements desirable for any language structure, ease of perception and ease of articulation, are achieved in consonant combinations with different place of articulation and different laryngeal specifications (see dissociation of like consonants (Krupa 1967)).²⁹ Consonants that are too dissimilar are dispreferred in the same way as consonants that are too similar, e.g. a [-voice, -back] consonant does not combine with a [+voice, +back] consonant. This could be due to the fact that they are too distant. The balance is achieved when features are specified in alternating way, e.g. a [-voice, +back] consonant combines well with a [+voice, -back] consonant. The Compensatory Principle that captures the optimal combinations of consonants is depicted in (74). A dotted line denotes the dispreferred consonant combinations, while a solid line denotes the preferred ones.

²⁹ Zubkova (1990) establishes the existence of a tendency: combinations of vowels are characterised by harmony (homogeneity), while combinations of consonants are characterised by contrast (heterogeneity). This holds not only for adjacent consonants, but also for consonants at a distance.



Melikishvili's proposal is comparable to the hypothesis proposed by Saporta (1955).

... the average frequency of a consonant cluster is a function of the difference between the phonemes in the cluster: low frequencies are expected for clusters which are either extremely similar or extremely dissimilar; high frequencies are expected for clusters which are at neither extreme (Saporta 1955:25).

The reasoning behind this is that the relative frequency of consonant clusters reveals a tendency on the part of any language system to produce speech in such a way as to consider the effort of both the speaker and the listener, the encoder and the decoder. Saporta says:

In any consonant cluster, the situation of least effort for the speaker is that in which the successive phonemes are not similar; but this requires maximum effort on the part of the listener, who is then forced to make a series of fine discriminations. For the listener, the optimal situation is that in which the phonemes of a cluster differ as much as possible; but this requires maximum effort on the part of the speaker (Saporta 1955:25).

The Compensatory Principle balances combination of feature classes in consonant sequences and disallows too similar or too dissimilar consonant combinations. Thus, the CP can be regarded as one of the instantiations of the Balancing Principle, just like the SSP and the OCP, discussed in Chapter 2. The CP operates in Georgian and the Kartvelian languages (as advocated by Melikishvili 1997), but the cross-linguistic validity of this principle remains to be verified.

3.6.3. My study

In Chapter 2, I proposed the following hypothesis:

- (75) *Hypothesis*
 If a language has C_iC_j clusters, then the language will have stems of the C_ivC_j type.

As predicted, the restrictions within clusters are more constrained than between consonants across a vowel; compare the two tables given in (62) and (70), for example.

In order to test the hypothesis in (75), I used a dictionary containing 60,000 words (Chikobava 1986). I extracted all stems of the C₁V₁C₂ type; afterwards all word-initial biconsonantal sequences were extracted. The co-occurrence patterns in biconsonantal clusters were compared with the co-occurrence patterns across a vowel within C₁V₁C₂ stems.

The findings of the comparison were that there are no biconsonantal clusters whose members are not found within a CVC-type stem domain,³⁰ for example, Georgian does not permit stems such as *dVp', *tV j, *t'aj, *dVt', and, consequently, *dp', *tj, *t'j, *dt' clusters are not allowed either. Notable exceptions are harmonic clusters, e.g. clusters such as p'k', p'χ' and t'χ' are attested while stems of the type *pV'k', *p'Vχ' and *t'Vχ' are not. This observation provides further justification for the treatment of the harmonic clusters as complex segments.

Having established generalisations on consonant combinations in adjacency and across a vowel, I can spell out some restrictions that hold for consonant combinations both in adjacency and across a vowel. Members of the following phonemic sets do not combine with each other:

- (76) a) Stops /b p p'/ /d t t'/ /g k k'/
 b) Affricates and fricatives /j c c' z s ʃ č č' ž š/
 c) Fricatives /χ x χ'/

The restrictions concerning fricatives and dental stops are as follows: a sequence of dental stop + fricative is disallowed, while the combination fricative + dental is permitted. One restriction concerns bilabial consonants (including /v/): /b p p' / + /m v/ is attested, but /m v/ + /b p p' / is disallowed. Thus, the constraints on C_iC_j in both types of structures, C_iC_j (cluster) and C_iVC_j, are similar.

3.6.4. Conclusions

Similarities have been found between the composition of stems and consonant sequences, which suggests the hypothesis in (75).

³⁰ As I have said, clusters seem to be the result of vowel deletion. This holds only for true, i.e. core, clusters. Some 'clusters' like harmonic clusters, sequences of C + /v/ and /s/ + obstruent are treated as single segments.

In order to test the hypothesis, I compared the consonant co-occurrence restrictions in adjacency and across a vowel. The co-occurrence patterns in both contexts are quite similar, e.g. both show a preference for rising sonority, i.e. show the effect of the SSP. In addition, both types of restrictions have similar OCP effects.

The plausibility of the hypothesis has been confirmed on the basis of data from Georgian. In order to demonstrate the cross-linguistic validity of this hypothesis, it would be necessary to test it on a large corpus of genetically distinct languages.

The hypothesis leads to another sub-hypothesis:

- (77) All consonant sequences (maximally biconsonantal) are derived, i.e. the result of vowel deletion.

I return to this claim in Chapter 6, which focuses on the analysis of the Georgian consonant sequences. This claim concerns only genuine biconsonantal clusters; as illustrated in this chapter, surface consonant sequences may be the result either of morpheme concatenation or of complex cluster formation, as is the case with harmonic clusters or with C + /v/ combinations. It is also possible that the language adopts morphologically complex loan words as monomorphemic ones. However such words would represent a restricted subset of the vocabulary and should be treated separately by taking into account the patterns of both host and donor languages.

3.7. General conclusions

The most salient observations about the Georgian word-level phonotactics are as follows.

Georgian, as a Grammatical language (i.e. a language with features of inflectional and agglutinative morphology), demonstrates the discrepancy between lexical and grammatical morphemes in terms of the use of the phonological inventory.

One way to explain the surface consonant sequences of Georgian is to carefully study the morphological structure of words where the sequences could be the result either of the deletion of a stem vowel (which generally happens when a vowel-initial affix is added to a root) or of mere addition of a consonantal affix to a root.

The disyllabicity of the Georgian minimal word is substantiated by phonological (accent assignment), phonetic (monosyllabic lengthening) and morphological (alienable/inalienable constructions) evidence. Additional evidence comes from the reduplication process discussed in Chapter 5. The Georgian minimal word can be formally presented as a general disyllabic template of the $C_1V_1C_2V_2$ type.

Georgian sonorants are syllabic in consonant sequences. This claim is based on phonological, ditributional, historical and comparative evidence. This especially concerns the most sonorant consonant /r/. A comprehensive phonetic study should be carried out to test this claim.

Combinations of $C + /v/$ can be treated as complex labialised segments. The claim is based on phonological, distributional and historical evidence. This claim is substantiated by reduplication data in Chapter 5.

Phonological processes, distributional facts and historical considerations suggest that harmonic clusters can be analysed as complex segments. Phonetic evidence discussed in Chapter 4 strengthens this claim.

The study of word-initial consonant sequences has shown that the complexity of consonant sequences is illusory in Georgian. Two-consonant clusters, which are the building blocks of much longer consonant sequences, are of the type obstruent + sonorant. Thus, they largely obey the SSP.

The most commonly attested clusters are combinations of obstruent + sonorant, harmonic clusters and consonants that share the laryngeal specification and are regressive ([front] + [back] place of articulation). Sequences not obeying the SSP are largely of morphological origin.

Similarities are found between the composition of stems and consonant sequences. These observations lead to the hypothesis that if a language has C_iC_j clusters, then the language will have stems of the C_iVC_j type.

The plausibility of the hypothesis has been confirmed on the basis of the Georgian data. The hypothesis entails that all consonant sequences (maximally biconsonantal ones) are derived, being the result of vowel deletion. I return to this claim in Chapter 6.

