WORD-INITIAL CONSONANT CLUSTERS IN ALBANIAN

Rachel Klippenstein
The Ohio State University

Abstract

Albanian has a wide but not unrestricted range of initial consonant clusters. This paper lays out some constraints on such clusters; e.g., there are no clusters of two voiced stops, nor of voiced obstruent + voiceless obstruent. Dictionary data is supplemented by phonetic data from a native Albanian speaker, which helps determine how well orthographic evidence reflects pronunciation. I find that vowel epenthesis in obstruent-obstruent clusters is rare; schwa is sometimes elided to form clusters that are not orthographically evident, but less often than expected; and clusters written with voiceless obstruent + voiced obstruent are pronounced as such, at least sometimes.

1.   Introduction

Albanian has a wide range of word-initial consonant clusters. Some of the more remarkable ones include *çd* [t͡ʃd] in *çdo* 'every', *tk* in *tkurrje* 'contraction', *zhvl* [ʒvl] in *zhvleresoj* 'devalue'; the longest contain four consonants, such as *zmbr* [zmbɾ] in *zmbraps* 'repel', and *çmpl* [t͡ʃmpl] in *çmpleks* 'untwine'. With such a rich variety of clusters, Albanian offers many interesting issues about syllabicity, the relationship between the phonology of morphologically simple and morphologically complex words, and the relationship between spelling and pronunciation. However, few people have yet addressed these challenges; work on the topic is essentially limited to lists of clusters (e.g., Buchholz & Fiedler 1987:46–50), and does not provide any substantial account. Before the interesting issues raised by Albanian clusters can be investigated in depth, a basic description and analysis of the facts is needed. In this paper I provide a step towards this goal by providing a description and basic analysis of two-consonant clusters (as well

as the first two consonants of longer clusters). I use dictionary-based evidence to provide an overview of constraints on consonant clusters, and supplement this with instrumental phonetic data to check how closely the spellings given in the dictionaries reflect the actual pronunciation of a native speaker. In addition to checking whether initial combinations of consonant letters are in fact pronounced as clusters, I also investigate the possibility that the elision of schwa in initial syllables results in consonant clusters that are not spelt as such. Finally, I investigate whether there is in fact a voicing mismatch in obstruent clusters spelt as *çb* and *çd*.

1.1   **Background on Albanian**

Albanian is an Indo-European language in its own branch of the Indo-European language family. It has two major dialect groupings: Geg, the northern variety, spoken primarily in northern Albania and Kosovo, and Tosk, the southern variety, spoken primarily in southern Albania, as well as in various diaspora communities in Italy, Greece, and elsewhere; there are many smaller dialect divisions within these two main groupings (Newmark et al. 1982:6–7, Friedman 2006:14). Standard Albanian is based primarily on Tosk, but has some Geg features (Newmark et al. 1982:8).

		labial	dental/alveolar[1]		palato-alveolar	palatal	velar	glottal
stop	voiceless	p	t			q /c/	k	
	voiced	b	d			gj /ɟ/	g	
fricative	voiceless	f	th /θ/	s	sh /ʃ/			h
	voiced	v	dh /ð/	z	zh /ʒ/			
affricate	voiceless		c /t͡s/		ç /t͡ʃ/			
	voiced		x /d͡z/		xh /d͡ʒ/			
nasal		m	n			nj /ɲ/		
tap			r /ɾ/					
central approximant			rr /r/			j[2]		
lateral approximant			l /l/	ll /ɫ/				

Table 1. Obstruents and nasals (orthography = IPA unless otherwise indicated)

As background for discussing consonant clusters, an overview of the consonant inventory of Standard Albanian may be helpful. The obstruents, nasals, and sonorants are shown in Table 1 above; when orthographic representation differs from IPA, the IPA equivalent is indicated.[3] The values given here are based on Buchholz & Fiedler (1987:37–42).

---

[1] Buchholz & Fiedler (1987) classify *th, dh* as apicodental, *t, d, ll* as alveolar-dental (dental in Geg, alveolar in Tosk), and *s, z, c, x, n, l, r, rr* as alveolar. Newmark et al. (1982:9–10) describe *t, d, th, dh, n* as apicodental, and *s, z, c* as apicoalveolar; *x* /d͡z/ appears to have a typo in its description and presumably is meant as apicoalveolar along with *z*. According to Friedman (2006:1), *t, d, c, x, ll, n, r, rr* are 'alveolar, NOT dental, (except in some Geg)'.

[2] Buchholz & Fiedler (1987) consider [j] an allophone of /i/, not a consonant phoneme. I follow Newmark et al. (1982:10, 13–14) in considering it a consonant.

[3] The vowel system of Standard Albanian consists of the vowels *i, y, e, a, o, u, ë*. With the exception of *ë*, the orthographic representation of vowels aligns reasonably closely with IPA; *ë* is a central unrounded vowel whose precise quality varies; it may be stressed (Newmark et al. 1982:11–12). I will refer to it as schwa and use the IPA symbol /ə/ when necessary. Albanian also has diphthongs *ie, ua, ye, ue* (Newmark et al. 1982:12).

1.2   Writing and speech

The relationship between writing and speech is a complicated one. Speech exists before writing both historically and in the life of a speaker. For good reason, therefore, linguists typically view written language as dependent upon spoken language, a reflection which is considerably distorted by constraints that the written medium imposes. In addition, once a written system is established, it tends to change slower than the associated spoken language, giving rise to additional mismatches between writing and speech. However, the relationship between writing and speech is not entirely unidirectional, with influence from speech flowing towards writing, and never the other way round. Writing also influences spoken language, as in the case of spelling pronunciations, where speakers come to pronounce a word based on its spelling rather than on its traditional pronunciation. The interaction between writing and speech means that there is a non-arbitrary relationship between them, and written language can help in understanding spoken language, so long as the limitations of the relationship are kept in mind and writing is not taken as a simple substitute for speech.

The present written standard for Albanian was developed fairly recently, in several stages over the course of the 20th century. At the Congress of Manastir in 1908, two ways of writing Albanian were established as acceptable: the present phonetically based Latin alphabet, and a system based on the Turkish/Arabic alphabet. The Literary Commission of Shkodër in 1916/1917 and the Educational Congress of Lushnjë in 1920 determined that the southern Geg dialect of Elbasan should be the basis for the standard, and this was taught at teachers' training school. However, this decision did not take hold, and people continued writing in both Geg and Tosk. Gradually, a shift occurred towards usage of Tosk with an admixture of Geg, encouraged in part by the use of Tosk in official documents. In the early 1950s, the Albanian Writers' Union and the National Conference on Orthography decided that the literary standard should be Tosk alone. The culmination of standardization efforts was the 1972 Congress of Albanian Orthography, which laid out the rules of orthography for a Tosk-based Standard Albanian, and lead to the publication of official orthographic works. This Tosk-based standard continued (and continues) to have some Geg features, especially in lexicon and morphology (Newmark et al. 1982:6–9, Moosmüller and Granser 2006:122–123).

With this basis, it is reasonable to hypothesize that written standard Albanian reflects carefully spoken standard Albanian to a considerable degree, since it was intentionally designed to be phonologically grounded and the time since standardization is quite short, so that drastic phonological changes are unlikely to have caused the spoken language to change significantly while the writing system remained constant. Nevertheless, this is no guarantee that the written language is an accurate guide to the phonetics of the spoken language, and even where it reflects careful speech, casual or fast speech is likely to differ. Thus, it is necessary to confirm orthographically based hypotheses with phonetic data.

2.   Standard Albanian consonant clusters from written sources

As mentioned above, Albanian has a wide range of clusters allowed word-initially, but possibilities are not unconstrained. In this section I work towards a full analysis of

Albanian word-initial clusters by examining the constraints governing which clusters do and do not occur, based on evidence from written sources. I limit my investigation to the initial two consonants of a word; primarily this includes two-consonant clusters, but I also include the initial two consonants of three-consonant clusters, where these do not seem to occur without a third consonant. I assume that in the occasional cases where two consonants are found as the initial two consonants of a three-consonant cluster, but do not occur as an initial two-consonant cluster, the lack of a two-consonant cluster is probably an accidental (i.e., nonsystematic) gap. I impose this limitation for two reasons: first, to keep the investigation to a manageable size, and second, because two-consonant sequences are the most basic level of complexity, which form the necessary background for any fuller account.

As far as I am aware, very little work has been done to investigate the constraints on Albanian consonant clusters. Buchholz & Fiedler (1987:47ff) provide a list of word-initial clusters, but give no analysis of the principles constraining them.

In this project, I assemble and analyze a list of consonant clusters based on Newmark (1998), supplemented by Stefanllari (1996). Newmark (1998) is a very thorough dictionary including many rare, obsolete, and dialectal words; non-standard words are easy to identify since they are marked with an asterisk. I restrict my analysis to clusters found in words not marked as non-standard, in order to avoid combining data from dialects that may have differing phonologies. Since Newmark (1998) contains many words that are probably not familiar to most native speakers,[4] I compare the list of clusters derived from Newmark with those found in Stefanllari (1996), a much smaller dictionary, almost all of whose words are probably familiar to most Albanian speakers. Clusters found in Newmark (1998) but not in Stefanllari (1996) are italicized; clusters found in only one word in Newmark (1998) and none in Stefanllari (1996) I consider marginal, and put in italics and parentheses. It is conceivable that the competence of some Albanian speakers might not include these clusters; however, the fact that they occur in some words, even if only rare ones, suggests that they are not systematically excluded by Albanian grammar.

Newmark (1998) indicates by means of italics that certain instances of schwa (written as <ë> in Albanian orthography) may be elided. In many cases, Newmark indicates that a schwa between a word-initial consonant and another consonant may be elided, creating a word-initial cluster. Some of the clusters that would be formed by elision are also found in cases without elision (e.g., [ps] in *pse* 'why', *pësoj* 'undergo, suffer'), while others are not (e.g., [kʃ] in *këshill* 'council'); in addition, there are some clusters that are possible without elision, but appear to be impossible as results of elision (e.g., [ʃt] in *shtet* 'state' but not in *shëtit* 'stroll around'[5]). I discuss clusters without elision and with elision separately.

2.1 Clusters without elision

---

[4] I encountered many that were not familiar to the speaker I worked with.

[5] Newmark (1998) indicates that the stress in *shëtit* is on the second syllable, so the failure of *ë* to elide is not attributable to it being stressed.

Clusters without elision from Newmark (1998) are shown in tables 2 and 3. The major constraints on consonant clusters on this chart are indicated by shading, as noted in the key to the table; these constraints are discussed in more detail below.

	p	t	q /c/	k	c /t͡s/	ç /t͡ʃ/	f	th /θ/	s	sh /ʃ/	b	d	gj /ɟ/	g	x /d͡z/	xh /d͡ʒ/	v	dh /ð/	z	zh /ʒ/
p									sp	shp										
t	*(pt)*						ft		st	sht										
q						çq	fq		sq	shq										
k		*tk*				çk			sk	shk										
c																				
ç																				
f					cf	çf			sf	shf										
th				*kth*						*shth*										
s	ps			*ks*																
sh	psh						fsh													
b						*çb*													zb	zhb
d						*çd*											*vd*		*zd*	*zhd*
gj																	*vgj*		*zgj*	*zhgj*
g																			*zg*	*zhg*
x																				
xh																				
v										*(shv)*		*(dv)*							zv	zhv
dh														*gdh*						*(zhdh)*
z																				
zh																				
j	pj	tj		kj	cj	çj	fj	thj	sj		bj	dj					vj	dhj	zj	
r	pr	tr		kr	cr	çr	fr	*thr*		shr?	br	dr		gr			vr	*dhr*		
rr	*prr*			*krr*		*çrr*								*grr*			*(vrr)*			
l	pl			kl	cl	çl	fl		sl	shl	bl			gl			vl			
ll	pll			kll			fll		*sll*	*(shll)*	bll			gll			vll		*(zll)*	
m		tm		*km*		çm			sm	shm									zm	
n	*(pn)*					çn		*(thn)*	*(sn)*	shn				gn						
nj						çnj														
h				*(kh)*		*(çh)*														

Table 2. Clusters where C_1 is an obstruent

Notes to Table 2:
- Underlying clusters from Newmark 1998 not marked as nonstandard.
- *Italicized* clusters not in Stefanllari 1996.
- *(Italicized and parenthesized)* clusters in only one word in Newmark 1998 and no words in Stefanllari 1996, counted as marginal; see Appendix 1 for list of marginal words.
- Only with a morpheme boundary between the consonants: *çb, çf, çh, çn, çnj, çq, çr, çrr, shth, zhb, zhdh, mv.*

Key to shading:

No obstruent clusters with voicing mismatch (exceptions: *çb, çd,* marginal *shv*); No clusters of two voiced stops	C1 not *x, xh, q, gj*; C1 not oral sonorant (exception: *rrj*); no clusters of *th, dh* + obstruent
No geminates	C2 not *c, ç, xh*; no clusters of obstruent + *x*

	j	r [ɾ]	rr [r]	l	ll [ɫ]	m	n [n,ɲ,ŋ]	h
p						mp		
t								
q								
k							*nk*	
c								
ç								
f								
th								
s								
sh								
b						mb		
d							nd	
gj							ngj	
g							ng	
x							nx	
xh								
v						*mv*		
dh								
z								
zh								
j			*rrj*			mj		
r						mr		*(hr)*
rr						*mrr*		
l						*ml*		
ll						*mll*		
m								
n								
nj								
h								

Table 3. Clusters where C_1 is a sonorant

2.2 Clusters without elision: overall patterns

Geminate consonants do not occur word-initially in Albanian. (In fact, Albanian does not have geminates word-internally either; orthographic *ll* and *rr* represent /ɫ/ and /r/ respectively.)

There are several consonants that do not begin clusters in Standard Albanian as represented in Newmark (1998): the voiced affricates *x, xh* /d͡z, d͡ʒ/ and the palatal stops *q, gj* /c, ɟ/. There are fairly plain historical explanations for at least some of these. The palatal stops have several sources. They developed by palatalization of pre-Albanian **k, *g* before **j* and front vowels, and from Proto-Albanian **kl, *gj*.[6] Additionally, *gj* developed from Proto-Indo-European **s* before a stressed vowel, and from pre-Albanian **j* (Demiraj 1996:196–200; Beekes 1995:261–263). Since *q/gj* always developed before a vowel (or before a glide which was then lost), it was never in a position to be followed by another consonant.

---

[6] The development of *q, gj* from **kl, gj*, is relatively recent, since some outlying dialects still have clusters: Standard Albanian *gjuhe* 'language' = *gluhë* in an Arvanitika dialect in Greece (Demiraj 1996:198).

The palato-alveolar affricate *xh* [d͡ʒ] has a special status in Albanian. It occurs primarily in loanwords from Turkish, as well as in some loanwords from other languages (including English), and some sound-symbolic forms (Curtis 2008). This sound would not begin a cluster in loanwords from Turkish, as the source sound in Turkish could not be the first member of an initial cluster, since clustering in Turkish is very limited. This does not fully explain why it does not begin clusters in sound-symbolic words. However, if *xh* at one time existed only in loanwords from Turkish, and was afterwards employed for sound-symbolic use, then it would first have become established in the language as a segment that occurred word-initially before vowels but not before other consonants; this may have become a phonotactic constraint that sound-symbolic forms adhered to when they were introduced.

The historical reasons that these consonants cannot begin clusters do not mean that there are no synchronic phonological reasons. The historical developments gave rise to a state of the language in which initial clusters beginning with these consonants do not appear. When a language learner is presented with the data of Albanian, there is nothing to encourage the learner to posit the possibility of such clusters. They are not sporadic gaps in an otherwise full range of possible clusters, but a consistently absent category. Given this consistent absence, it is plausible that speakers would exclude it from their grammars. The hypothesis in the preceding paragraph about the reason for the absence of *xh*-initial clusters in sound-symbolic forms is an example of how this could apply.

In addition to consonants that cannot begin clusters, there are also consonants that cannot end clusters. Specifically, there are no clusters ending with three of the four affricates: *ç, c, xh* /t͡ʃ, t͡s, d͡ʒ/ (the fourth affricate, *x* /d͡z/ is found after *n*, as mentioned later).

2.3   Obstruent-obstruent clusters

In clusters of two obstruents, voicing mismatches are avoided. There are no cases of a voiced obstruent followed by a voiceless obstruent, and clusters of a voiceless obstruent followed by a voiced obstruent are very restricted. Newmark (1998) gives the clusters *çb* [t͡ʃb] (e.g., *çbind* 'dissuade'), *çd* [t͡ʃd] (e.g., *çdo* 'every'), and marginally, *shv* [ʃv], which occurs in the word *shvenk* 'flash-pan' (a cinematographic term), which appears to be a borrowing from German *Schwenk* 'pan' (also in a film context; Messinger et al. 1993). This is the first case where *ç* [t͡ʃ] seems to combine more freely than other consonants. Newmark et al. (1982:19) indicate that this may be a place where orthography is misleading; they state that the negative prefix *ç* becomes *zh* before a voiced consonant (e.g., *ç-* + *duk* = *zhduk* 'cause to disappear'). However, in the examples they give, the assimilated form is spelt with a *zh-*, leaving a mystery of whether the few forms spelt with *çd-* and *çb-* are also assimilated or not.[7] This question will be investigated further in the phonetic study below.

Clusters of two stops do occur, but they are very restricted. The only such clusters consist of two voiceless stops; the cluster *tk* occurs in *tkurrje* 'contraction' and a few

---

[7] While *çboj* and *çbind* have the negative prefix in question, *çdo* does not, but rather a morpheme meaning 'what'; this may not follow the same morphophonemic rules as the negative prefix.

other words from the same root, and *pt* occurs marginally in the interjection *ptu* 'stylized spitting to represent spite/contempt for someone'. In addition to the consonants discussed earlier, obstruent-obstruent clusters cannot begin with the dental fricatives *th, dh* /θ, ð/, and cannot end with *x* /$\widehat{\text{dz}}$/.

Some of the clusters described in this section, particularly the clusters with voicing mismatch and those with two stops, raise questions about underlying and surface forms, and the relationship of orthography to these. Assuming that these spellings represent some level of phonological reality, which level do they represent? Do they represent an underlying level, or something close to it, with phonological processes affecting the cluster so that it surfaces in a different form, such as [dʒb] or [kət]? Or do they represent a surface level, aligning well with the phonetic realization of these clusters? Does this have any relation to broader tendencies in the relationship between writing and speech? The question of the phonetic realization of these clusters will be taken up below in section 3.2.

2.4   Obstruent-sonorant clusters

There are fewer clearly definable restrictions on obstruent-sonorant clusters than on obstruent-obstruent or sonorant-sonorant clusters: most sonorants (especially *j, r* /ɾ/*, l, ll* /ɫ/*, m*) cluster quite freely after a wide range of obstruents. Three of the four liquids, *rr* /r/*, l, ll* /ɫ/*,*[8] do not follow most coronals, but there are exceptions: the clusters *çrr, çl, sll, shl, shll, zll* [t͡ʃr, t͡ʃl, sɫ, ʃɫ, zɫ] do occur.

The palatal nasal *nj* [ɲ] occurs only after *ç* [t͡ʃ] (e.g., *çnjerëzor* 'inhuman') — once again, *ç* patterns more freely than other consonants. In all cases of *çnj*, the *ç* is a separate morpheme, a negative prefix. The ability of *ç* to cluster more freely than most consonants, (and especially than other affricates), combined with the fact that in these cases the *ç* is a prefix, raises questions about how the phonology of morphologically complex words relates to that of morphologically simple words.

2.5   *h* in clusters

Generally, *h* does not occur in clusters either as the first or second member. However, Newmark et al. list a few marginal cases. As a first member it occurs marginally in the cluster *hr*, in the clearly borrowed word *hrushovian* 'Khrushchevian'.[9] As a second member it occurs marginally in the cluster *kh* the onomatopoeic word *khu-khu* 'coughing sound of someone choking'[10] and in the linguistic term *çhundorëzim* 'denasalization' (in which the *ç* is a negative prefix.)

---

[8] The fourth liquid is *r* [ɾ].

[9] The speaker I worked with did not know this word; it would be interesting to find out how it is pronounced by speakers who do use this word—is it [hɾ], [xɾ], [r̥], or something else?

[10] The speaker I worked with knew this word but said the initial consonant was simply a *k*, and would spell it accordingly. More investigation would be needed to find out whether other Albanian speakers pronounce it with an onset other than [k], whether a cluster or fricative [x] or something else.

2.6   Sonorant-initial clusters

The only cluster beginning with an oral sonorant in Standard Albanian[11] is *rrj* [rj] (e.g., *rrjedhim* 'result').[12] There are clusters of *m* followed by all of the oral sonorants. The other nasals, *n* and *nj* [ɲ], however, are not followed by oral sonorants. There are no clusters of two nasals in Standard Albanian[13]. There are also clusters of nasal + obstruent, including voiceless obstruents. These include *mv* (the only cluster of nasal+fricative) in e.g. *mvehtësi* 'independence, individuality', and (what appear to be) clusters of a nasal followed by a homorganic stop or affricate, including some cases with voiceless stops, such as *mp* in *mposht* 'defeat'. These nasal-stop clusters can occur as part of larger clusters, such as *zmbr* [zmbɾ] in *zmbraps* 'repel', and *çmpl* [t͡ʃmpl] in *çmpleks* 'untwine'. Such clusters raise questions about syllabicity and the role of sonority in syllabification: are clusters such as *mp* and *mv* onsets, or do they involve a syllabic nasal? If such sonorant-obstruent clusters are onsets, what about obstruent-sonorant-obstruent-sonorant clusters such as *çmpl*; are they also onsets? If so, what does this say about the role of sonority in syllabification? In addition to questions about the syllabification of clusters such as *mb* and *ng*, there is also a question whether they are clusters at all, or whether they are in fact prenasalized stops. In the next section, I discuss some facts related to these possibilities, though I come to no firm conclusion.

2.7   #NC

There are three main possibilities for the phonological identity of what appears to be a word-initial nasal-stop cluster (NC). These possibilities are phonetically quite similar, though there may be subtle differences, but phonologically they are distinct, and are tied to differences in whether the NC is tautosyllabic or heterosyllabic, and whether it is monosegmental or bisegmental. These possibilities are outlined in (1).

(1)   Syllabic nasal+stop   bisegmental   heterosyllabic /#n̩.d/[14]
       Onset cluster   bisegmental   tautosyllabic   /#nd/
       Prenasalized stop   monosegmental tautosyllabic   /#^{n}d/

There are two ways to try to determine which of the possibilities in (1) occurs in a given case: phonetic and phonological. While the different possibilities are phonetically similar, there may be subtle differences which make it possible to distinguish between them phonetically; such a method for distinguishing clusters from prenasalized stops will be discussed below. A phonological method, instead, attempts to determine how an NC fits in to the phonological system of the language. For instance, a language might have suffix

---

[11] Other clusters beginning with oral sonorants occur in other dialects which are outside the scope of this paper; Geg, for instance, has clusters of *rrn-*, as in the title of the book *Rrno vetëm për me tregue* by Zef Pllumi (1995).

[12] It is worth noting that all instances of *rrj* were followed by the vowel *e*.

[13] Tosk dialects spoken in Greece, known as Arvanitika, show a development of *mj-* into *mnj-* in e.g. *mnjekrë* 'beard, chin', under the influence of Modern Greek (Brian Joseph, p.c.)

[14] A variant on this possibility is that the two consonants are phonetically preceded by or broken up by a vowel: [#ən.d] or [#nə.d]. Phonetic data would reveal if there is a vowel present; my phonetic study found no such evidence.

with a disyllabic allomorph that attaches monosyllabic roots and a monosyllabic allomorph that attaches to words of two or more syllables; in such a language, if roots like *ndal-* take the monosyllabic allomorph, they must be disyllabic, and the *n* must therefore be syllabic.

Riehl (2008) both details a phonological method for distinguishing the different types of NCs (not only in initial position) and finds a phonetic distinction between prenasalized stops and clusters. In Riehl's phonological method, NCs which are heterosyllabic are necessarily clusters, while NCs that are not 'separable' — whose stop component does not occur as an independent segment outside NCs — are necessarily prenasalized stops. Tautosyllabic separable NCs are considered clusters by default, and are identified as prenasalized stops only if other phonological evidence points that way — that is, 'if NC sequences appear to be treated as single segments by the phonology, in contrast to clear consonant clusters' (Riehl 2008:24). Riehl (2008:52–62) also argues that there are no cases of prenasalized voiceless stops, and that the occasional reports of them have been confounded by other factors, e.g., failing to distinguish between tautosyllabicity and monosegmentality, or analyzing phonetically voiced prenasals as voiceless to economize on the language's feature inventory.

Phonetically, Riehl (2008) finds that prenasalized voiced stops are roughly equivalent in duration to plain nasals, while NC clusters (even when tautosyllabic) are substantially longer. (This applies in particular to NCs where the C is a voiced stop; NCs where the C is an affricate or voiceless stop are substantially longer than plain nasals even when they are monosegmental (Riehl 2008:272–275).)

Albanian NCs are clearly separable: their stop/affricate components occur as independent segments; thus they are not obvious cases of prenasalized stops. According to Buchholz & Fiedler (1987:43–44), Albanian NCs are tautosyllabic, both word-initially and word-internally. For the speaker I was working with, this was not intuitively clear. The speaker found it hard to decide whether words like *mposht* 'defeat' (with initial [mp-]) and *ngjall* 'revive' (with initial [ɲɟ-]) were one syllable or two. When asked how many notes she would sing them on, she replied with no hesitation that she would sing each word on one note.[15] However, intuition may be a misleading guide to syllabification (Riehl 2008:21). Ideally there would be not only intuitions, but more directly phonological evidence indicating whether these clusters are tautosyllabic or heterosyllabic, and whether the nasal forms a syllable of its own or not. I will tentatively follow Buchholz & Fiedler and my speaker's singing syllabification and assume that Albanian NCs are tautosyllabic. This leaves them in the 'inseparable tautosyllabic' category, where by default they are to be considered clusters, though they may be considered prenasalized stops if further evidence warrants. Further investigation would be needed to determine whether phonological evidence exists. Phonetic evidence about

---

[15] The speaker's syllabification judgments were sometimes problematic in other places; e.g., when first asked how many syllables there were in *mbiemri* 'adjective, surname', she said that there were two: *mbi* and *emri*. In these cases, asking her how many notes she would sing them on gave a more expected syllabification; however, it is possible that this reflects conventional singing patterns rather than directly representing phonology. (On a later occasion, asked again how many syllables there were in *mbiemri*, she responded that there were 3: *mbi*, *em*, and *ri*.)

the duration of NCs relative to nasals would also be informative, but is not incorporated into the present study.

2.8　　Clusters with elision

	p	t	q	k	c	ç	f	th	s	sh	b	d	gj	g	x	xh	v	dh	z	zh
p																				
t				kët			fët													
q																				
k																				
c																				
ç																				
f																				
th														gëth						
s	pës			kës																
sh	pësh			kësh																
b																				
d																				
gj																				
g																				
x																				
xh																				
v																				
dh																				
z																				
zh														gëzh						
j																	vëj			
r																	vër			
rr	përr			kërr							bërr	dërr		gërr			vërr			
l		tël		kël																
ll	pëll			këll			fëll			shëll		dël					vëll			
m				këm			fëm													
n				kën																
nj																				
h																				

Table 4. Clusters where C₁ is an obstruent

- Clusters formed by elision from Newmark 1998 not marked as nonstandard

Key to shading (same as table 2, except clusters starting with oral sonorants not shaded):

No obstruent clusters with voicing mismatch (exceptions: *çb, çd,* marginal *shv*); No clusters of two voiced stops		C1 not *x, xh, q, gj*; no clusters of *th, dh* + obstruent	
No geminates		C2 not *c, ç, xh*; no clusters of obstruent + *x*	

	j	r	rr	l	ll	m	n	nj	h
p						mëp			
t						mët			
q						mëq			
k				lëk		mëk			
c									
ç									
f									
th									
s						mës			
sh						mësh			
b			rrëb	lëb					
d						mëd			
gj						mëgj			
g						mëg			
x									
xh									
v						mëv			
dh				lëv		mëdh			
z						mëz			
zh						mëzh			
j	■								
r		■				mër			
rr			■			mërr			
l				■		mël			
ll					■	mëll			
m				lëm		■			
n				lën		mën	■		
nj						mënj		■	
h						mëh			■

Table 5. Clusters where C_1 is a sonorant

In addition to words beginning with orthographic consonant clusters, there are words where consonant clusters may be formed if orthographic schwa (*ë*) is not pronounced. Newmark (1998) italicizes *ë* in some words to indicate that it may be unpronounced. In some words, a schwa is italicized between a word-initial consonant and another consonant. In these cases, if the vowel is omitted, a consonant cluster would result. For example, Newmark (1998) gives *mësim* 'education'; if the *ë* is omitted, the result would be an initial [ms] cluster. In the following discussion, I parenthesize cases of *ë* that Newmark italicizes. Table 4 shows all the initial clusters formed by elision according to Newmark (1998). This differs from Table 2 in the following significant ways. First, there are a wider range of clusters beginning with oral sonorants—mostly *l*, e.g. in *l(ë)vere* 'rag', *l(ë)mosh(ë)* 'alms', *l(ë)bardh* 'give a white appearance', but also *rr* in *rr(ë)byth* 'force backwards'. (Consonant clusters still do not begin with *j, r,* and *ll.*) There are no clusters beginning with any of the nasals except *m*, but *m* can begin a cluster followed by almost any consonant, with the exceptions of the affricates (*c, ç, x, xh* /t͡s, t͡ʃ, d͡z, d͡ʒ/), which do not participate in any clusters formed by elision, and *f, th, b, j*. The consonants *s*, *z* and *zh* do not begin clusters with elision in Newmark, and there are almost no clusters with two voiced obstruents (the only exception is *g(ë)zh-* in e.g. *g(ë)zhoj(ë)* 'shell').

Schwa elision raises questions about the relevance for phonological description and analysis of formal or careful speech pronunciations vs. informal or fast speech pronunciations; in some sense, careful speech forms seem more basic and fundamental, with fast speech forms being secondary and derived from them. Yet, it seems probable that people hear far more fast and informal speech than careful and formal speech, and so people must often learn words from informal speech rather than from formal speech, making informal speech in a sense more basic at least for certain types of things. What role does each of these types of speech play in a speaker's knowledge of their language, and thus how should it be taken into account in phonological analysis?

3.   Phonetic study
3.1   Speaker

The study involved a single speaker who was female, in her early 30s, and came from Prishtina in Kosovo. She spoke both Geg and Geg-influenced Standard Albanian; she reported that her parents spoke to her in Standard Albanian and Standard Albanian was the language of her schooling. This study investigated only her Standard Albanian. Geg influence was evident in the fact that she did not distinguish *r* and *rr* (standardly [ɾ] and [r]), and merged the palatal stops *q* and *gj* (standardly [c] and [ɟ]) with the palato-alveolar affricates *ç* and *xh* ([t͡ʃ] and [d͡ʒ]).

3.2   Stimuli and recording process

Based on dictionary work and elicitation sessions, 84 words were selected to be recorded. The words discussed here included 17 words with orthographic obstruent-obstruent clusters, and 20 words with *ë* between the first two consonants (which might be elided); 15 words with initial NCs (e.g., *ndal*), 15 with initial stops (e.g., *dal*), and 11 with initial nasals (e.g., *nam*) were recorded in order to investigate NCs, but results from these are not discussed here, as I decided after recording that the stimuli and recording conditions were not sufficiently controlled to reliably measure timing. 6 words with other clusters were recorded but not used in analysis. The obstruent-obstruent clusters tested in this study were selected as follows. For a complete list of words, see Appendix 2; for a list of words recorded but not analyzed, see Appendix 3.

(2)   *çb, çd, çk, çf* (x2)   ([t͡ʃb, t͡ʃd, t͡ʃk, t͡ʃf])
        These clusters represent all the available clusters of an affricate followed by a consonant. Newmark has examples with *cf* [t͡sf] and *çq,* [t͡ʃc] but the speaker in this study was not familiar with these words.

(3)   *shp, sht, shk, shf*   ([ʃp, ʃt, ʃk, ʃf])
        These clusters are equivalents of those in (2) with a voiceless fricative in place of the affricate, and voiceless consonants where those in (2) had voiced consonants.

(4)   *zhb, zhd, zhg, zhv*   ([ʒb, ʒd, ʒg, ʒv])
        These clusters are the voiced equivalents of those in (3); thus, in relation to the clusters in (2) they have voiced fricatives in place of the affricates, and voiced second members where those in (2) had voiceless ones.

(5)   *vd, tk, kf, gdh*   ([vd, tk, kf, gð])
These clusters were chosen for a variety of reasons: *tk* is the only stop-stop cluster familiar to my speaker, and I reasoned that this was one of the most likely contexts for epenthesis;[16] *vd* and *gdh* were the only all-voiced obstruent clusters familiar to my speaker that did not begin with *z* or *zh*;[17] *kf* does not occur in Newmark (1998), but was found by elicitation in *KFOR-i* 'Kosovo Forces'.

Obstruent-obstruent clusters that were not tested include clusters beginning with *s* (*sp, st, sq, sk, sf*), *z* (*zb, zd, zgj, zg, zv*), *f* (*ft, fq, fsh*), *k* (*kth, ks*) and *p* (*ps, psh*).

The 20 words chosen to test the possibility of elision included some in which Newmark indicated that *ë* could be elided, and some in which he did not; they also included some in which pre-recording elicitation with the speaker indicated that she thought she would pronounce it (at least sometimes) with an initial consonant cluster, and some in which she did not. The two classifications did not line up: there were words where Newmark indicated elision and the speaker did not, and vice versa. For a full list of words and details about whether Newmark and the speaker indicated elision for each word, see Table 6 below.

The speaker was asked to say these words in frames designed to elicit various rates of speech, as given in (6) and (7), where _____ represents the slot into which the speaker was to insert the word. (Note the presence of an isolated instance of the target word before the instances in sentence contexts.)

(6)   'them' frame
Albanian

_____.

Tani do të them _____ përsëri.
Tani do të them _____ shpejt.
Tani do të them _____ ngadalë.
Translation

_____.   ('isolated repetition')
Now I'll say _____ again.   ('again repetition')
Now I'll say _____ quickly.   ('fast repetition')
Now I'll say _____ slowly.   ('slow repetition')

(7)   'përsërisë' frame[18]
Albanian

_____.
Tani do të përsërisë _____ përsëri.
Tani do të përsërisë _____ shpejt.
Tani do të përsërisë _____ ngadalë.

---

[16] Newmark (1998) gives an interjection with *pt*; to my speaker this cluster did not seem like a possible word beginning.

[17] Newmark (1998) gives words with *dv* and *vgj*, but these clusters were not familiar to my speaker.

[18] The final *ë* in *përsërisë* was rarely pronounced.

Translation
_____.   ('isolated repetition')
Now I'll repeat _____ again.   ('again repetition')
Now I'll repeat _____ quickly.   ('fast repetition')
Now I'll repeat _____ slowly.   ('slow repetition')

The frames and the target words were presented on separate pieces of paper. The target words were organized in 8 groups of 10 or 11 words each, with two groups on each piece of paper. Each group of words was recorded into a separate file, using the computer program Praat (Boersma 2008). Before recording, the speaker was instructed to do what she said she would do when saying the words in the frames. This was effective in eliciting different speech rates.

3.3   Results: Do obstruent-obstruent sequences have epenthetic vowels?

The first question investigated by this study was whether epenthetic vowels break up sequences of two orthographic obstruents. I found that this occurs very rarely, and did not find any contexts where it happened consistently.

As discussed above, 17 words with orthographic obstruent-obstruent sequences were recorded. In investigating this question, I have only analyzed those in the 'them' frame, due to the difficulty in distinguishing between the final *s* in *përsërisë* and an initial fricative in the target word; thus, 4 repetitions of each word are analyzed here, giving a total of 68 tokens. Among these were two tokens with clear epenthetic vowels, listed in (8).

(8)   Tokens with clear epenthetic vowels
        One token of [t͡ʃəb] in *çboj* 'I undo'[19] in isolation
        One token of [ʒəv] in *zhvillim* 'development' in isolation

Phonetically, there is no clear-cut division between a long release and a brief vowel. A short release may be clearly a release, and a non-brief vowel may be clearly a vowel, but the area in between is gradient and not categorical. Among the words recorded, I found two tokens with a release that comes close to being a vowel, listed in (9).

(9)   Tokens with near-vowel releases
        çboj in *Tani do të them 'çboj' përsëri* 'now I'll say *çboj* again'
        vdekja 'the death' in *Tani do të them 'vdekja' ngadalë* 'now I'll say *vdekja* slowly'[20]

---

[19] For this word, Newmark (1998) gives standard *zhbën* (3sg; 1sg = *zhbëj*) and nonstandard *çbën* (3sg; 1sg = *çbëj*); the speaker's *çboj* (1sg; 2sg *çbon*) is not listed, but I have used it as an example of the cluster *çb* which does occur in other standard words in Newmark, particularly *çbojatis* 'discolor' and *çbind* 'dissuade'.

[20] Kelly Maynard has called to my attention that the near-vowel release in *vdekja* may be due to the fact that the speaker is from Kosovo, since in Kosovo dialects, this word begins with plain *d*, so *vd* may be an unfamiliar cluster.

Spectrograms of *çboj* said without epenthesis, with an epenthetic vowel, and with a release that is almost a vowel are given in 10–12.
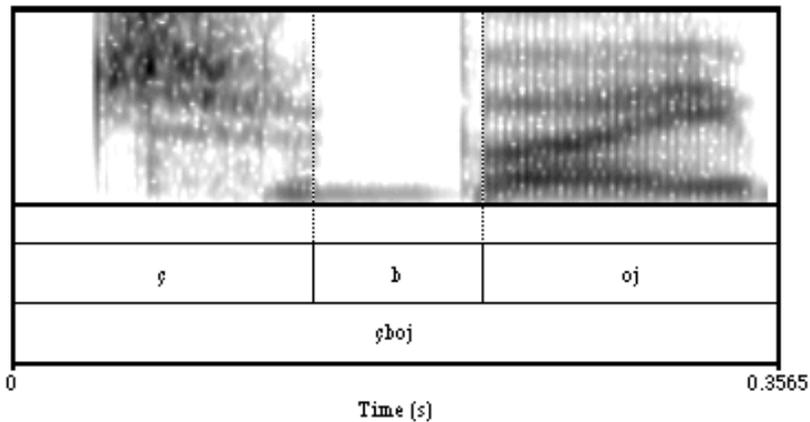
(10)   *çboj* with voiceless *ç* [t͡ʃ], voiced *b*, and no epenthetic vowel:
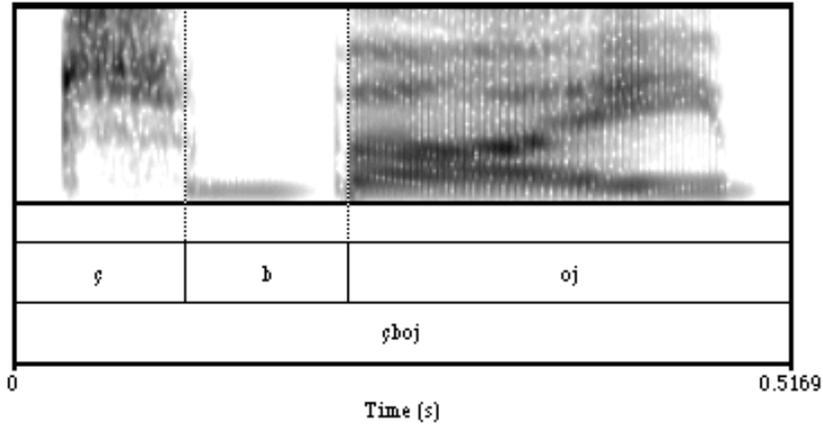
(11)   *çboj* with an epenthetic vowel:
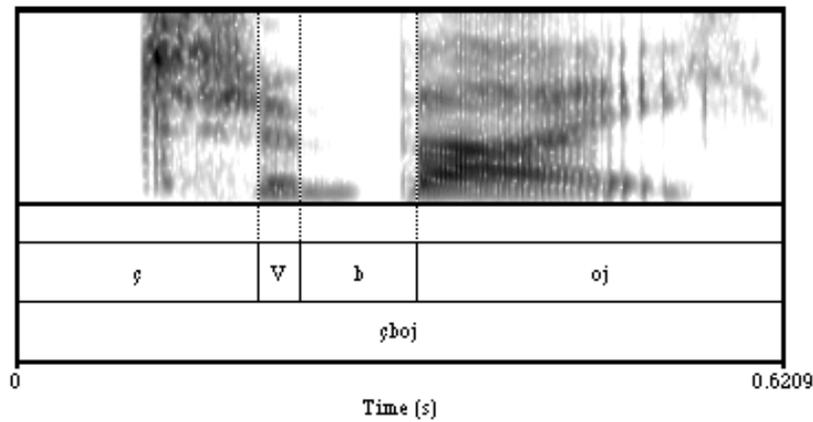
(12)   *çboj* with a release that is almost a vowel:

Although there were a few cases of epenthesis in words said in isolation, it is absent in the vast majority of cases, even in isolation. Although not all types of obstruent-

obstruent clusters were tested, the results from those that were tested suggest that epenthesis in obstruent-obstruent clusters is uncommon, and that the occasional instances are more likely to happen in slow or careful speech than in casual or fast speech.

3.4   Results: Can the existence of clusters formed by elision be confirmed acoustically?

The second question investigated in this study is whether it is possible to confirm acoustically that elision of schwa produces clusterswhich are not indicated orthographically.[21] The data showed that this does take place, but less than anticipated.

Pre-recording elicitation had determined a number of words with *ë* in the first syllable, some of which the speaker said she would say with an initial cluster (without the *ë*) in spoken Albanian, and some of which she indicated that she would not omit the vowel in.[22] These data did not align very well with what was found in recordings. Clusters were not found in any cases where the speaker said she would not have them, and were also only occasionally found in cases where she said she would have them in spoken Albanian. All cases of elision occurred in fast repetitions. 20 words were investigated for elision. Of the 40 tokens in fast repetitions, one was disfluent, leaving 39 fluent tokens, 22 of which belonged to words that the speaker indicated elision in, and a further 11 of which Newmark indicated elision in, while 6 tokens belonged to words that neither Newmark nor the speaker indicated elision in.

Elision occurred once in the frame *tani do të them ___ shpejt* (with the word *dëlirë* 'pure') and 3 times in the frame *tani do të përsërisë ___ shpejt* (with the words *mësuesja* 'the teacher', *bërryl* 'elbow', and *mësim* 'education'). A spectrogram of *mësim* with elision is given below in (13). There was also one case where not only the vowel, but also the initial consonant was omitted in the word *kështjella* 'the castle'. Thus, out of the 22 fast speech tokens of words where the speaker indicated that she would elide, less than one-fourth showed elision. In addition, there were 3 tokens where *ë* was very reduced but not entirely gone. A full chart of cases with and without elision is given in Table 6; this table also indicates whether the speaker said that she would elide the vowel, and whether Newmark (1998) indicates that elision is possible.

---

[21] Strictly speaking, this study does not prove that these cases are elision, since it does not prove that schwa is underlyingly present phonologically. However, the fact (discussed below) that it is usually present strongly suggests that it is phonologically present.

[22] The speaker was asked questions like 'do you know any words that start with *kt*?', and often volunteered words that she said had the particular cluster in question in spoken Albanian, but were written with an intervening *ë* in written Albanian. On occasion she was asked particularly how she would pronounce a given word with *ë* in the first (orthographic) syllable, and would say it and indicate whether she would omit the vowel or not.

Albanian	gloss	Tani do të them ___ shpejt	Tani do të përsërisë ___ shpejt	Speaker said would elide?	Newmark shows elision?
kësaj	this (infl.)	x	x	(not asked)	yes
këmishë	shirt	x	x	would	yes
mëshova	I did put pressure on	x	x	would	yes
kënetë	marsh	x	x	wouldn't	yes
këtu	here	x	x	would	yes
mëtejshëm	later	x	x	(not asked)	yes
gëzim	joy	x	x	would	no
dënim	punishment	x	x	wouldn't	no
gjëmim	thundering sound	x	x	wouldn't	no
lëkundje	shaking	x	x	?	no
gëzuar	joyous	x	x	would	no
kështjella	the castle	x	ELISION; k also elided	would	no
kësulkuqe	Little Red Riding Hood	x	x	would	no
lëmoshë	alms	x	x	?	yes
mësuesja	the teacher	x	ELISION	would	yes
gëzhojë	shell	x	(very reduced)	would	yes
bërryl	elbow	x	ELISION	(not asked)	yes
mësim	education	(very reduced)	ELISION	would	yes
mëditje	daily wages	(very reduced)	(disfluency; no data)	wouldn't	yes
dëlirë	pure	ELISION	x	would	yes

Table 6. Elision data: Elision in fast speech; words listed in the order they were recorded. Note increasing elision towards the end of the recording session.

The fact that there were more elisions in the *përsërisë* frame than in the *them* frame is probably due to several factors that likely increased speech rate. First, since *përsërisë* is longer than *them*, the whole frame was longer, which could contribute to its components being pronounced faster. Second, the *përsërisë* repetitions followed the *shpejt* repetitions, and the speaker's rate of speech increased over the course of the task.

While the phonetic data provide evidence that *ë* may sometimes be elided, they do not provide evidence that it is usually elided in some words. This could be due to one of two things. First, it may be that (at least for this speaker), *ë* is not in fact elided in these words, despite speaker intuitions to the contrary. Second, it may be that the task did not elicit the speech registers in which *ë* is elided. For example, since the task depended on writing, it may be that even at relatively fast speech rates, an orthographically-based pronunciation was elicited, rather than a pronunciation that would be used in everyday speech.

(13)   *mësim* with elision of schwa
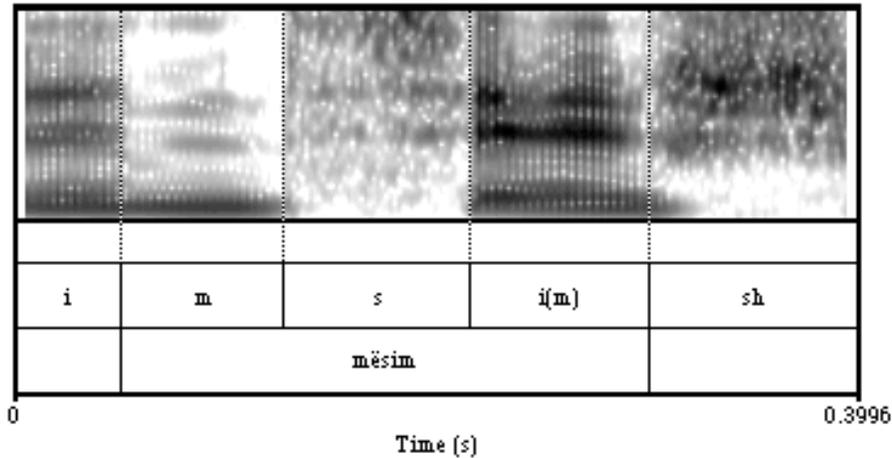
It is worth drawing attention to the fact that while epenthesis sometimes occurred in orthographic clusters, and a vowel was sometimes elided in words with *ë* between the first two consonants, the cases with orthographic clusters usually did not show epenthesis, and the cases where *ë* was in a position to be elided usually did not show epenthesis. This gives evidence that the orthographic presence or absence of *ë* between the first two consonants does in fact reflect a phonological reality.

3.5   Results: Are *çb, çd* phonetically voiceless followed by voiced [t͡ʃb, t͡ʃd] as spelling suggests?

The third question investigated in this study is whether Albanian does in fact have clusters of a voiceless obstruent followed by a voiced obstruent—specifically, are *çb* and *çd* phonetically [t͡ʃb, t͡ʃd] as the spelling suggests, or not? The phonetic evidence showed that they can be phonetically voiceless + voiced, and are in several tokens, but that they are not always.

The initial *ç* is generally pronounced voicelessly, as [t͡ʃ]; however, in the fast repetition of both *çboj* and *çdo* after *përsërisë*, some assimilation occurs: the final *s* of *përsërisë* is pronounced as voiced [z], and the initial *ç* of the following word is pronounced as [ʒ], producing a cluster across the word boundary of [z#ʒ].[23]

The second consonant in the cluster, *b* or *d* was sometimes voiced, but was voiceless in many tokens. It is worth noting that this was true also of the voiced stops in the 'all-voiced' clusters *zhb, zhd,* and *zhg*. A spectrogram of *çboj* with the initial cluster pronounced [t͡ʃb] has already been given in 10.

4.   Conclusion

The rich, largely unexplored clustering possibilities of Albanian consonants provide many interesting questions and avenues for investigation. In this paper I have done

---

[23] The final *ë* in *përsërisë* was elided in these cases, as in most cases at all speech rates.

necessary groundwork for the investigation of Albanian word-initial consonant clusters. I have not only provided a list of clusters, but have determined some of the constraints on clusters not formed by elision, which I summarize as follows:

- Clusters of voiced obstruent followed by voiceless obstruent are not permitted.
- Clusters may not begin with palatal stops (*q, gj*) or voiced affricates (*x, xh*).
- Affricates (*c, ç, x, xh*) may not be second members, except that *nx* is allowed.
- Obstruent-obstruent clusters may not begin with dental fricatives (*th, dh*).
- The only voiceless obstruent that may precede a voiced obstruent is *ç*.
- The only consonant that may precede *nj* is *ç*.
- *rr, l, ll* do not follow most coronals, but there are exceptions.
- *h* does not normally participate in clusters.

Based on phonetic data, I have determined several things. First, obstruent-obstruent clusters are not normally broken up by an epenthetic vowel, though they occasionally are, Second, elision may form consonant-consonant clusters that are not orthographically indicated; however this was observed less than anticipated, possibly due to the effects of reading pronunciation even in fast speech. Third, the clusters *çb* and *çd* may be pronounced with a voicing mismatch, with a voiceless affricate followed by a voiced consonant, but sometimes assimilation takes place and they are either all voiced (in certain fast speech conditions) or all voiceless.

Further research is needed in several areas. Phonetic and phonological motivations for the clustering constraints that I have described should be investigated. Determining when and under what conditions schwa-elision occurs requires work with more speakers and probably a more natural elicitation task — or a range of types of elicitation in order to determine what types of language use promote and inhibit elision. In addition, research needs to be extended beyond the initial two consonants into clusters of three and four consonants; for example, are the constraints on C_1C_2 the same in three and four consonant clusters as in two-consonant clusters? Are the constraints on C_2C_3 at all related to the constraints on C_1C_2? A conclusive determination of whether *mp, mb* etc. represent clusters or prenasalized stops also requires further work; Riehl's (2008) phonetic criteria provide a promising method of investigation; combining phonetic and phonological investigation could help both to determine the status of NCs in Albanian and to test whether Riehl's findings are supported in languages besides the ones she studied.

Appendix 1: Newmark's (1998) entries for words counted as marginal in 0

pn:	**pneumoni**ˈ**a** *nf* [*Med*] pneumonia
pt:	**ptu** *interj* stylized spitting to represent spite/contempt for someone
kh:	**khu-khu** *onomat* coughing sound of someone choking
çh:	**ç**|**hund**|**or**|*ë***z**|**i**ˈ**m** *nm* [*Ling*] denasalization
shv:	**shvenk** *nm* [*Cine*] Swiss-pan, flash-pan
dv:	**dv***ier•* *vt* = **degjener***o*ˈ*•n*
vrr:	**vrro**ˈ**m***ë* *nf* slut, slattern = **mola**ˈ**r***e*
zll:	**zllo**ˈ**ti** *nm* zloty (Polish money)
zhdh:	**zh**|**dhjam**|**o**ˈ**s•** *vt* to take the (excessive) fat off of []
hr:	**hrushovia**ˈ**n** *nm* Khrushchevian
thn:	**thneg***ë***l** *nf* [*Entom*] ant

Appendix 2: Stimulus words: obstruent-obstruent clusters

(Abbreviated definitions based on Newmark 1998)
Obstruent-obstruent clusters: No elision

çboj	'undo' (nonstandard; standard *zhbëj*)
çdo	'every'
çka	'what'
çfarë	'what'
çfejoj	'make X break off X's engagement' (nonstandard; standard *shfejoj*)
zhbllokoj	'unblock'
zhdukje	'disappearance'
zhgunë[24]	a kind of white wool fabric; a cloak made of this fabric
zhvillim	'development'
shpoj	'pierce'
shtoj	'increase'
shkoj	'go'
shfajësoj	'exonerate'
vdekja	'the death'
tkurrje	'contraction'
KFOR-i	'Kosovo Forces'
gdhiu	'it dawned'

---

[24] Newmark (1998) spells this word *zhgun*; the spelling above was offered by the speaker in my study.

Appendix 3: Stimulus words not analyzed in this paper

Sonorant-obstruent clusters		*Stop/affricate*		*Nasal*	
mposht	'defeat'	posht	'low'	mos	'don't'
mposhti	'defeat'(infl.)	posti	'the post'	mosha	'the age'
mplaket	'he gets old'	plaket	'he gets old'	mjetet	'the tools'
mbiemri	'the surname'	bileta	'the ticket'	minuta	'the minute'
mbështetje	'support'	bërthama	'the kernel'	mërzitem	'get bored'
mbledhje	'meeting'				
ndal	'stop'	dal	'go out'	nam	'reputation'
ndoshta	'maybe'	dosja	'the file'	nofka	'the nickname'
nxënës	'pupil'	xixa	'the spark'	nëna	'the mother'
nxehje	'heating'	xixat	'the sparks'	nënat	'the mothers'
				nisje	'departure'
ngadalë	'slowly'	gazetë	'newspaper'	natyrë	'nature'
ngushëllime	'condolences'	gazolina	'the gasoline'		
		gabimisht	'mistakenly'		
ngjall	'resurrect'	gjallë	'alive'		
ngjak	'in (your) blood'	gjak	'blood'		
ngjarje	'event'	gjendje	'state'		

Miscellaneous

tmerr   'terror'
çmim   'price'
mllef   'anger'
Xrxa   (Village name) The speaker seemed to say this with a syllabic *r*; when asked, she said it was probably Xërxa in written Albanian.
tlynë   According to my speaker, this was something similar to butter but not the same. This appears to be the same word that Newmark (1998) gives as t*ë*lyen 'butter', with an elidable *ë*.
msheftas   'hide-and-seek' (Nonstandard; see *mshef-* in Newmark (1998))

References

Beekes, Robert S. P. 1995. *Comparative Indo-European linguistics.* Amsterdam: John Benjamins.

Boersma, Paul & David Weenink. 2008. Praat: Doing phonetics by computer (Version 5.0.21) [Computer program]. Retrieved April 28, 2008, from http://www.praat.org/

Buchholz, Oda & Wilfried Fiedler. 1987. *Albanische Grammatik.* Leipzig: VEB.

Curtis, Matthew. 2008. *Xhorxh, xhuxhmaxhuxh*, and the *xhaxhalar*: The allolinguistic status of Albanian /xh/. Paper presented at the 16th Balkan and South Slavic Conference, May 1–4, Banff, Alberta, Canada.

Demiraj, Shaban. 1996. *Fonologjia historike e gjuhës shqipe.* Tirana: Toena.

Friedman, Victor A. *Albanian grammar*. 2006. Online publication of The Slavic and East European Language Research Centre.
Available at http://www.seelrc.org:8080/grammar/pdf/albanian_bookmarked.pdf .
Last accessed Sept. 20, 2008.

Messinger, Heinz, Gisela Türck & Helmut Willmann. 1993. *Langenscheidt compact German dictionary.* Berlin: Langenscheidt.

Moosmüller, Sylvia & Theodor Granser. 2006. The spread of Standard Albanian: An illustration based on an analysis of vowels. *Language Variation and Change* 18(2): 121–140.

Newmark, Leonard. 1998. *Albanian-English Dictionary.* Oxford: Oxford University Press.

Newmark, Leonard, Philip Hubbard & Peter Prifti. 1982. *Standard Albanian: A reference grammar for rtudents.* Stanford: Stanford University Press.

Pllumi, Zef. 1995. *Rrno vetëm për me tregue.*

Riehl, Anastasia Kay. 2008. *The phonology and phonetics of nasal obstruent sequences.* Dissertation: Cornell University.

Stefanllari, Ilo. 1996. *Albanian-English English-Albanian Dictionary* (Hippocrene Practical Dictionary). New York: Hippocrene Books.