TOPICS IN THE PHONOLOGY AND MORPHOLOGY OF TUVAN

K. DAVID HARRISON
YALE UNIVERSITY

A Dissertation submitted in partial satisfaction
of the requirments for the Doctorate of Philosophy in Linguistics.

September 1, 2000

Professor Stephen R. Anderson, Chair

Assistant Professor Abigail R. Kaun, Co-chair

Professor Louis Goldstein

TABLE OF CONTENTS

Acknowledgements

This work is a testament to the cultural wealth of the Tuvan people as embodied in their traditions, world view, and language. This wealth has been preserved and transmitted under the most adverse of conditions: decades of oppressive Soviet policy, cultural assimilation, and political and economic turmoil. Materially, most Tuvans are far from wealthy; a good number of them make a subsistence living as pastoral nomads. Yet, during my field work in 1996, 1998 and 2000, I was without fail welcomed at every village house or nomad camp I visited. I was provided with food, shelter, and the opportunity to participate in daily activities of the nomads' camps. More importantly, my Tuvan friends shared their intellectual and cultural treasures with me most generously and without reservation. This is a debt I feel I can never repay.

In my field work and travels in Siberia and Mongolia in 1996-2000 I have been financially supported by the Department of Linguistics, the Yale Graduate School, The International Exchanges and Research Board (IREX), The National Geographic Society, The Silk Road Project, and Volkswagen-Stiftung. In Tuva, I received logistical support from the Tuvan Institute for Humanitarian Researches. For intellectual support and collaboration I am indebted to my colleague and co-author Greg Anderson.

In my graduate career at Yale, I have been blessed with a series of outstanding mentors and teachers who have molded my intellectual outlook and challenged me to do my best work. The first of these was Alexander Schenker, who brought me to Yale to study Slavic Philology. During my first year at Yale, I was fortunate enough to take Steve Anderson's seminar on Morphology. Here I discovered the joys of formal linguistics, to which I decided to devote my academic career. The following year, I followed Steve Anderson's advice to 'make a home for myself' in the Linguistics department. Here I have benefited from outstanding teaching and mentoring by Larry Horn, Stanley Insler, Sergei Avrutin, Hugh Stimson, and Dianne Jonas. I express my heartfelt thanks to each of them.

Most crucially for my academic career, I was fortunate to find an excellent mentor, advisor and friend in Abigail Kaun. Abigail has most generously shared her time and talents with me for these past four years. She patiently apprenticed me through the steps to becoming a professional linguist, prepared me for conferences, coached me in my writing, and introduced me to people. Above all, she took me on as a collaborator in our common research projects, giving me the confidence I needed to establish myself. This thesis is a tribute to Abigail's tireless mentoring and coaching.

The highlight of my time at Yale has been the opportunity to teach undergraduates. I am most grateful to Steve Anderson, Lettie Naigles and Joe Errington for the privilege of serving as their teaching assistant. I am also grateful to the graduate students and Linguistics faculty at SUNY-Stony Brook, who gave me the opportunity to teach a phonology seminar and sharpen my own ideas in the process. The people and activities at Haskins Laboratories have provided a stimulating intellectual environment. I am particularly grateful to Cathi Best for taking me on as research assistant, and to Doug Whalen for his mentoring in the area of endangered languages.

At a more personal level, I thank my friends at Yale and elsewhere who provided me with a social life as well as intellectual stimulation. My parents, David and Catherine Harrison can take the lion's share of credit for my academic achievements, as they taught me to be curious about the world at an early age. My siblings, Doreen and Nathan, provided me with the best possible companionship in my early years. My entire family has encouraged me through every step of my academic career, even when it was not clear where this would take me. I thank them for their continuing faith in me.

Finally, I thank my partner, Khiem Tang, for his moral and emotional support and his patience in seeing me through five years of the ups and downs of graduate school. This work is dedicated to Khiem.

Preface

This thesis sets out, in some detail, the grammatical structure of the Tuvan language. Examples given herein are nearly all taken from original data collected by the author during fieldwork in 1996, 1998, 1999 and 2000. Chapter one provides a general overview of the grammar, including phonology, morphology and syntax.

Chapter two presents an acoustic study of the unique pitch accent system. We develop a hypothesis that the phonologically salient feature is pitch *contour*. Other acoustic qualities such as quantity are determined, we argue, by the dynamics of producing the right contour.

Chapter three considers two complex phonologcial processes and their interaction. First, Tuvan is shown to have a kind of hiatus resolution in which low vowels dominate regardless of their relative position. This type was previously thought to be rare among the world's languages. Second, a process of velar deletion is shown to 'feed' hiatus resolution. Velar deletion is shown to be a robust phonological process, yet one that is blocked in a non-uniform subset of environments. Blocking of velar deletion arises, we demonstrate, to enhance recoverability of an identifiable class of 'small' morphological elements (i.e. short stems and short suffixes).

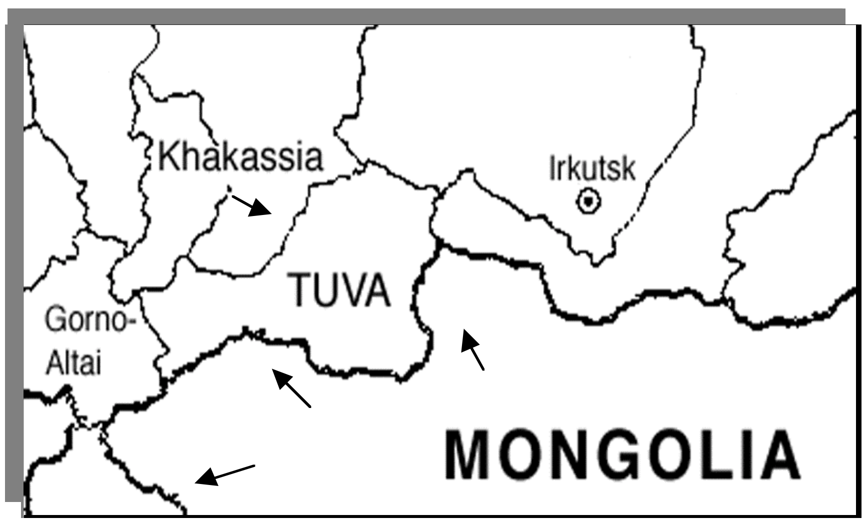
Chapter four discusses the basic patterns of vowel harmony in Tuvan and situates these in a theoretical model. In chapter five, we explore previously undocumented reduplication patterns. Further, we show the complex interaction of these with vowel harmony. The operation of harmony in reduplication yields strong evidence for the necessity of underspecification in the grammar. We argue that underspecification is a natural and expected outcome of harmony systems and is not dependent on the presence of predictable alternations. Rather, it is driven by the presence of pervasive patterns of vowel occurrence, which may or may not include surface alternations.

In the Summary, we discuss some outstanding problems and suggest areas of research that might benefit from our discoveries about Tuvan.

Chapter One: A Grammatical sketch of Tuvan

0. Introduction

The Tuvans (also Tuvinians, native ethnonym *tɨvalar*) live in the Republic of Tyva, a constituent of the Russian Federation located in southern Siberia just north of western Mongolia. Tuvan is also spoken by small communities in several administrative districts of Mongolia and Russia (denoted by arrows on map).



Tuvan speaking communities in teh Altai-Sayan region of Siberia

The Tuvan language (*tɨva dɨl*) belongs to the eastern branch of Turkic family, along with Yakut (Sakha) and Uighur. Due to the constant migrations of nomadic peoples, there has been considerable contact among divergent branches of the Turkic family. As a result, areal features are often a more significant basis for classification than are genetic features (Deny 1952). A number of areal features are shared by the Turkic languages of the Altai-Sayan region: Tofalar (now moribund), Xakas, Altai (actually a diverse group of dialects), Tuvan, and Shor. These languages may show substrate influence of the now extinct Yeniseyan language, Kamas, and they clearly show the strong influence of Mongolian, which is still widely spoken in the region.

The present work is based on field work carried out in Tuva during the period 1996 through 2000 by the author. Data was elicited from many dozens of speakers of various ages, backgrounds, and regions of Tuva as well as Tuvans in Mongolia. The material reflects the rich dialectal diversity of the contemporary spoken language. In general, we eschew a prescriptivist approach in the interest of giving a fuller view of the grammatical and stylistic possibilities of Tuvan. In this chapter, we present an outline of Tuvan grammar, based primarily on data collected in the course of our field work. Primary field sites were the Milk Lake (söt-xöl) region of central Tuva, the Mugur-Aksy region in southwestern Tuva on the Mongolian border, and Kyzyl, the capital city which is inhabited by speakers of various dialects.

1.0 Review of literature

The earliest linguistic work that mentions Tuvan (at that time called *Uriankhay*) is by V. Radloff (1882), who collected comparative lexical and phonological data from several Siberian Turkic languages. An early work by N. Katanov (1903) traces the historical proto-Turkic phonemes and their reflexes in modern Siberian Turkic. K.H. Menges (1955, 1956, 1958, 1959) also explored the historical phonology of south Siberian Turkic dialects.

The foremost scholar of Tuvan was A. A. Pal’mbax, who worked on Tuvan from the 1930’s until his death in 1963. Together with F. G. Isxakov, he published the first grammar of Tuvan (Isxakov & Pal'mbax 1961). Inspired by Pal'mbax, a number of native Tuvans have since taken up the linguistic study of their own langauge. These include A. Ch. Kunaa (1970), who investigated sentence structure; Sh. Ch. Sat (1960, 1973), who wrote on participles and on language codification; and D. A. Mongush (1958, 1983, 1987, 1995) who described Tuvan tense, aspect, modality, and auxiliary nouns. The first acoustic studies of Tuvan vowels were carried out by K. Bicheldei (1980, 1984, 1986), who in 1994 became the first freely-elected president of the Republic of Tuva.

Tuvan data has also been collected by Russian field linguists from the Novosibirsk school, for example, Cheremisina (1984, 1995); Ubryatova & Litvin (1986), Shamina 1987.

These works have been especially valuable in documenting regional, dialectal variation and non-standard, colloquial speech. The Tozha dialect, which differs most radically from other varieties of Tuvan, was documented by native Tuvan scholar Z. B. Chadamba (1974). All the above cited works (except Radloff 1882) were published in Russian. A few native Tuvan scholars have, however, published brief linguistic articles in Tuvan: Sat (1983), Mongush & Sat (1967), Bicheldei (1993).

The standardized Tuvan language was codified between 1928 and 1950. Based on a geograhically central dialect, it was close to the language spoken by most Tuvans. Dialects spoken on the periphery of Tuva differ considerably. These include the dialect of the former reindeer herding Tozha people of northeastern Tuva (Chadamba 1974), the strongly Mongolianized dialects of the southeast (e.g. Kungurtux, Erzin), and mountain dialects of Western Tuva (e.g. Mugur-Aksy region).

The earliest Tuvan alphabet was developed by a Tuvan Buddhist monk named Mongush Lopsan-Chimit. In the 1930's and 1940's a number of Tuvans, including many children, became literate in Tuvan by using primers written in this alphabet. After 1945, a modified Cyrillic alphabet was introduced, and quickly became the only accepted form of writing. Literacy is now nearly universal in Tuva, due to compulsory schooling. There are a fair number of Tuvan publications, including school textbooks, novels, collections of songs, poetry and shamanic chants, religious (Buddhist) texts, newspapers and journals. A series of Tuvan language primers and textbooks for schoolchildren of grades 1 to 10 is particularly noteworthy in this respect (e.g. Mongush and Kuular 1992). The only translations of Tuvan literary works to date are a two-volume collection of folktales in Russian,*Tuvinskoe narodnoe skazki* (Samdan 1994), and a collection of shamanic songs in English (Kenin-Lopsan 1993).

1.1 Phonology

The phonology of Tuvan will be discussed in detail in following chapters. In this introductory chapter, we provide only a segment inventory.

1.2 Vowels

(1)

		front		***back***	
		unround	***round***	***unround***	***round***
short	***high***	i	ü	ɨ	u
	non-high	e	ö	a	o
long	***high***	ii	üü	ɨɨ	uu
	non-high	ee	öö	aa	oo

In addition to eight vowel qualities and distinctive length, Tuvan vowels in word-initial syllables may have a distinctive low pitch . Chapter 3 presents detailed analysis of distinctive pitch. We indicate the presence of this feature with the IPA low pitch diacritic: [è] [ǜ]. Vowels that bear low pitch may contrast minimally with both short and long vowels.

(2)

<u>short</u>		<u>low pitch</u>		<u>long</u>	
et	'leather'	*èt*	'meat'	*eet*	'estuary'

We employ the following vowel archiphonemes herein: /I/ represents a high vowel with four surface variants: [i] [ü] [ɨ] [u]. Archiphoneme /A/ represents a non-high, unrounded vowel with surface variants [e] and [a]. The surface variation of these archiphonemes results from backness harmony and rounding harmony, which are discussed in detail in chapter 3.

1.3 Consonant phonemes [1]

(3)

	Labial	Labio-dental	Alveolar	Palatal	Velar
Plosive	p b		t d		k[2] g
Nasal	m		n		ŋ
Trill			r		
Fricatives			s z	š ž	x
Affricates				č	
Lateral			l		
Approximant		v [3]		y	

Consonants undergo predictable surface alternations. We list here the basic form and surface alternations. In addition, we employ archiphonemic symbols in this chapter. In so doing, we follow turcological convention to point out the predictable quality of the alternations. We do not mean to imply that consonants are underspecified, as we will argue for vowles in chapter 3 and 4.

<u>Basic form</u>	<u>surface forms</u>	<u>archiphoneme</u>
p	p b v m	/B/
t	t d n	/D/
k	k g	/G/
č	č ǰ	/J/
n	n t d	/N/
l	l t d n	/L/
l [4]	l n	/L/
z	s z	/Z/

/n/ surfaces as [n], [d] or [t] and may be found in the accusative suffix /-NI/

(4) *ača-nɨ* ‘father’-ACC
xam-nɨ ‘shaman’-ACC
šil-di ‘bottle’-ACC
xat-tɨ ‘wind’-ACC

/t/, which appears in the locative suffix /-DA/, shows the same alternations: [n], [d] or [t].

(5) *ača-da* ‘father’-loc
*xam-n*a ‘shaman’-loc

---

[1] Note: [f] also occurs, but only in borrowings; [c] (or [ts]) appears in a few Russian loanwords.

[2] [k] alternates predictably with [q] in back vocalic contexts.

[3] For most speakers this is an labio-dental approximant with only very weak frication or no frication. We use the symbol [v] herein.

[4] Whn [l] appears in the onset of an encltic, its range of surface alternations is limited by virtue of its location at a morpheme boundary.

	šil-de	'bottle'-loc
	xat-ta	'wind'-loc

/L/ surfaces as [l], [n], [t], or [d], as in the plural marker /-LAr/.

(6)	*ača-lar*	'father'-PL
	xam-nar	'shaman'-PL
	šil-dar	'bottle'-PL
	xat-tar	'wind'-PL

The surface forms of archiphonemes /N/, /D/ and /L/ show the same phonological conditioning.

(7)	/N/		[d]	/	[+lateral] __ [5]
	/D/	→	[t]	/	[-voice] __
	/L/		[n]	/	[+nasal] __

Elsewhere (after all other voiced segments) these archiphonemes surface in their unmarked form; [n], [d] or [l], respectively.

Full assimilation of /L/ applies only within words. When /L/ is in the onset of a simple clitic it only undergoes partial alternation.

(8)	*ača=la*	'father'-EQU
	xem=ne	'river'-EQU
	xaan=na	'khan'-EQU
	šil=le (šil=de)*	'bottle'-EQU
	xat=la (xat=ta)*	'wind'-EQU

1.2.1 Consonant Allophones

Several consonant sounds exhibit positionally determined variability in Tuvan:

	surface alternation	front vocalic context		back vocalic context	
(9)	k ~ q	[k] *inek*	'cow'	[q] *ayaq*	'bowl'
	g ~ ɣ	[g] *igil*	'fiddle'	[ɣ] *deɣ*	'as if'
		ögden	'yurt'-ADL	*daɣ*	'mountain'
		algɨš	'shamanic chant'		

Labio-dental approximant [ʋ] is an independent phoneme: *aʋay* 'mother', vs. *ačay* 'father' but [ʋ] is also an allophone of [p] in intervocalic position: *xep* 'clothing' *xeʋi* 'clothing'-

[5] The dissimilatory rule that changes /l/ to [d] after /l/ is attested widely across Turkic. Tuvan, like many Turkic languages, bans geminate [ll] within words. [ll] occurs only across a word/enclitic boundary. Cf. *kol=la* 'main'=emph, but *xol-da* 'hand'-loc.

3. Surface [ǰ] is a variant of [č] when it follows a sonorant: *aŋǰɨ* ‘hunter’ *emǰi* ‘doctor’ *qonǰuɣ* ‘very’.

1.2.2 Plosives

Bilabial stops [p] [b] contrast only in word-initial position. The contrast for some speakers is one of voicing, [p] / [b], [t] / [d]. For most speakers, however, the contrast relies on the presence or absence of aspiration, which in turn determines voice onset time. We thus assume aspiration to be the basis of the phonological contrast for these segments at least for some speakers. [6] The following are minimal pairs:

	aspirated / voiceless		unaspirated / weakly voiced	
(10)	*p*^{h}*ar*	‘tiger’	*par*	‘is/are’
	p^{h}*aj*	‘a clay pigeon for shooting’	*paj*	‘rich’

Alveolar show a similar contrast in word initial position:

	aspirated / unvoiced		unaspirated / weakly voiced	
(11)	*t*^{h}*us*	‘separate’ (adj.)	*tus*	‘salt’
	t^{h}*araa*	‘bread’	*taraa*	‘next’

1.2.5 Consonant Clusters

No onset clusters exist in native Tuvan words. The only coda clusters are [rt] and [yt]. The following heterosyllabic clusters are attested in the native lexicon (we exclude clusters found only in Russian borrowings but include some clusters--notably *rb* and *lb*--that were probably borrowed with Mongolian words). In the examples below, we classify possible clusters according to the first segment.

[6] The transcription used in the examples of Tuvan words throughout the remainder of this chapter and dissertation reflects the orthography of Tyvan: Unaspirated stops [p] and [t] are symbolized by ‘b’ and ‘d’ and aspirated stops [t^{h}] and [p^{h}] are symbolized by ‘t’ and ‘p’.

(12) Native consonant clusters

	m	p	b	t	d	s	š	č ǰ	z	ž	y	n	r	l	ŋ	k	g
m			+		+			+	+			+					+
p		+		+				+						+		+	
b	+														+		
t		+		+		+								+		+	
d																	
s		+		+				+								+	
š		+		+				+								+	
č																	
z														+			+
y	+			+	+						+	+		+			+
n	+				+			+	+			+			+		+
r			+	+	+					+		+		+			+
l			+		+			+	+					+			+
ŋ	+				+			+	+			+					+
k		+		+			+	+						+			
g			+		+					+				+			

(13) Plosive: *pp, pt, pč, pl, pk (~ q)*
aptaraa, 'wooden chest', *xepče* 'clothing'-ADL , *xeple* 'clothing'-EMPH, *xepke* 'clothing'-DAT, *pelekke* 'gift'-DAT

(14) Plosive: *tp, tt, tč, tl, ts, tk (~ q)*
udutpa 'sleep'-CAUS-NEG, *attar* 'name'-PL, *otčugaš* 'fire'-DIM, *at=la* 'name'=EMPH, *atqa* 'name'-DAT

(15) Velar (~uvular): *kp, kt, kš, kč, kl (~ q)*
idiktig 'boot'-ADJ, *oq=pile* 'bullet'-INSTR, *soqtaar* 'fact'-P/F , *akša* 'money', *aq=če* 'white'-ALL, *aq=la*'white'-EMPH

(16) Velar: *gb, gd, gž, gl*
ugbay 'elder sister', *ögden* 'yurt'-ALL, *ögže* 'yurt'-ADL, *ögler* 'yurt'-PL

(17) Nasal: *mb, md, mn, mz, mǰ, mg*
am=be 'now'-QUES, *amdan* 'taste', *xamnaar* 'shamanize'-P/F , *emzig* 'medicine'-EQU, *amgɨ* 'present', *emǰi* 'doctor', *čemnen* 'eat'-IMPER

(18) Nasal: *nd, nm, nn, nz, nǰ, ng*
mende 'I'-LOC, *onmašqa* 'othewise', *san-na* 'number'-EQUA, *onza* 'special', *qunǰuu* 'husband's mother', *xanga* 'blood'-DAT

(19) Velar Nasal: *ŋd, ŋm, ŋn, ŋz, ŋǰ, ng*
aŋdarɨq 'turned inside out', *aŋmeŋ* 'animals (collective), *aŋnaar* 'hunt'-FUT, *aŋzaq* 'hunting'-ADJ, *aŋǰɨ* 'hunter', *aŋga* 'animal'-DAT

(20) Sonorant: *rb, rt, rd, rn, rž, rl, rg*
borbaq 'round', *ertine* 'treasure', *erden* 'male'-ADL, *arn-ɨm* 'face'-1SG, *aržaan* 'spring', *aldarlɨɣ* 'famous', *arga* 'means'

(21) Lateral: *lb, ld, ll, lz, lǰ, lg*
alban 'definitely', *aldar* 'fame', *kol=la* 'main'-EMPH, *bolzumza* 'be'-FUT-1SG-CON, *aal=ǰe* 'yurt'-ABL, *aalga* 'yurt'-DAT

(22) Alveolar Fricative: *sp, st, sč, sk ~ sq*
aas=pile, 'mouth'-INSTR, *aastan,* 'mouth'-ADL, *aas=če,* 'mouth'-ABL, *öskeer,* 'to stray'-FUT

(23) Palatal Fricative: *šp, št, šč, šk ~ šq*
bɨštaq 'cheese', *čestey* 'elder sister's husband', *čaraš-paj* 'beautiful'-DIM, *ešter* 'friend'-PL, *ɨšqaš* 'as if', *ašqa* 'money'

(24) Glide: *yt, yd, ym, yn, yl, yg*
aydan 'moon'-ADL, *aytɨrɨɣ* 'question', *moynum* 'neck'-1SG, *oynaar* 'play'-P/F , *xoytpaq* 'fermented milk', *ayga* 'moon'-DAT

In borrowings, onset clusters that are not tolerated are typically broken up by vowel epenthesis. Coda clusters in borrowings are typically reduced by consonant elision: *rasis* 'racist' *spesalis* 'specialist' (R. *spetsyalist*).

1.3.2 Metathesis

Sequences of labial and velar (or uvular) stops in Russian loanwords frequently undergo metathesis. When a sequence [pk] appear in loanwords, it is often reanalyzed as underlyingly /kp/. However, some bilingual speakers fluctuate between /pk/ and /kp/, using both forms:

(25) *pro**qp**a (~ pro**pq**a)* 'cork' (from Russian *propka*)
*trya**qp**a (~ trya**pq**a)* 'rag' (from Russian *tryapka*)

Metathesis of /pk/ sequences is also observed in native words. Although labial / velar sequences never occur word-internally, they frequently result from the juxtaposition of converb plus auxiliary forms:

(26) *bižip kaan men*
'read'-cv aux 1
'I already wrote'

These [pk] sequences are invariably metathesized to [kp] in rapid speech. For some younger speakers in certain dialect regions they are rendered [kp] even in slow speech. In metathesis, [p] shifts from coda to onset position, where it may become (partially) voiced. Examples in (27) column one represent careful speech, while those in column two show the same utterances with [pk] metathesizing to [kp] in fluent, rapid speech.

		Slow or careful speech	Rapid or fluent speech
(27)	a.	*bo nom körü***p q***aan men* 'that' 'book' 'see'-CV AUX-PAST.I 'I' 'I have already seen that book'	*bo nom körü***<u>k p</u>***aan men*
	b.	*xoy ölü***p q***aan* 'sheep' 'die'-cv AUX-PAST.I 'The sheep has (already) died'	*xoy ölü***<u>k b</u>***aan*
	c.	*kaynaar čoru***p q***aan* 'where.to' 'go'-cv AUX-PAST.I 'Where has s/he gone?'	*kaynaar čoru***<u>q b</u>***aan*
	g.	*kažan olar čedi***p k***eer* 'they' 'reach'-CV CLOC-P/F -DAT 'When they came'	*kažan olar čedi***<u>k p</u>***eerge*

Stop fricative sequences metathezise in fluent speech. Like [pk] metathesis, this is both a historical process that has already been completed in many borrowings, and an active process that shows variation between slow vs. rapid speech. It does not occur across word boundaries in the way that p/k metathesis does (28c). It may affect any new loanwords coming into the language.

(28) a. *ša***spË***al* (< *ša***psË***l*) — name of mountain
b. *a***sË***a* (orthographic *aqša*). — 'money'
c. *aq šil* *aš kil* — 'white bottle'

The desiderative affix /-ks-/ exemplifies the fluid state of underlying stop fricative sequences. Speakers of several regional dialects have fully reanalyzed this sequence as /-sx-/, and always render it as [sk] ~ [sx] at the surface. A few dialects and idiolects seem to have both surface forms in more or less free variation. Other dialects (e.g. Süt-xöl) have underlying /-ks-/, rendered as [ks] in careful speech but typically metathesized to [sx] or [sk] in rapid speech: *maŋna***-<u>sx</u>***-ap tur men* 'I want to run' *biʒi***-<u>sk</u>***-ep tur men* 'I want to write'. To complicate matters further, most speakers have enough cross-dialect exposure so that they recognize both [ks] and [sk] as common renderings. When speaking with people from another region, speakers may be observed to over- or under-apply metathesis to accommodate their speech style to that of their interlocutor.

1.4 Syllable Structure

The following syllable types are attested in native Tuvan words:

(29)	V	*a*	'and'	*a.zɨ*	'if'
	VV	*öö*	'yurt'-3	*oo.žum*	'slow'
	VC	*am*	'now'	*ot*	'fire'
	VVC	*aar*	'heavy'	*eet*	'delta'
	VCC	*àrt*	'mountain pass'	*ìrt*	'ram'
	CV	*te*	'male wild goat'	*ča*	'bow'
	CVV	*xuu*	'private'	*čaa*	'new'
	CVC	*xap*	'sack'	*xar*	'snow
	CVVC	*šuut*	'completely'	*suur*	'village'
	CVCC	*dört* [7]	'four'	*bö̀rt*	'hat'

2.0 Morphology

Tuvan, like all Turkic languages, has a primarily agglutinative morphology. The language is almost exclusively suffixing, with each suffix bearing only one piece of grammatical information. Tuvan also has a few morphological processes that are non-agglutinating; these include elision, vowel lengthening, and (partial or full) reduplication. These will be discussed briefly here and in greated detail in later chapters. In this section, we provide an overview of all salient features of Tuvan nominal and verbal morphology, both inflectional and derivational.

---

[7] [rt] and [yt] are the only tautosyllabic consonant clusters in native Tuvan words: *d̥òrt* 'straight', *bö̀rt* 'hat', *xoytpaq* 'fermented milk'.

2.1 Nominal morphology

In Tuvan, number, case and possession are realized morphologically in the noun. We begin with an overview of the productive derivational morphology of nouns, then move on to inflectional morphology.

Tuvan utilizes a wide array of deverbal and denominal noun-formants. Some commonly used suffixes are /-Ig/, which forms deverbal nouns; /-čI/ and /-kčI/ the agentive noun formants; and /-IIškIn/ which forms abstract nouns of action.

(30)

Basic noun or verb		Derived noun	
aytɨr-	'to ask'	*aytɨrɨg*	'question'
balɨk	'fish'	*balɨkčɨ*	'fisherman'
kör-	'to see'	*körüüškün*	'view(ing)'

Tuvan has also borrowed deverbal noun formants from Mongolian, for example: /-čIn/, /-lgA/, /-mAl/, /-ldA/ and /-l/. These appear with both Mongolian and Turkic roots.

(31)

Basic noun or verb		Derived noun with Mongolian suffix	
mal	'livestock'	*malčɨn*	'herder'
boda-	'to think'	*bodal*	'thought'
čemnen-	'to eat'	*čemnenilge*	'nourishment'
kör-	'to see'	*körülde*	'observation'

Possessed forms of modifiers also frequently function as abstract nouns in Tuvan:

(32)

modifier		possessed form	
dürgen	'fast'	*dürgeni*	'speed'
čaraš	'beautiful'	*čaražɨ*	'beauty'

Tuvan also makes use of several derivational diminutive suffixes:

(33)

/-čIɣAš/	*at-čɨɣaš*	'little horse'
	xol-čuɣeš	'small hand'
/čIk/	*xem-čik*	'little river'
/-BAy/	*dürgen-mey*	'quick(ly)'-DIM
	čaraš-pay	'beautiful'-DIM
	diispey	'little cat'

2.1.1 Number

The plural suffix in Tuvan is /-LAr/, which has eight phonologically conditioned variants.

(34)	*bürüler*	'leaves'	*xadɨlar*	'pines'
	emner	'medicines'	*nomnar*	'books'
	xölder	'lakes'	*xoldar*	'hands'
	inekter	'cows'	*àttar*	'horses'

The PL marker precedes possessor markers and case suffixes:

(35)	PL < POSS	*nom-nar-ɨm*	'book'-PL-1
		nom-nar-ɨŋ	'book'-PL-2
	PL < Case	*bürü-ler-nɨ*	'leaf'-PL-ACC
		bürü-ler-nɨŋ	'leaf'-PL-GEN
	PL < POSS < Case	*idik-ter-im-ni*	'boot'-PL-1-ACC
		idik-ter-im-niŋ	'boot'-PL-1-GEN

2.1.2 Nominal case

Six cases are marked morphologically: locative (LOC), ablative (ABL), dative (DAT), accusative (ACC), genitive (GEN) and allative (ADL). The basic form of each case suffix can be considered the form that suffixes to vowel-final stems. Case markers are fully integrated phonologically to the host word and subject to vowel harmony. The exception is the allative case, which fails to undergo back harmony. It may be viewed as an enclitic with a fully specified vowel rather than a true suffix with an undpecified vowel. Archiphonemic representations of the Tuvan case suffixes are as follows:

(36)

Case	Archiphonemic	Surface forms			
ACC	/-NI/	*ni*	*-nü*	*-nɨ*	*-nu*
		-di	*-dü*	*-dɨ*	*-du*
		-ti	*-tü*	*-tɨ*	*-tu*
GEN	/-NIŋ/	*niŋ*	*-nüŋ*	*-nɨŋ*	*-nuŋ*
		-diŋ	*-düŋ*	*-dɨŋ*	*-duŋ*
		-tiŋ	*-tüŋ*	*-tɨŋ*	*-tuŋ*
DAT	/-GA/	*-ge*	*-ga*		
		-ɣe	*-ɣa*		
		-ke	*-qa*		
Loc	*/-DA/*	*-de*	*-da*		
		-te	*-ta*		
Abl	*/-DAn/*	*-den*	*-dan*		
		-ten	*-tan*		
Adl	*/-Je/*	*-če*	*-ǰe*	*-že*	

Sample paradigms with the full range of case markings are as follows:

(37)

	'leaf'	*'pine'*	*'book'*	*'lake'*	*'horse'*	*'boot'*
Nom	*bürü*	*xadɨ*	*nom*	*xol*	*àt*	*idik*
Acc	*bürünü*	*xadɨnɨ*	*nomnu*	*xoldü*	*àttɨ*	*idikti*
Gen	*bürünüŋ*	*xadɨnɨŋ*	*nomnuŋ*	*xoldüŋ*	*àttɨŋ*	*idiktiŋ*
Dat	*bürüɣe*	*xadɨɣa*	*nomga*	*xolge*	*àtka*	*idikke*
Loc	*bürüde*	*xadɨda*	*nomda*	*xolde*	*àtta*	*idikte*
Abl	*bürüden*	*xadɨdan*	*nomdan*	*xolden*	*àttan*	*idikten*
Adl	*bürüže*	*xadɨže*	*nomǰe*	*xolǰe*	*àtče*	*idikče*

Cases fulfill a wide range of semantic and syntactic functions in Tuvan. We summarize here the most common functions of each case. For expository purposes in this section only, we have set off the case suffixes with hyphens.

Nominative (NOM) case in Tuvan serves as the unmarked case for the subject of a sentence. Since NOM has no morphological marker, we do not include it in interlinear glosses.

(38)	*men ɨrlaar men*	*Mergen keldi*
	'I' 'sing'-P/F 1	'Mergen' 'come'-PAST.II-3
	'I will sing'	'Mergen came'

The first nominal in two-part *izafet* (possessive) and post-positional constructions is in the nominative case (see 2.1.3).

(39) *inek südü*
'cow' 'milk'-3
'cow milk'

ög ištinde
'yurt' 'inside'-3-LOC
'inside the yurt'

Nominative marks indefinite direct objects of transitive verbs.

(40) *men iyi čagaa aldɨm*
'I' 'two' 'letter' 'take'-PAST.II-1
'I got two letters'

sen inek saan be
'you' 'cow'-NOM 'milk'-PAST.II Q
'Have you milked a cow?'

Nominative also marks motion towards a place or object in non-emphatic constructions (cf. allative case, which expresses emphatic motion towards a place/object).

(41) *ol kɨzɨl baar*
's/he''Kyzyl' 'go'-P/F
's/he will go to Kyzyl'

Ablative (ABL) case marks motion away from a location, either in time or in space.

(42) *olar xoray-dan kelgen*
'they' 'town'-ABL CLOC-PAST
'They've come from town'

bir-ay-dan turar men
'one-month'-ABL AUX-P/F
'I've been here since January'

Ablative may also mark the source, or material of which something is made.

(43) *damdɨ-dan dalay büder*
['drop'-ABL] ['sea'] ['form'-P/F]
'from (rain)drops the sea is formed'

(44) *šay ayak-tan aža berdi*
'tea' 'bowl'-Adl 'spill'-cv inch-Past.ii
'tea began to spill from the bowl'

(45) *idiimni xöm-den kɨlgan*
'boot'-1-ACC 'leather'-ADL 'make'-PAST.I
'(He) made my boots from leather'

Ablative marks the object of comparison in a comparative construction.

(46) *ol kiži men-den uluɣ*
'that' 'person' 'I'-ADL 'big'
'that person is taller than me'

Ablative may express a partitive meaning on a direct object noun.

(47) *šay-dan ižer men* — 'tea'-ADL 'drink'-P/F 1 — 'I'll drink some (of the) tea'

aška-dan berem — 'money'-ADL 'give'-POL — '(please) give me some money'

Ablative is governed by certain postpositions, e.g. *öske* 'besides'

(48) *orus dɨl-dan öske kandɨg dɨldar bilir sen ?*
'Rus.' 'language'-ADL 'other' 'which' 'langauge'-PL 'know'-P/F 'you'
Besides Russian, what languages do you know?

The accusative (ACC) case in Tuvan performs several functions. In direct object position, only definite nouns receive accusative case marking. Indefinite nouns remain in the unmarked (Nominative) case.

(49)

Indefinite object	Definite object
nom ekkel	*nom-nu ekkel*
'book' 'bring'	'book'-ACC 'bring'
'bring **a** book!'	'bring **the** book!'

(50) *daɣ-nɨ tuneldep erter bis*
'mountain'-ACC 'tunnel'-CV 'pass.through'-P/F 'we'
'We will go through the mountain by tunnel'

Pronouns, when they serve as direct objects, are always ACC marked.

(51) *bo ulus men-i bilir*
'this' 'people' 'I'-ACC 'know'-P/F
'these people know me'

The Allative (ALL) case suffix is /+Je/. This case is limited in function: it marks direction (either real or abstract) towards or around a location. It is generally omitted in non-emphatic contexts (52) and reserved for emphatic or assertive contexts (53).

(52) *daarta Turan čoruur sen be*
'tomorrow' 'Turan' 'go'-P/F 'you' Q
'Are you going to Turan tomorrow?'

(53) *čok daarta men Sukpak-če čoruur men*
NEG'tomorrow' 'I' 'Sukpak'-ALLgo'-P/F 1
'No, tomorrow I'm going to Sukpak' (emphatic)

Allative may also signal motion in the proximity of a location (54), or action directed towards an object (55).

(54) *bis süt-xol-če dolganɨp čoraan bis*

'we' 'Milk.lake'-ALL 'go.around'-CV 'go'-PAST.I 'we'
'We went around Milk lake'

(55) *čuruk-če körbüšaan men.*
‘picture’-All ‘look.at’ ‘I’
‘I was still looking at the picture.’

The dative (DAT) case serves primarily to mark a recipient, beneficiary or indirect object.

(56) *(men) kiži-ge nom berdim*
‘I’ ‘person’-dat ‘book’ ‘give’-Past.ii-1
‘I gave a person a book’

men àt-ka sigen berdim
‘I’ ‘horse’-dat ‘hay’’give’-Past.ii-1
‘I gave hay to the horse’

Dative marks the causative of a transitive verb.

(57) *ačam-ga tool čugaaladɨp alɨyn*
‘father’-dat ‘story’ ‘say’-caus-cv aux-f
‘I make my dad tell a story’

Dative also marks an overtly expressed agent in passive formations.

(58) *akɨm-ga etedip aldɨm*
‘elder.brother’-dat ‘beat’-caus-cv sben-Past.i
‘I was beaten by my elder brother’

Dative is required to mark physical location in past and future tenses (59). Present tense location is marked by the locative (60).

(59) ***Past tense, Dative case***
avam bažɨŋ-ga turgan
‘mother’-1 ‘house’-dat aux-Past.i
‘my mother will be at home’

Future tense, Dative case
avam bažɨŋ-ga turar
‘mother’-1 ‘house’-dat aux-P/F
‘my mother was at home’

(60) ***Present tense, Locative case***
avam bažɨŋ-da
‘mother’-1 ‘house’-loc
‘my mother is at home’

The dative case also marks the experiencer subject with certain adjectival predicates.

(61) *bis-ke bildingir eves*
‘we’-DAT ‘known’ ‘not’
‘(it) is unknown to us’

meŋee nom xerek
‘I’-DAT ‘book’ ‘necessary’
‘I need a book’

(62) *iyi-ge iyi deŋ*
‘two’-DAT ‘two’ ‘equal’
‘Two equals two’

iyi-ge üš deŋ eves
‘two’-DAT ‘three’ ‘equal’ ‘not’
‘Three does not equal two’

Certain verbs and modifiers may govern the dative: e.g. ‘to like’, ‘to be able’.

(63) *men kɨš-ka ɨnak eves men*
1 ‘winter’-DAT ‘like’ ‘not’ 1
‘I don’t like winter’

(64) *ol bažɨŋ kɨlɨrɨn-ga salɨmnɨɣ*
‘he’ ‘house’ ‘make’-ACC-DAT ‘ability’-ADJ
‘He knows how to build houses’

The genitive (GEN) case marks the possessor in a possessive construction, followed by a noun in the possessed form (indexing the person/number of the possessor).

(65) *baškɨ-nɨŋ nomu*
‘teacher’-GEN ‘book’-3
‘the teacher’s book’

ačam-nɨŋ boozu
‘father’-1-GEN ‘gun’-3
‘my father’s gun’

(66) *daɣ-nɨŋ bažɨ*
'mountain'-GEN 'head'-3
'the mountain's peak'

The LOC case expresses location in time or space in the present tense only.

(67) *men amda Kɨzɨl-da čurttap tur men*
‘I’ ‘now’ ‘Kyzyl’-LOC ‘live’-CV AUX ‘I’
‘I live in Kyzyl now’

Location in space in the past or future is marked by the dative case (as noted above).

(68) *men Kɨzɨl-ga törtüngen men*
‘I’ ‘K’-DAT ‘be.born.-PAST.I 1
‘I was born in Kyzyl’

men Kɨzɨl-ga čurttaar men
‘I’ ‘Kyzyl’-LOC ‘live’-P/F ‘I’
‘I will live in Kyzyl’

Location in time takes the LOC in all tenses.

(69) *kelir ay-da keer men*
‘come’-P/F ‘month’-LOC come-P/L 1
‘I’ll come next month’

biray-da keldim
‘one’ ‘month’-LOC come’-PAST.II-1
‘I came in January’

The LOC has been extended to denote simple possession, as in Russian:

(70) *men-de akɨlarɨm čok*
‘I’-LOC ‘older.brother’-PL-1 ‘not’
‘I don’t have older brothers’

men-de xöy nom bar
‘I’-LOC ‘many’ ‘book’ ‘there.is’
‘I have many books’

Case marking in conjoined, multi-word, and appositional nouns generally is realized only on the rightmost element.

(71) *akkɨm=bile uɣbam-nɨ kördɨŋ be-*
‘elder.brother’-1=INS ‘elder.sister’-1-ACC ‘see’-PAST.II-2 Q
‘Did you see my older brother and sister?’

(72) *men Moskva baza Kɨzɨl-ga ažɨldadɨm*
‘I’ ‘Moscow’ ‘and’ ‘Kyzyl’-DAT ‘work’-PAST.II.1
‘I worked in Moscow and Kyzyl’

2.1.3 Possession

The person/number of the possessor is morphologically marked on the possessum. The form of the suffix is determined by the phonological shape of the stem.

(73)

	singular	***plural***
first person	/-(I)m/	/-IvIs/
second person	/-(I)ŋ/	/-IŋAr/
third person	/-(z)I/	/-(z)I/

(74)

noun	1.POSS	2.POSS	3.POSS	gloss
xar	*xarɨm*	*xarɨŋ*	*xarɨ*	‘snow’
àt	*àdɨm*	*àdɨŋ*	*àdɨ*	‘horse’
teve	*tevem*	*teveŋ*	*tevezi*	‘camel’
	1PL.POSS	2PL.POSS	3PL.POSS	
	xarɨvɨs	*xarɨŋar*	*xarɨ*	‘snow’
	àdɨvɨs	*àdɨŋar*	*àdɨ*	‘horse’
	tevevis	*teveŋer*	*tevezi*	‘camel’

In possessive constructions both possessor and possessum are overtly marked. The possessor takes the genitive case. It is followed by the possessum in the possessed form, which indicates the person/number of the possessor.

(75) *(meeŋ) tevem*
‘I’-GEN ‘camel’-1
‘my camel’

(bistiŋ) tevevis
‘we’-GEN ‘camel’-1
‘our camel’

Possessive markers precede case markers. Suffixes show phonologically conditioned variation (e.g. ABL *-dan* ~ *-tan*).

(76)

	‘my book’	***‘your book’***	***‘her book’***	***‘our book’***	***‘your book’ (pl.)***
NOM	*nomum*	*nomuŋ*	*nomu*	*nomuvus*	*nomuŋar*
ACC	*nomumnu*	*nomuŋnu*	*nomun*	*nomuvustu*	*nomuŋarnɨ*
GEN	*nomumnuŋ*	*nomuŋnuŋ*	*nomunuŋ*	*nomuvustuŋ*	*nomuŋarnɨŋ*
DAT	*nomumga*	*nomuŋga*	*nomunga*	*nomuvuska*	*nomuŋarga*
LOC	*nomumda*	*nomuŋda*	*nomunda*	*nomuvusta*	*nomuŋarda*
ABL	*nomumdan*	*nomuŋdan*	*nomundan*	*nomuvustan*	*nomuŋardan*
ALL	*nomumǰe*	*nomuŋǰe*	*nomunǰe*	*nomuvusče*	*nomuŋarže*

Some additional constructions can mark possession. The derivational suffix /-LIg/ derives an adjectival modifier form any noun, e.g. *bàš* ‘head’ *bàštɨg* ‘having a head’

(77) *ɨ̀dɨm čok, (men) diistiɣ men*
‘dog’-1 ‘not’ 1 ‘cat’-ADJ 1
‘I don’t have a dog, I have a cat’

(men) beš àttɨg men
‘I’ ‘5’ ‘horse’-ADJ 1
‘I have five horses’

A possessive formation using the locative case is used to mean ‘at X’s home/place’ or ‘possessed by X’.

(78) *mende bir akkɨm bar*
‘I’-LOC ‘one’ ‘older.brother’-1 COP
‘I have one older brother’

Possessive markers also play a role in *izafet* constructions. In the Turkological literature *izafet* refers to a a type of compounding used to indicate possession. Such compounds consist of two nominal stems, the first of which is in the unmarked (nominative) form, and the second in the third singular possessive form.

(79) *čüve adɨ*
'thing’ ‘name’-3
‘a noun’

xam xevi
‘shaman’ ‘clothing’-3
‘shamanic clothing’

These contrast with 2-part noun + noun possessive formations. In these, unlike in *izafet*, the first component is marked by Genitive case and a possesive suffix.

(80) *kižiniŋ adɨ*
‘person’-GEN ‘name’-3
‘a person’s name’

ačamnɨŋ mašɨnazɨ
‘father’-1-GEN ‘car’-3
‘my father’s car’

3.0 Pronouns

Tuvan expresses two numbers and three persons. The singular pronouns show morpho-phonological irregularities in their declension, while the plural pronouns behave as regular nouns. The third singular pronoun *ol* 's/he' is identical to the demonstrative pronoun *ol* ‘that’. Dative singular pronouns have optionally alternating long and short forms.

(81)

	Singular 1	2	3	Plural 1	2	3
NOM	*men*	*sen*	*ol*	*bis(ter)*	*siler*	*olar*
ACC	*meni*	*seni*	*onu*	*bisti*	*silerni*	*olarnɨ*
GEN	*meeŋ*	*seeŋ*	*ooŋ*	*bistiŋ*	*silerniŋ*	*olarnɨŋ*
DAT	*mee* (*meŋee*)	*see* (*seŋee*)	*aa* (*aŋaa*)	*biske*	*silerge*	*olarga*
LOC	*mende*	*sende*	*ɨnda*	*biste*	*silerde*	*olarda*
ABL	*menden*	*senden*	*oon*	*bisten*	*silerden*	*olardan*
ALL	*menje*	*senje*	*olje*	*bisče*	*silerje*	*olarje*

3.1 Demonstratives

The basic set of demonstratives in Tuvan consists of the lexemes *bo* 'this one', *ol* 'that one' (distal), and *döö* (~ *düü* ~ *duu* ~ *doo*) 'that one over there' (more distal). Definiteness may be marked on subjects and indirect objects by using one of these demonstratives. On direct objects, definiteness is expressed by accusative case marking (see 2.1.2).

3.2 Reflexives

Reflexive pronouns are formed from the stem *bot-* 'self', plus a person marker (§8.3). On verbs, reflexivity is expressed morphologically (§8.1)

(82)

	Singular	Plural
1	*bod-um*	*bod-uvus (~ bot-tar-ɨvɨs)*
2	*bod-uŋ*	*bod-uŋar*
3	*bod-u*	*bot-tarɨ*

(83) *boduŋ qadɨɨŋ boduŋ xoluŋda*
'self'-2 'health'-2 'self'-2 'hand'-2-LOC
'your health is in your (own) hands'

(84) *kiži-bürüzü bodun bay dep sanaar*
'everyone' 'self'-3-ACC 'rich' COMP 'count'-P/F
'Everyone considers himself rich'

They may also bear number and case marking.

(85) *bot-tar-ivis-tan*
'self'-PL-POSS-ABL
'from our own'

bod-um-nuŋ xolumda
'self'-1-GEN 'hand'-1-LOC
'in my own hand'

3.3 Interrogative pronouns

The basic interrogative pronouns are as follows.

(86)	*kɨm*	'who'	*čüü*	'what'
	čüge	'why'	*čüden*	'from what'
	čeže	' how many'	*kàš*	'how much/many'
	kandɨg	'how, which'	*kayɨ ~ kay(ɨ)zɨ*	'which'
	kayda	'where'	*kayaa ~ kaynaar*	'to where'
	kažan	'when'	*kayɨɨn ~ kayɨɨrtan*	'from where'

(87) *kayɨɨn kelgen siler*
'whence' 'come'-PAST.I 'you'
'Where have you come from?'

kaynaar baarɨŋ ol
'where.to' 'go'-P/F -2 DISC
'Where are you going?'

(88) *sen kàš xarlɨg sen*
'you' 'how.many' 'year'-ADJ 2
'How old are you?'

bažɨŋ kayɨ-xire uluɣ
'house' 'how.much' 'big'
'How big is the house?'

3.4 Indefinite pronouns

Tuvan possesses several types of indefinite pronominal forms. *One* form employs bir 'one' plus the emphatic enclitic /=lA/, followed by a noun.

(89) *bir-le kiži*
'one'=EMPH person
'someone'

bir-le čüve
'one'=EMPH thing
'something'

An interrogative pronoun followed by enclitic [=bir] 'one' followed by a noun also yields an indefinite meaning.

(90) *kay(ɨ)=bir kiži*
'what.like=one' 'person'
'anyone (at all)'

kay(ɨ)=bir čüve
'what.like=one' 'thing'
'anything (at all)'

An interrogative pronoun or noun plus the emphatic clitic [=daa] (which exhibits negative polarity) can be used in this function as well.

(91)	*kɨm=daa*	'anyone', 'noone'
	kiži=daa	'anyone', 'noone'
	kaynaar=daa	'to anywhere', 'to nowhere'
	čüü=daa	'anything', 'nothing'

3.5 Quantitative pronouns

Common quantitative pronouns include *čamdɨk* ‘some’, *bürü* ‘each, every’ and a universal quantifier *šuptu* ‘all’. These may bear possessive and case marking.

(92) *čamdɨɨvɨs* — ‘some’-1-PL — ‘some of us’

čamdɨɨvɨstan — ‘some’-1-PL-ABL — ‘from some of us’

When the quantified object is overt, it is typically genitive case marked. The quantifier always bears 3-SG possessive marking that indexes the quantified object.

(93) *süttuŋ šuptuzu* — ‘milk’-GEN ‘all’-3 — ‘all of the milk’

ešternɨŋ šuptuzu — 'friend'-GEN 'all'-3 — 'all of the friends'

Alternatively, the quantified object may index a non-overt possessor.

(94) *südüŋ šuptuzu* — ‘milk’-2 ‘all’-3 — ‘all of your milk’

ešterɨŋ šuptuzu — 'friend'-PL-2 'all'-3 — 'all of your friends'

4.0 Adjectives

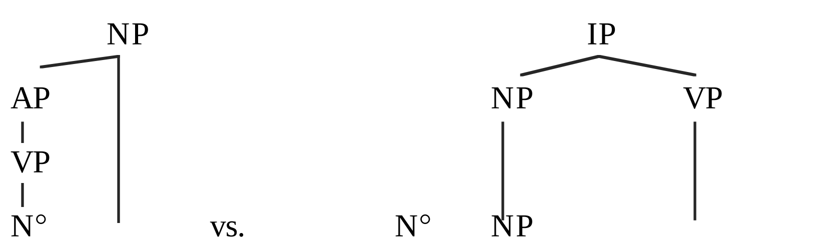
Adjectives precede nouns in Tuvan. Many lexemes double as both nouns and modifiers, e.g. *demir* ‘iron’ (noun or adjective), *daš* ‘stone’, *ɨyaš* ‘wood(en)’, *tɨva* ‘Tuva(n)’.

(95) *tɨva kiži* — ‘Tuvan’ ‘person’ — ‘a Tuvan’

bičii ѐrgi bažɨŋ — ‘small’ ‘old’ house’ — ‘a small, old house’

Modifier phrases take a variety of forms. Verbs or verb phrases in any tense serve as modifiers when they precede a noun, but as verbs when they follow a noun.

(96)



saar kiži — ‘milk’-P/F ‘person’ — ‘a milking person’ *or* ‘a person who milks’

kiži — *saar* — ‘person’ ‘milk’-P/F — ‘a person will milk’

The most common means of deriving adjectives includes the suffix /-LIg/ which forms adjectives denoting possession of the noun or its qualities: *aškalɨg* ‘having money’, *čigirlig* ‘sweet’ *baškɨlɨg* ‘with a teacher’.

(97) *aldar*
‘fame’

aldarlɨg kiži
'fame'-ADJ 'person'
‘famous person’

(98) *aldar čok kiži*
'fame' NEG 'person'
‘an unfamous person’

(99) *sen aškalɨg sen be*
‘you’ ‘money’-ADJ ‘you’ Q
‘Do you have money?’

bis mašɨnalɨg bis
‘we’ ‘car’-ADJ ‘we’
‘We have a car’

Adjectives may be modified in various ways. The superlative is formed by preposing *eŋ* ‘most’ or *eŋ=ne* ‘most’=EMPH before the noun: *šɨɨrak* ‘strong’, *eŋ(=ne) šɨɨrak* ‘strongest’. The suffix /-zImAAr/ denotes an ameliorative: *sarɨɣ* ‘yellow’, *sarɨɣsɨmaar* ‘yellowish’, *türgen* ‘fast’, *türgenzimeer* ‘somwehat fast’. The comparative (CMP) enclitic /-L^2A/ adds the meaning ‘-like’: *xavan* ‘pig’ *xavan=na* ‘pig-like’, *süt* ‘milk’, *süt=le* ‘milk-like’. Finally, Tuvan intensifies some modifiers by (C)Vp- reduplication: *sap-sarɨɣ* ‘very yellow’, *tüp-türgen* ‘very fast’, *kap-kara* ‘completely black’ (see also 8.4.2.5).

Adjectival stems are negated with the particle *eves* 'not'; while nouns that function as modifiers are generally negated with the particle *čok* 'not'.

(100) *bay eves kiži*
‘rich’ NEG 1 ‘person’
‘a not rich person’

aška čok kiži
‘money’-NEG ‘person’
‘a person without money’

5.0 Adverbs

Adverbs in Tuvan are uninflectable words that appear before the adjective or verb they modify: e.g. *ara* ‘not completely’, *aray* ‘barely’, *dɨkka ekki* ‘very well’, *končuɣ* 'rather'. Complex adverbs may be composed of compound nouns or noun+enclitic: *küš=bile* 'with force' *bir=le* 'sometimes' *iyi-čaŋgɨs* 'sometimes' (literally 'two'-'one').

(101) *ol oruk dɨkka ekki bilir*
'he' 'road' 'very' 'well' 'know'
'He knows the road very well'

ol doraan keldi
'he' 'right.away''come'-PAST.II-3
'He came right away'

(102) a. *dürgen maŋnaar kiži*
'fast' 'run'-P/F 'person'
'a fast running person'

b. *dürgen eves maŋnaar kiži*
'fast' NEG 'run'-P/F 'person'
'a not fast running person'

(103) *mašɨnavɨs bir-le ažɨldap tur, bir-le ažɨldavas.*
'car'-1PL 'one'-EMPH 'work'-CV 'one'-EMPH 'work'-NEG-P/F
'Our car sometimes works, sometimes doesn't'

Examples (100, 102b., and 103) all contain negation. The full range of negation types are discussed in 8.4.

6.0 Auxiliary nouns

Auxiliary nouns—a characteristic feature of Tuvan morpho-syntax—perform directional, causal, and other functions filled by prepositions and postpositions in other languages. Common auxiliary nouns include:

(104)	*aldɨ*	'underside'	*išti*	'inside'
	murnu	'front'	*üstü*	'top'

Auxiliary nouns may bear case marking. They appear following a noun which is in either NOM (78) or GEN (79) case. The combination of case markings allowing for a wide range of complex locational and directional relations to be expressed.

(105)	*ot ištinde*	'inside (of) the fire'
	ot ištinǰe	'to the inside (of) the fire'
	ot üstünde	'on top (of) the fire'
	ot üstünden	'from on top (of) the fire'
(106)	*ottuŋ üstünde*	'in the fire's top'
	ottuŋ üstünǰe	'to(ward) the fire's top'

When the auxiliary noun is used without a preceding lexical noun, the resulting form functions as an adverbial:

(107) *ištinden köör*
'inside'-ABL 'look'-P/F
'(to) look from inside'

7.0 Postpositions

Many postpositions are verbal in origin in Tuvan, representing lexicalized uses of participial forms, converb forms, or causative forms of verbs. Postpositions differ from auxiliary nouns in two ways. First, they govern one of several different cases. Second, they do not bear case marking. Common postpositions include *čedir* '(up) to', *deeš* 'because of', *keži̇ldir* 'through, across', *dolgandɨr* 'around' *öske*, 'besides', *ažɨr* 'across'.

(108) *dag ažɨr*
'mountain' 'across'
'across the mountain'

sen deeš keldim
'you' POST 'come'-PAST.II-1
'I came because of you'

(109) *orus dɨldan öske kandɨɣ dɨldar bilir sen*
'Russian' 'language'-ALL 'besides 'which' 'language'-PL 'know' 'you'
'Besides Russian, what languages do you know?'

8.0 Verbal morphology

Tuvan verbs have a rich morphology that reflects the person and number of the subject, as well as tense, mood, aspect, and negation. Each verb also has several distinct converb and participial forms. The auxiliary verb system is discussed in (8.7).

8.1 Verbal categorizers

The most common and productive verbal stem formant (VF) in Tuvan is /-LA/. It attaches to almost any noun stem to create a verb expressing some action associated with the noun. Some forms with /-LA/ are fully lexicalized and have non-compositonal semantics (e.g. *aŋna-* 'hunt'). However, speakers readily generate new verb forms with /-LA/.

(110)

noun		derived verb	
aŋ	'animal'	*aŋna-*	'to hunt'
baškɨ	'teacher'	*baškɨla-*	'to teach'
moskva	'Moscow'	*moskvala-*	'to go via Moscow'
örge	'squirrel'	*örgele-*	'to hunt squirrels'
šay-	'tea'	*šayla-*	'to drink tea'
xap-	'sack'	*xapta-*	'to put in a sack'

Other, less productive, verb-forming suffixes include /-A/, /-IrgA/, /-ZIrA/, /-Ar/, /-rA/, /-Gar/, /-sI/, /šI/, /čI/, and /-ksA/:

(111)	noun / adjective		derived verb	
	at	'name'	*ad-a-*	'to name, call'
	dus	'salt'	*duz-a-*	'to salt'
	ba̖š	'head'	*ba̖š-kar-*	'to lead'
	anɨyak	'young'	*anɨyak-sɨ-*	'to become young'

8.2 Verbal morphology

Verb stems may be morphologically marked for the following categories: causative, passive, reciprocal and reflexive. The most productive causative suffix is /-D(Ir)/ which has an allomorph [-t]. The full form, /-DIr/, appears after C-final stems, while [-t] is used following vowel-final stems or stems already marked with the full form of /-DIr/. Multiple instances of causativity may be marked on a single verb, with each contributing its own semantic content and morphological material.

(112)	*kɨl-*	'do'
	kɨl-dɨr-	'make someone do something'
	kɨl-dɨr-t-	'make someone make someone do something'
	kɨl-dɨr-t-tɨr-	'make someone make someone make someone do something'

(113)	*nomču-*	'read'
	nomču-t-	'make someone read'
	nomču-t-tur-	'make someone make someone read'

The reciprocal suffix /-(I)š/ marks reciprocal, mutual or accompanied action.

(114)	*duzala-*	'help'	*duzala-ž-*	'help each other'
	biži-	'write'	*biži-ž-*	'write (to) each other'
	kut-	'pour'	*kut-š-*	'pour for each other'

The reciprocal suffix may combine with the causative.

(115) *biži-š-tir-* 'make people to write (to) each other'

The reflexive is marked by /-(I)n/, which surfaces with the vowel after C-final stems.

(116)	*kut-*	'pour'	*kutt-un-*	'be poured'
	beletke-	'prepare'	*beletk-en-*	'prepare self'
	öökte-	'button'	*öökte-n-*	'be buttoned'

The passive is marked by the suffix /-(I)l/.

(117)	*kɨzɨr-*	'abbreviate'	*kɨzɨr-l-*	'be abbreviated'
	tip-	'find'	*tɨv-ɨl-*	'be found'

8.3 Person suffixes

There are two sets of verbal inflectional markers, one suffixal (Class-I) and the other enclitic (Class-II). We characterize Class-II markers as enclitics for two reasons: (i) they do not stand alone, and (ii) their initial consonants are do not alternate as do those of suffixes.

(118)

Class-I	Singular	Plural
1	*-(I)m*	*-(I)vIs*
2	*-(I)ŋ*	*-(I)ŋAr*
3	---	*(-LAr)*

Class-II	Singular	Plural
1	*men*	*bis*
2	*sen*	*siler*
3	—	*(-LAr)*

All main clauses (except those containing verbs in the PAST.II) use Class-II markers. Subordinate clauses take Class-I markers.

Both Class-I and Class-II markers are always obligatory, except for the third plural (-LAr), which is optional and marks a stylistic register. Personal pronouns, by contrast, are subject to pro-drop and are frequently omitted. For a full pronominal paradigm see (§3.0).

Class I marked verbs

(119) *(men) čagaa aldɨ-m*
'I' 'letter' 'take'-PAST.II-1
'I got a letter'

(bister) amda kelbeedi-vis
'we'-PL 'yet' 'come'-NEG-PAST.II-1PL
'We haven't come yet'

(120) *(sen) süt išti-ŋ be?*
'milk' 'drink'- PAST.II-2 Q
'Did you drink milk?'

(siler) uštu-ŋar
'you'-PL 'fly'- PAST.II-2.PL
'You flew'

Class II marked verbs

(121) *(men) baškɨlap tur men*
'I' 'teacher'-VSF-CV AUX 1
'I teach'

(bis) čanaar bis
'we' 'go.home'-P/F 1PL
'We're going home'

(122) *(sen) kažan törtüngen sen*
'you' 'when' 'be.born'-PAST.I 1
'When were you born?'

olar oray kelgen(ner)
'they' 'late' 'come'-PAST.I(-PL)
'They came late'

8.4 Tense, aspect, mood

Tuvan expresses various temporal, aspectual, and modal nuances through the use of suffixes and auxiliary verbs. In many instances, a given mood or aspect may expressed by both means. Approximately two dozen common auxiliary verbs are used to index a wide range of temporal, aspectual, and modal distinctions. As in all Turkic languages, a small set of verbs (namely, 'stand' 'go' 'sit' and 'lie') has been grammaticalized for the most common auxiliary uses. These are discussed further in section 8.7.

8.4.1 Tense

8.4.1.1 Past tense

The tense system opposes a non-assertive (or 'indefinite') past and an assertive (or 'definite') past. We refer to these respectively as PAST.I and PAST.II.

The non-assertive or indefinite PAST.I is marked by the suffix /-Gan/, followed by a Class-II person marker. Generally, the PAST.I serves as the unmarked past tense, referring to a point in the past that is indefinite, unspecified, or more distant from the spech event. It is somewhat closer in meaning to the English present perfect, e.g. 'I have seen' than it is to the simple past tense 'I saw'.

The assertive or definite PAST.II is marked by the suffix /-DI/, followed by a Class-I person marker. PAST.II carries a more assertive connotation and may also refer to a (comparatively more) definite time in the past. PAST.II also carries a deictic function in that it typically refers to events already introduced in the discourse or known to the interlocutor.

The table shows a full paradigm of the two past tenses with their respective person markers.

(123)

	PAST.I + Class-II		PAST.II + Class-I	
(men)	*kelgen men*	'I have come'	*keldim*	'I came'
(sen)	*kelgen sen*	'you have come'	*keldiŋ*	'you came'
(ol)	*kelgen*	's/he has come'	*keldi*	'he came'
(bis)	*kelgen bis*	'we have come'	*keldivis*	'we came'
(siler)	*kelgen siler*	'you have come'	*keldiŋer*	'you came' (PL)
(olar)	*kelgen(ner)*	'they have come'	*keldi(ler)*	'they came'

8.4.1.2 Present/Future tense

A non-past action is marked by the present/future (P/F) suffix /-Ir/. With stems ending in a low vowel, the suffix vowel also becomes low (98). This follows the general pattern of hiatus resolution (cf. chapter 3).

(124)	*nomču-*	'read'	*nomčuur*	'reads', 'will read'
	čɨt-	'lie down'	*čɨdɨr*	'lies down', 'will lie down'
	kel-	'comes'	*kelir*	'comes', 'will come'
(125)	*čuɣaala-*	'talk'	*čuɣaalaar*	'talks' , 'will talk'

The present/future form denotes either present or future. The interpretation depends on the semantics of the verb and on the speech context. Some examples are ambiguous.

(126) *bo ulus meni bilir*
'this' 'people' 'I'-ACC 'know'-P/F
'these people know me'

(127) *ažɨldaar men*
'work'-VSF-P/F 1
'I work' *or* 'I will work'

Present tense is more often expressed with a verb stem in the /-(I)p/ converb form (see 8.7.1) followed by a P/F auxiliary verb (see also 8.7). Auxiliaries include *tur*, *turar, olur*, or *čor* and mark a non-past action in Tuvan. These may have either a progressive or non-progressive meaning. By far the most common auxiliary in this function is *tur* 'stand'.

(128) *süt ižip tur men*
'milk' 'drink'-CV AUX 1
'I am drinking milk'

süt ižip turar men
'milk' 'drink'-CV AUX 1
'I (regularly) drink milk'

(129) *am texnikumda baškɨlap tur men*
'now' 'technicum'-LOC 'teacher'-VF-CV AUX 1
'now I work as a teacher at the Technical Institute'

8.4.2 Aspectual morphology

There are at least four aspectual constructions realized affixally in contemporary Tuvan: iterative, perfective, resultative (or assertive) and cessative. The first two of these (iterative/partitive and perfective) are basically derivational in that they occupy a position between the lexical stem and the tense-markers. The latter two (resultative and cessative), on the other hand, occupy the same position in the verb as the tense markers and are thus mutually exclusive with them.

8.4.2.1 Iterative / Partitive

The iterative (ITER) aspect in Tuvan is marked by the suffix /-GIlA/. In current colloquial usage this suffix also marks partitive (PART) or rapid actions, e.g. 'to X a little bit', 'to do X quickly'.

(130) *kɨldɨr-gɨla*
'do'-PART
'(You) work (a little)!'

(men) šaydan kut-kula-an
'I' 'tea'ABL 'make'-ITER-PAST.I
'(I) (quickly) made some tea'

8.4.2.2 Perfective aspect

The perfective (PERF) suffix is /-IvIt/, but is often reduced to /-Ipt/, or /-Ip/ alone. The suffix /-IvIt/ is used most often after a bare (imperative) stem. This combination of PERF and IMP contrasts with the plain imperative in signaling limited duration, rapidity or perfectivity of the action.

(131) *meni mana* vs. *meni manavɨt*
'I'-ACC 'wait'-IMP
'wait for me' (unspecified time)

'I'-ACC 'wait-IMP-PERF
'wait for me' (for a limited time)

(132) *yablaqtɨ či* vs. *yablaqtɨ čivit*
'apple'-ACC 'eat'-IMP
'(You) eat the apple!'

'apple'-ACC 'eat'- IMP -PERF
'(You) eat the apple right now!'
or 'Eat the apple up!'

8.4.2.3 Perfective -p plus -t-

An action that is done done quickly, unexpectedly or for a very short time can be expressed with the PERF.II. It is formed from the /-(I)p/ converb (see 8.7.1) plus the formant [-t-], followed by any tense suffix.

(133) *men ɨrla-p-t-ar-ɨm*
1 'sing'-PERF.II-P/F-1
'I'm gonna sing for a bit'

doozu-p-t-uŋ be
'finish'-PERF.II-PAST.II Q
'Did you finish (it)?'\'

(134) *xar bolza, doraan eskeri-p-t-er men*
'snow' 'be-COND 'right.away' 'notice'-PERF.II-P/F 1
'If there's snow, I'll notice (it) right away'

Compare this to the simple present/future (106):

(135) *xar bolza, eskerer men*
'snow' 'be-COND 'notice'-P/F 1
'If there's snow, I will notice (it)' *or* 'If it snows, I notice'

8.4.2.4 Resultative aspect

The resultative (RSLTV) suffic /-ČIk/ has been described (I&P 1961) as marking actions completed in the past, with emphasis on the result. In contemporary spoken Tuvan, the resultative is used in a limited fashion to pose a rhetorical question (136), or to assert the truth of a statement which the hearer doubts (137).

(136) *sen meni düün köržük sen be*
'you' 'me'-ACC 'yesterday' 'see'-RSLTV 'you' Q
'Didn't you see me yesterday?' *or* 'Don't you realize you saw me yesterday?'

(137) *sen mančɨŋnɨ čop čiveen sen* *čižik men*
'you' 'dumplings'-ACC DISC 'eat'-NEG.PAST.I 2 'eat'-RSLTV 1
'you didn't eat your dumplings, did you?' 'I *did* eat (them)!'

8.4.2.5 Cessative

The cessative (CES), which conveys the meaning 'to stop doing X', is morphologically marked with the negative suffix /-BA/, followed by the stem-forming aspectual suffix /-stA-/.

(141) *škola soonda orus nomnar nomčuvastap kalgan men.*
'school' 'after' 'Rus.' 'book'-PL 'read'-NEG-CES-CV AUX-PAST.I 1
'After (finishing) school, I stopped reading Russian books'

(142) *nomčuvastaanɨm užurundan šilgeldeni baɣay dužadɨm.*
'read'-NEG-CES-PAST.I-1 'reason'-ABL 'exam'-NOM 'badly' 'meet'-PAST.II-1
'Due to my stopping reading, I did badly on the exam'

8.4.3 Mood

A number of suffixes mark different modal categories. These include the following: conditional, concessive, evidential past and present, imperative, optative, desiderative, possibilitive, and subjunctive.

8.4.3.1 Conditional mood

The conditional (COND) in Tuvan is marked by the suffix /-ZA/. The conditional is double-marked in first and second person forms, with the person/number suffix coming between the two conditional suffixes, the first of which appears as /-ZI-/:

(143)	*(men) kelzimze*	'if I come'	*(bis) kelzivisse*	'if we come'
	(sen) kelziŋze	'if you come'	*(siler) kelziŋerze*	'if you come'
	(ol) kelze	'if s/he comes'	*(olar) kelze(ler)*	'if they come'

The conditional has several functions in Tuvan. First, it marks conditional clauses, both past/irrealis (143) and future (144) ones.

(144) *aškaŋ četse čemden sadɨp al*
'money'-2 'suffice'-COND 'food'-ABL 'buy'-CV SBEN-IMP
'if your money is sufficient, buy some food'

(145) *čaaskan bolzumza*
'alone' 'be'-COND-1
'if I am alone'

The conditional may combine with the auxiliary verb *bol-* 'be'.

(146) *čorbas bolza xoržok*
'go'-NEG-FUT 'be'-COND 'impossible'
'It's impossible not to go'

8.4.3.2 Concessive mood

The concessive (CON) is formed from the conditional plus the emphatic (focus) enclitic *=daa*. The vowel of this enclitic is invariant. The CON corresponds to English clauses introduced by 'even though', 'although'.

(147) *men čaaskan bolzumza=daa*
'I' 'alone' 'be'-CON-1-CON=EMPH
'even though I am alone'

nomčuvaanɨm=daa
'read'-NEG-PAST.I-1=EMPH
'Even though I didn't read'

8.4.3.3 Conciliatory mood

The form known as the 'conciliatory' mood (CNCL) in the Tuvan linguistic literature and the 'optative' in general Turkological studies is marked by the suffix /-GAy/. It frequently carries a meaning of an agreement on the part of the subject to perform an action (148). It most frequently appears in the first person.

(148) *am čigey bis*
'now' 'eat'-CNCL 'we'
'okay, we'll eat now'

vs. *am čiir bis*
'now' 'eat'-P/F 'we'
'We'll eat now'

(149) *kɨm xöy man ǰɨ čiir=il kör-gey-le bis*
'who' 'many' 'dumpling' 'eat'-I/D 'see'-CV-CNCL-EMPH 'we'
'We'll see who can eat the most dumplings!'

8.4.3.4 Desiderative mood

The suffix /-ksA/ is used to express a desiderative (DES) meaning. Note that /ks/ often metathesizes to [sk] or [sx] in rapid speech. In some dialects it has been reanalyzed as underlying [-sk].

(150) *men cagaa bižiksedim*
‘I’ ‘letter’ ‘write’-DES-PAST.II-1
‘I wanted to write a letter’

maŋnasxap tur men
‘run’-DES-CV AUX 1
‘I want to run’

8.4.3.5 Evidential mood

The evidential mood (EVID) is marked by the suffix /-DIr/, which may also attach to nouns and modifiers, and has a wide range of uses as an evidential marker.

(151) *men dir men*
‘I’ EVID 1
‘It’s me’

čoktur
NEG-EVID
‘There isn’t any’

bardɨr
AUX-EVID
‘There is/are’

It can also signal reported speech or the inadvertant, involuntary or unexpected nature of an action.

(152) *men àrtar men*
1’stay’-P/F 1
‘I’m staying’

vs.

men àrtardɨr men
1’stay’-P/F-EVID 1
‘I’m supposed to stay’

(153) *möngün taygaga sarlɨk körgendir men*
‘M.T.’-DAT’yak’’see’-PAST.I-EVID 1
‘In Mongun-Taiga I saw a yak’ (unexpectedly)

(154) *ayalganɨ dɨŋnaydɨr men*
‘music’-ACC ‘hear’-EVID.P/F 1
‘suddenly, (it seems) I hear music’

8.4.3.6 Imperative mood

The singular form of the second person imperative (IMP) mood is the simple bare stem, lacking any other tense-aspect-mood, person/number marker (155). In the plural the second person adds the suffix /-IŋAr/.

(155) *čemnen*
‘eat’
‘eat!’

čemneniŋer
‘eat’-PL
‘eat!’ (plural)

In the first person singular the form is /-Ayn/; this form has both a hortatory function (156a) as well as an intentional future meaning (156b).

(156) a. *arbustu čipteyn*
'watermelon' 'eat'-PERF-IMP
'I'm gonna eat the watermelon!'

b. *am čedeveyn*
'now' 'come'-IMP
'I'll come now'

Third person forms are in /-ZIn/ and /-ZIn(nAr)/.

(157) *bodanzɨnnar*
'ponder'-3.IMP-PL
'may they think it over'

bilzin
'know'-3.IMP
'may he know'

In addition to the forms above, Tuvan speakers also may use a first dual (1DL) (inclusive) imperative forms in /-(A)Al(I)/ The dual contrasts with the plural imperative in /-(A)AlIŋAr/.

(158) Dual vs. Plural

(bis) beel(i)
(1PL) 'give'-1DL-IMP.II
Let's (we two) give.

(bis) beeliŋer
(1PL) 'give'-1PL-IMP.II
Let's give.

8.5 Negation

Verbal negation is expressed with the suffix /-BA/. This appears in verbs in various tenses, as well as in the unaccomplished aspect and imperative mood.

(159) *nomču-va-dɨm*
'read'-NEG-PAST.I.1
'I didn't read'

nomču-va-an men
'read'-NEG-PAST.I 1
'I didn't read' *or* I haven't read'

(160) *ɨrla-va-ŋar*
'sing'-NEG-PL
'don't sing!' (plural)

düne kɨlašta-va
'at.night' 'walk'-NEG
'don't walk at night!'

In the present/future, the negative suffix takes the form /-BAs/.

(161) *čemnen-mes sen be*
'eat'-NEG-P/F 'you' Q
'Won't you eat?'

seni ut-pas men
'you'-ACC 'forget'-NEG-P/F 1
'I won't forget you'

The negative marker *čok* '(is) not', also functions as a kind of copula. It may stand alone as the answer to a question: *čok* 'No.' It may also be used in existential and possessive functions corresponding to positive *bar* 'is', 'are'.

(162) *nom bar be*
'book' COP Q
'Is there a book?'

nom čok *[nomǰok]*
'book' NEG-COP
'There's no book'

The two negation markers thus differ in their ability to serve as a copula.

(163)	*Igor čok* 'Igor' NEG-COP 'Igor is not here'	*Igor eves* 'Igor' NEG 'not Igor'

8.6 Verbal modifiers

8.6.1 Converbs

Converbs constitute a crucial element of the Tuvan verb system. Several different converb types are used, some with an extensive range of functions and others quite restricted. In the Turkological literature converbs are traditionally distinguished from participles by the lack of ability on the part of converbs to be inflectable elements. In Tuvan, however, the distinction is not so clear. Some forms that have traditionally been considered converbs in Tuvan studies (I & P 1961, Menges 1956, 1959, Katanov 1903), may, in fact, bear inflection (e.g. /-GAš/; see below).

The most common converb is /-(I)p/. Lexical verb stems appear in this form in almost every auxiliary verb construction, including the present tense formation and the evidential past .

(164)	*ɨrlap kaan men* 'sing'-CV AUX-PAST.I 1 'I already sang'	*inek saap tur men* 'cow' 'milk'-CV aux 1 'I'm milking a cow'
(165)	*(men) kurs doozup tur men* 'I' 'course' 'finish'-CV AUX 1 'I am finishing a course'	*suɣlap čor men* 'fetch.water'-CV AUX 1 'I'm fetching water'

A second important converb is the /-A/ converb. The converb surfaces as a glide [-y] after vowel-final stems:

(166)	<u>verb stem</u>	<u>converb</u>	
	iš-	*iž-e*	'drinking'
	olur-	*olur-a*	'sitting'
	či-	*či-y*	'eating'
	dile-	*dile-y*	'asking'

(167)	*diley aldɨm* 'ask.for'-CV SBEN-PAST.II-1 'I asked for'	*čiy berdim* 'eat'-CV INCH-PAST.II-1 'I began to eat'

The following minimal pair shows that the /-(I)p/ converb and /-A/ converb form of the same verb yield different meanings.

(168) <u>/-(I)p/ converb</u>
bižip kaaptɨm
'write'-CV AUX-PAST.II-1
I already wrote (it).

<u>/-A/ converb</u>
bižiy kaaptɨm
'write'-CV AUX-PAST.II-1
I already wrote (it) (quickly).

Converbs may be restricted in how they co-occur with auxiliaries. Here, the /-(I)p/ converb plus the auxiliary *ber-* yields the benefactive meaning, while the /-A/ converb with *ber-* yields the inchoative meaning.

(169) <u>Benefactive</u>
ol oynap berdi
's/he' 'play'-CV BEN-PAST.II-3
'S/he played for (someone)'

<u>Inchoative</u>
ol oynay berdi
's/he' 'play'-CV AUX-PAST.II-3
'S/he began to play'

In the Tuvan grammatical tradition (e.g. I&P 1961), the converb in /-GAš/ is called the 'past converb'. It may express an action that has preceded the action of the main verb (170), or a progressive action simultaneous with the main verb (171).

(170) *xoy(nu) körüp kaaš ap aldɨm*
'sheep'(-ACC) 'see'-CV AUX-CV 'take'-CV SBEN-PAST.II-1
'having seen the sheep, I took it'
Note: [kaaš] is underlying /kag + GAš/. See chapter 3 on [g] deletion.

(171) *kɨzɨl čorup orgaš orukka xoynu kördüm*
'K.' 'go'-CV AUX-CV 'road'-DAT 'sheep'-ACC 'see'-PAST.II-1
'while going to Kyzyl, I saw a sheep in the road'

Suffix /-GAš/ may also suffix to the /-(I)p/ converb to give the meaning of an action that is done quickly or to only a small degree.

(172) *čugle čaŋgɨs yablak čipkeš üngen men*
'only' 'one' 'apple' 'eat'-PERF-CV 'go out'-PAST.I 1
'Having eaten quickly just one apple, I left'

(173) compare: *čem čigeš, üngen men*
'food' 'eat'-CV 'go out'-PAST.I 1
'Having eaten food, I left'

Converbs are described in the Turcological literature as uninflectable elements. In Tuvan, however, the /-GAš/ converb may take Genitive case marking. We also found that the /-

GAš/ converb suffix may follow the /-(I)p/ converb. The distributional facts of /-GAš/ suggests that it doesn't neatly fit into either the converb or participle category.

(174) *kelgeštiŋ, menče dolgavɨt*
'come'-CV-GEN 'me'-ALL 'phone'-IMP-PERF
'when you return, call me'

Simultaneous (175) or durative (176) action is expressed with a third converb type: /-BIšaan/.

(175) *nomčuvušaan sen be*
'read'-DUR 2 Q
'Are you still reading?'

(176) *avam ɨrlavɨšaan ineen saap olur*
'mother'-1 'sing'-DUR 'cow'-3-ACC 'milk'-CV AUX
'My mother milks her cow while singing'

8.6.2 Participles

Participles play a major role in Tuvan morphosyntax. Most types of main clauses, as well as subordinate, complement, and relative clauses, involve participial verb forms. There are several participial forms. The most common of these are the PAST.I in /-GAn/ (177) and the perfective/future in /-Ir/, which has a negative form /-BAs/ (178).

(177) *bir tuɣaar kelgen kiži*
'one' 'time' 'come'-PAST.I 'person'
'a person who came for the first time'

eŋ türgen maŋnaan kiži
'most''fast''run'-PAST.I 'person'
'the person who ran the fastest'

(178) *čorbas mašɨna*
'go'-NEG-P/F 'car'
'a car that doesn't run'

ažɨldaar ulus
'work'-P/F 'people'
'working people'

Some participles take different agreement markers when they appear in main vs. subordinate clauses. Both the PAST.I in /-Gan/ and the P/F in /-Ir/ are distinguished by their agreement class. When these verb forms appear in main clauses, they take Class-II person markers (e.g. *men, sen,* etc.). In subordinate or relative clauses, the person of the subject of the verb is realized suffixally through the use of Class-I markers (e.g. *-(I)m, -(I)ŋ,* etc.).

(179) <u>main clauses</u>

(men) ažɨldaar men	*(men) ažɨldaan men*
'I' 'work'-P/F 1	'I' 'work'-PAST.I 1
'I will work'	'I worked'

(180) <u>suboordinate clauses</u>

(meeŋ) ažɨldaarɨm deeš	*(meeŋ) ažɨldaanɨm deeš*
'I' 'work'-P/F -1 'due.to'	'I' 'work'-PAST.I -1 'due.to'
'due to my working'	'due to my having worked'

Participles are also commonly found in a range of case forms to create subordinate/complement clauses of various types. These are discussed in 9.3 and 9.4 below.

8.7 Auxiliary verbs

As in most Turkic languages, auxiliary verbs play an important role in the Tuvan verbal system, performing a wide range of functions. At least two dozen semantically bleached auxiliary verbs occur. They index various aspectual and modal categories. Auxiliaries follow a converb form of the lexical stem (the /-(I)p/ converb or the /-A/ converb).

The most common auxiliary is *tur,* which has the basic lexical meaning 'stand'. As an auxiliary, *tur* contibutes no lexical meaning to converbs with which it occurs:

(181)

saktɨp tur men	*xalɨp tur men*
'remember'-CV AUX 1	'run'-CV AUX 1
'I remember'	'I'm running'

Note, however, that the long form of the same verb, *turar* 'stand', yields a continuative present meaning when it serves as an auxiliary:

(182)

<u>simple present</u>		<u>continuative present</u>
ol Moskvada čurttap tur	vs.	*ol Moskvada čurttap turar*
'he' 'M.'-LOC 'live'-CV AUX-P/F		'he' 'M.'-LOC 'live'-CV AUX-P/F
'he lives in Moscow'		'he is living in Moscow'

Some auxiliary verbs, usually high-frequency ones, yield a different meaning when used with different converb types:

	verb	lexical meaning	auxiliary meaning
(183)	*al-*	'take, get'	1. self-benefactive voice (/-(I)p/ converb) 2. capabilitive mood (/-A/ converb)

(184) *ber-* 'give' 1. inchoative (with /-A/ converb)
2. benefactive voice (with /-(I)p/ converb)

		with /-(I)p/ converb	with /-A/ converb
(185)	***al-***	*ol bižip* **al***gan*	*ol bižiy* **al***gan*
		'he' 'write'-CV SBEN-PAST.I	'he' 'write'-CV CAP-NEG.FUT
		'he wrote it down'	'he was able to write'

		with /-(I)p/ converb	with /-A/ converb
		Inchoative mood	***Benefactive mood***
(186)	***ber-***	*ažɨldap* **ber***dim*	*ažɨlday* **ber***dim*
		'work'-CV INCH-PAST.II.1	'work'-CV BEN-PAST.II.1
		'I began to work'	'I worked' (on behalf of someone)

Other auxiliaries have the same meaning with either converb type, but take on different meaning in conjunction with different semantic classes of verbs:

(187) *bar-* 'go' 1. completive/perfective action
2. translocative action (across space)

	completive/perfective	translocative
(188)	*ol bižip bargan*	*ol canɨp bargan*
	's/he' 'write'- CV AUX-PAST.I	'he' 'go.home' AUX-PAST.I
	'he wrote' *or* 'he finished writing'	'he went home'

Most auxiliaries express just one meaning:

	verb	lexical meaning	auxiliary meaning
(189)	*bol-*	'be'	possibilitive mood
	čit-	'sit'	capabilitive mood
	čoru-	'go'	imperfective or durative aspect
	egele-	'begin'	inchoative aspect
	kag-	'leave'	'already'
	kel-	'come'	cislocative
	kir-	'enter'	completive or terminative aspect
	kör	'see'	attemptive mood
	olur-	'sit'	imperfective aspect

8.8 Auxiliary nouns

Tuvan has a small class of auxiliary nouns (AUXN) which perform a copular function. They are: *čüve* 'thing' *kiži* 'person', and *ulus* 'people'. Note that the following sentences lack a sentential verb. They end in a noun which does not signify its usual referent (e.g. 'thing' or 'people'), but functions purely as an auxiliary.

(190) *bis silerge kelgenner ulus bis*
'we' 'you'-PL-DAT 'come'-PAST.I-PL AUXN 1-PL
'We've come to (visit) you'

(191) *ol yeyl unifersitɨ kayɨ xoorayda čüvel*
'that' 'Yale' 'univ.' 'which' 'town'-LOC AUXN-DEIC
'What town is that Yale university in?'

9.0 Syntax

The basic constituent order within a sentence is subject-object-verb.

(192) *sen süt ižer sen*
'you' 'milk' 'drink'-P/F 2
'You drink milk'

Within the noun phrase, the maximal expansion is [Demonstrative-Possessive-Adjective Phrase-Noun].

(193) *bo meeŋ eŋ uluɣ akkɨm*
'this' 'I'-GEN 'most' 'big' 'elder.brother'-1
'this my biggest elder brother'

Within the verb phrase, the verb comes in final position, preceded by subject, indirect object, and direct object in that order.

(194) *men bo nomnu silerniŋ bažɨŋɨŋarda kördüm*
'I' 'this' 'book' 'your(pl)'-GEN 'house'-2PL-LOC 'see'-PAST.II-1
'I saw this book in your house'

(195) *ol menden taakpɨ aldɨ*
'he' 'I'-ABL 'tobacco' 'take'-PAST.II
'he took tobacco from me'

Some freedom of order is permitted among the arguments for purposes of focus. Generally, objects that move closer to the verb become more focused (compare 195 and 196). The verb, however, remains always in final position.

(196) *ol taakpɨ menden aldɨ*
'he' 'tobacco' 'I'-ALL 'take'-PAST.II
'It was from me that he took tobacco'

9.1 Copulae

With nominal and copular sentences present tense forms are marked with either a copula, an evidential particle, a negative marker, an interrogative/deictic enclitic or some combination of these. The distribution of these three items and the fact that they are obligatory for most speakers in existential sentences suggest that they individually or in combination perform the function of a copula. See also auxuliary nouns which may function as copulae (§6.0)

(197) *xlep bar be* / 'bread' 'be'-P/F Q / 'Is there bread?'
(xlep) bar(dɨr) / 'be'-P/F -(EMPH) / 'there is (bread)'
(xlep) čok(tur) / NEG-(EMPH) / 'there isn't (bread)'

(198) *xlep dɨr* / 'bread'-EVID / 'Is there bread?'
xlep čok(tur) / 'bread'-NEG-(EMPH) / 'there isn't (any) bread'
čok-tur / NEG-EVID / 'there isn't any'

(199) *böxün iziɣ dir* / 'today' 'hot' EVID / 'It's hot today'
** böxün iziɣ* / 'today' 'hot'

(200) *bažɨŋda üš kiži bar*
'house'-LOC 'three' 'people' 'there.is'
'there are three people in the house'

(201) *čayɨm čok* / 'free.time'-1 NEG-COP / 'I have no (free) time'
aškam čok / 'money'-1 NEG-COP / 'I have no money'

In all non-present indicative constructions the appropriately inflected form of an auxiliary, e.g. *tur* or *bol,* follows the nominal element.

(202) *čayɨm čok bolgan* / 'free.time'-1 'no' 'be'-PAST.I / 'I had no (free) time'
aškam čok turgan / 'money'-1 'no' AUX-PAST.I / 'I had no money'

9.2 Coordination

Coordinators (e.g. *bolgaš* 'and', *azɨ* 'or', etc) are found in both nominal and verbal constructions.

(203) *ačam aŋnaar bolgaš balɨktaar* / 'father'-1 'hunt'-P/F 'and' 'fish'-P/F / 'My father hunts and fishes'
aŋnar bolgaš balɨktar bar / 'animal'-Pl 'and' 'fish'-Pl COP / 'There are (both) animals and fish'

The /-GAš/ converb may be used with semantically conjoined verbs, requiring a sequential interpretation.

(204) *Kyzyldaaš čedip keldim*
'go to Kyzyl-CV 'return'-CV 'come'-PAST.II-1
'(After) Having gone to Kyzyl, I returned.

Simple juxtaposition of two or more verbs in the /-(I)p/ converb form also may be used in a coordinative function. In such cases two or more converbs share a single overt auxiliary.

(205) *ol turup, xeptenip, čip turdu*
'he' 'stand'-CV 'get.dressed'-CV 'eat'-CV AUX-PAST.II-3
'He got up, got dressed, and ate'

9.3 Subordination

The primary means of forming subordinate clauses is through the use of case marked verbal forms. Most often, these are participles and converbs. Participles in subordinate clauses may appear in an uninflected form, or may appear with case-marking, person-marking, auxiliary nouns, postpositions, etc.

(206) *(men) tɨvalaar bilir men*
'I' 'speak.tuvan'-P/F 'know'-P/F 1
'I know how to speak Tuvan'

inek saarɨ, seeŋ xereeŋ
'cow' 'milk'-P/F-3 'you'-GEN'duty'-2
'It's your duty to feed the dog'

Present/future verb with ACC case marking and no person marking:

(207) *(men) čuɣaalaarɨn bilir men*
'I' 'speak'-P/F -ACC 'know'-P/F 1
'I know how to speak'

The subject of a subordinate clause is indexed by using Class-I markers (*-ɨm, ɨŋ,* etc.) on the participle, as in all the following examples. The most common means of forming temporally subordinate clauses is by Past.II participles with LOC case marking.

(208) *(sen) kelgeniŋde kördüŋ*
'you' 'come'-PAST.I-2-LOC 'see'-PAST.II-2
'When you came, you saw'

(209) *men am=daa törütümeenimde*
'I' 'now'-EMPH 'be.born'-NEG-PAST.I-1-LOC
'When I wasn't yet born' (B 204)

Past.II participle with Class-I person marker and DAT case marking:

(210) *seeŋ ɨrlaarɨŋge*
'you'-GEN 'sing'-P/F -2-DAT
'At/for your singing'

(211) *meeŋ kelgenimge ačam amɨraar*
'I'-GEN'come'-PAST.I-1-DAT 'father'-1 be.pleased'-P/F
'My father will be happy when I come'

Past.II participle with Class-I person marker and ALL case marking:

(212) *(meeŋ) kelgenimden, eki udaan men*
'come'-PAST.I-1-DAT 'well' 'sleep'-PAST.I 1
'Since arriving, I've slept well'

Past.II participle with Class-I person marker, followed by postposition with ABL case marking:

(213) *nomčuvastaanɨm užurundan šilgeldeni baɣay dužaadɨm.*
'read'-NEG-CES-PAST.I-1 'reason'-ABL 'exam'-NOM 'badly' 'meet'-PAST.II-1
'Because I stopped reading, I did badly on the exam'

An overt subject NP in such constructions may appear either in the NOM (214), the GEN (215) or rarely, the ACC case.

(214) *(men) kelgenimde ažɨldaar men*
1-NOM 'come'-1-LOC 'work'-P/F 1
When I come, I work.

(215) *(meeŋ) kelgenimde ažɨldaar men*
1-GEN 'come'-1-LOC 'work'-P/F 1
When I come, I work.

Note that these sentences (186, 187) *do not* refer to an action in the *past*, even though the participle appears in the PAST.I.

Participles are also frequently used with a range of postpositions, forming both causally and temporally subordinate clauses.

(216) *bodu kelir užun bižik ɨtpadɨ*
'self'-3 'come'-P/F 'for' 'letter' 'send'-NEG-PAST.II
'because he intended to come himself, he didn't send a letter' (Sh 1993: 72)

Converb forms constitute another important means of forming subordinate clauses in Tuvan. For most of the formations traditionally considered converbs, clausal subordination is their primary or only function in the language.

(217) *xem uglatkaš, parom čorbastaan*
'river' 'flood'-CV 'ferry' 'go'-NEG.P/F -VSF-PAST.I
'because the river flooded, the ferry stopped running'

A temporally subordinate clause may be formed with the /-GAš/ converb plus GEN case marking.

(218) *kelgeštiŋ, menče dolgavɨt*
'come'-CV-GEN 'me'-ALL 'phone'-IMP-PERF
'when you return, call me'

This concludes the grammatical overview of Tuvan. Further examples and a more descriptive detail may be found in Anderson & Harrison 1999, of which this chapter represents an abridged version.

Chapter Two: Distinctive pitch in Tuvan

0. Introduction

Tuvan employs an unusual vowel contrast of a kind not found in genetically or areally related languages. In the linguistic literature, this contrast has been labeled 'pharyngealization,' implying that it derives from a pharyngeal constriction. However, Tuvan vowels have not previously been subjected to an instrumental analysis that could shed light on their true acoustic nature. The analysis presented here clarifies for the first time the special properties of these vowels. Data presented herein are used to argue that it is contrastive *low pitch*, not pharyngeal constriction, that serves as the salient acoustic characteristic of these vowels.

The patterning of low pitch raises a number of theoretical issues regarding gestural timing and duration. First, low pitch interacts with vowel duration by effectively neutralizing the long vs. short distinction normally available to Tuvan vowels. Second, the timing of supra-laryngeal gestures (e.g. stop closures) seems to be critically coordinated with respect to—and therefore controlled by—the pitch event itself. We demonstrate that low pitch has an intrinsic duration that is about as long as the duration of a typical syllable in the language. We argue that low pitch should thus be thought of as an autonomous *articulatory gesture* (Browman and Goldstein 1989, 1992a). Further, evidence that the pitch gesture has intrinsic duration speaks to the more general issue of whether duration must be globally specified as a scaffolding (that is, a skeleton), or whether duration may also may be purely-event driven. The Tuvan pitch gesture seems to require regulated, event-driven duration.

Lastly, we raise, but do not definitively resolve, the question of whether Tuvan should be thought of as having a pitch accent or a tone system. We suggest that that low pitch is more akin to tone in its acoustic and phonological behavior. Tuvan would seem to represent, on this view, a paradoxical one-tone language. If so, low pitch might represent an incipient tone system as well as the limiting case of what a tone system can be. Namely, it would have to be construed as a system with just a single ‘tone’, set in a privative opposition to no tone at all.

1.0 Vowel system

Tuvan has a typically Turkic eight-vowel inventory, plus contrastive length. In this chapter, we employ the IPA length diacritic rather than the geminate vowel symbols used elsewhere throughout this dissertation We also employ IPA symbols [y], [ø], [ɯ], instead of the turcological symbols [ü], [ö], [ɨ] used in other chapters.

(1) Tuvan vowels

	front		***back***	
	unround	***round***	***unround***	***round***
high	i	y	ɯ	u
non-high	e	ø	a	o

Besides contrastive length, Tuvan and the closely-related Tofa (Rassadin 1971) employ a vowel contrast that has been termed ‘pharyngealization’ in the Russian linguistic literature. The ‘pharyngealized’ vowels have been considered to constitute a distinct series, thus increasing the Tuvan vowel inventory from sixteen phonemes (including both short and long vowels) to twenty-four (Iskhakov & Pal’mbakh 1961). We shall refer to these vowels as low pitch and adopt the IPA diacritic for low pitch [è] in forms cited herein.

There is no obvious phonological conditioning to which the emergence of these vowels might be attributed. The origin of ‘pharyngealization’ has been ascribed rather speculatively to influence from a now extinct Yeniseian substratum (Verner 1972, Menges 1955). Tuvan has low pitch vowels where related Turkic languages such as Xakas, Sakha and Dolgan have, for the most part, short vowels.

(2)

Dolgan	Tuvan	
at	àt	‘horse’
et	èt	‘meat’

The label ‘pharyngealized’ was first applied in the Russian descriptive tradition to the Tungus language Even (Nadelaev 1986, Novikova 1960). Even reportedly has a distinct series of pharyngeally *constricted* vowels. The label ‘pharyngealized’ was apparently extended to Tuvan mainly on the basis of external, auditory observation, and perhaps on the analogy of Even. No instrumental phonetic or acoustic studies were carried out to justify the use of the label. The term has since been widely adopted in the turcological literature (e.g. Shcherbak 1994) and has persisted in subsequent discussions of Tuvan (Kreuger 1972, Tatarintsev 1986).

An early Tuvan grammarian (Lopsaŋ 1941) described these special vowels as being articulated with a degree of *kargyra*, referring to a style of Tuvan throat-singing that involves sustained periods of controlled, very low pitch or creaky voice (Levin & Edgerton 1999). An alternative term, ‘glottalized’—apparently meant to imply the presence of creaky voice or vocal fry—has also been used occasionally, and sometimes even interchangeably, with the term ‘pharyngealized’ (Nadelaev 1986). Clearly, the confusion over the correct label for Tuvan vowels has only increased in the absence of any instrumental investigation of the actual acoustic nature of these vowels. Moreover, the appropriateness of these labels has gone unquestioned.

We note that the term ‘pharyngealized’ in the Russian tradition functions as a cover term for a wide range of pharyngeal/laryngeal phenomena. When Novikova (1960) adopts the term in reference to Even vowels, she is clearly referring to a constriction in the pharynx. Moreover, she provides mid-saggital x-ray images to justify the use of the term. No such data has been collected for Tuvan, and a pharyngeal constriction cannot thus be completely ruled out. Nonetheless the acoustic data we collected show none of the expected acoustic correlates of a pharyngeal constriction (e.g. higher F1 or higher F2 for back vowels). They do show clearly that these vowels have distinctive F0. On the basis of both positive and negative acoustic evidence, we propose that pitch alone, not pharyngeal constriction, accounts for the distinctive nature of these vowels. The pitch contrast described herein for Tuvan has not been correctly identified as such nor has it been previously analyzed in the linguistic literature. The vowels are referred to hereafter as low pitch vowels.

2.0 Contrastive nature of low pitch

Contrastive low pitch is restricted in its distribution to initial syllables and to a subset of the lexicon. According to Bicheldei (1993), several thousand words in Tuvan have a ‘pharyngealized’ vowel in the first syllable. This figure may be accurate in a philological sense, for literary Tuvan, but we estimate the lexicon of most Tuvan speakers contains no more than a few hundred words with distinctive pitch. These constitute a semi-closed class of lexemes, though new words have appeared fairly recently in the form of loanwords from Russian (e.g. sàpɯk ‘boot’, mèdaːl ‘medal’). Speakers of certain regional dialects of Tuvan (e.g. Milk-Lake dialect) tend to add low pitch much more frequently. Anecdotal evidence suggests that speakers of geographically peripheral Tuvan dialects (e.g. those dialects spoken in Mongolia) may lack it entirely (Seren 1996). Low pitch is therefore a somewhat variable and unstable feature across Tuvan dialects (Martan-ool 1986, Sat 1987).

Low pitch constitutes a minimally contrastive feature of monosyllabic and polysyllabic lexemes. Vowels of the initial syllable only may exhibit a ternary contrast: short / low pitch / long. The following exemplify (near) minimal triplets.

	short		**low pitch**		**long**	
(2)	et	‘leather’	èt	‘meat’	eːt	‘delta’
	edi	‘leather’-3SG	èdi	‘meat’-3SG	eːdi	‘delta’-3SG
	at	‘name’	àt	‘horse’	aːs	‘mouth’
	at	‘name’	àrt	‘mountain pass’	aːr	‘heavy’
	ege	‘completely’	ège	‘beginning’	eːgi	‘rib’
	arga	‘means’	àrga	‘forest’	aːrɨg	‘illness’
	tʃoqta	‘go uphill’-IMP	tʃòqta	‘miss’-IMP	tʃoːqta	‘not far’
	udup	‘sleep’-CV	ùdup	‘win’-CV	uːttun	‘break’-IMP
	uze	‘tear apart’-CV	ùze	‘bump’-CV	uːʒe	‘frozen meat’

We note that Tuvan has very few monosyllabic words that are open syllables (CV). None of these are reported to bear low pitch. Many polysyllabic lexemes bear low pitch, again only on the first syllable (3).

(3)	àdɯɣ	‘bear’
	kìʒi	‘person’
	kàmgalal	‘defense’

òskunar	'to let go'
kỳʃkyl	'thrush'
kèdeːr	'to sneak up on'-P/F
ø̀skeːr	'to stray'-P/F
mègelep	'cheat'-CV

Some polysyllabic lexemes contrast minimally with homophones containing short vowels.

(4)	bùduq	'twig'
	buduq	'ink'
(5)	qàdɯq	'porridge'
	qadɯq	'health'

Lexemes that bear distinctive low pitch generally retain it in suffixed forms in standard Tuvan (6) This may be subject to dialect and contact-induced variation.[8]

(6)	àt	'horse'-1SG
	àdɯm	'horse'-1SG
	àtka	'horse'-DAT
	àdɯʊɯstan	'horse'-1PL-ABL
	àttarɯʊɯstan	'horse'-PL-1PL-ABL

3.0 Acoustics of low pitch

Low pitch is generally at the low end of a speaker's modal range. For this study, a 26-year old male speaker recorded three hundred vowel tokens spoken in isolation. These were then played back to the speaker and to another subject to ensure that they were judged to be well-formed. From this corpus, four tokens of each vowel quality in short, long and low-pitch varieties were selected for further analysis. Using SoundEdit™, we calculated F0 by measuring glottal periods at successive intervals of 20 milliseconds beginning from the first regular glottal pulse of each waveform. The first and last measurements for each waveform were taken to calculate the mean beginning and ending F0 for each class of vowels. For all vowels, the first measurement was made at a point 20 ms from the beginning

[8] Greg Anderson (Anderson & Harrison 1999) reports that in the speech of younger, urban Tuvan speakers, the low pitch vs. modal pitch distinction is often neutralized in inflected forms. Such neutralization is evident in the following forms spoken by a 14 year old girl: *àt* 'horse' vs. *ad-ɯm* ['horse'-1SG] 'my horse' *at-ka* ['horse'-DAT] 'for the horse', *at-tan* ['horse'-ABL] 'from the horse'. Compare these to the same forms spoken by an 87 year old woman, all of which retain low pitch: *àt, àd-ɯm, àt-ka, àt-tan*.

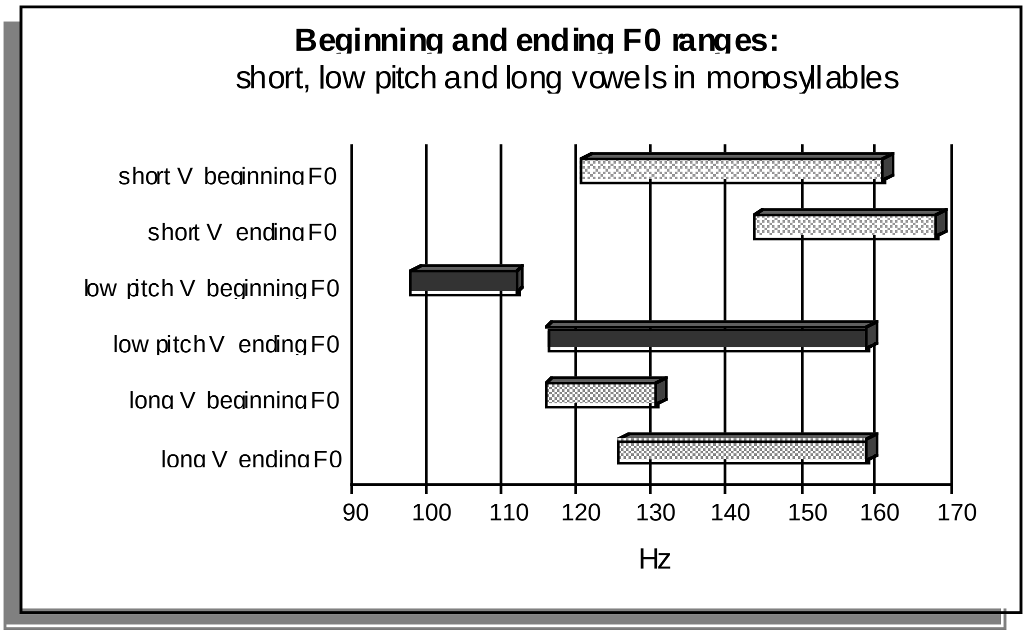
of the waveform. The last measurement was made at 80 ms for short vowels and 260 ms for long vowels. Average beginning and ending pitch measurements were as follows:

(7)

	mean beginning F0	mean ending F0
short vowels	141 Hz	152 Hz
low pitch vowels	106 Hz	138 Hz
long vowels	122 Hz	144 Hz

Somewhat more revealing than the mean F0 values are the range of F0 values for the beginning and ending of each vowel type, as shown in the following bar graph.

(8) Table A

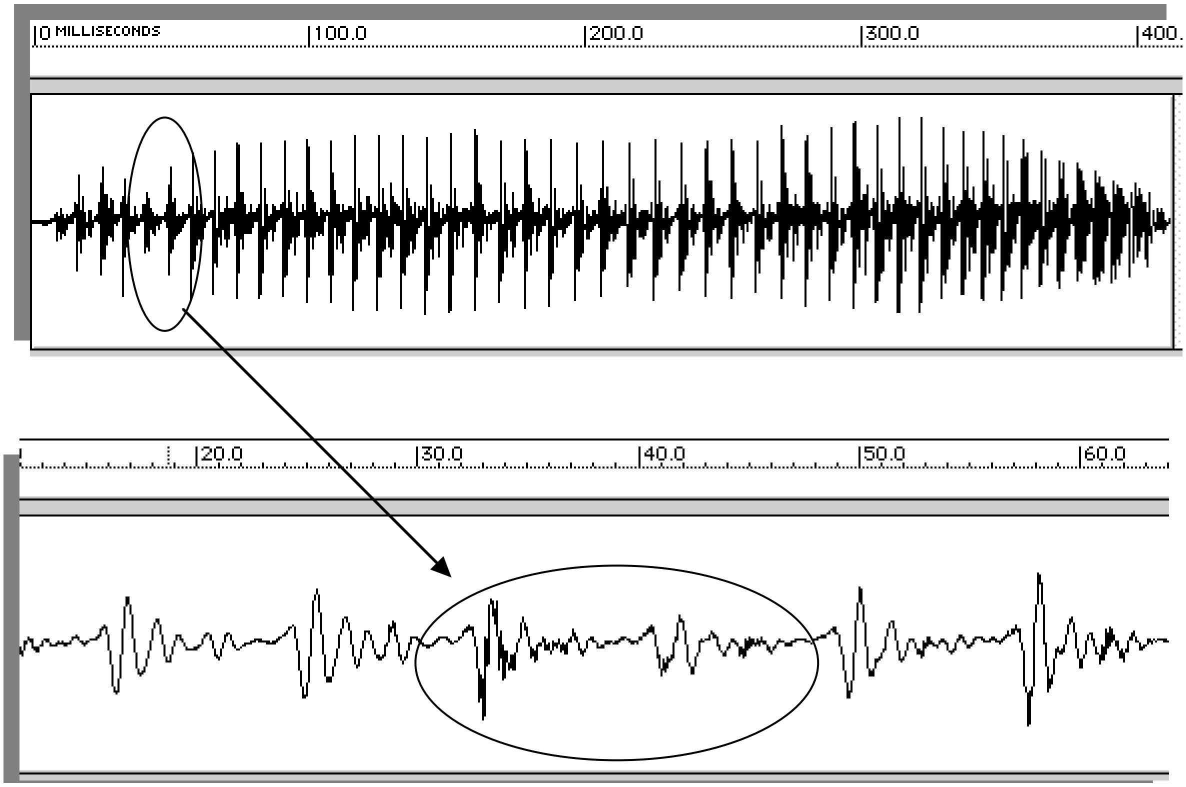


All vowel types show a rise in pitch in monosyllables. The three types show considerable overlap in their ending F0. They differ dramatically, however, in their beginning F0 values. Beginning F0 for short vowels is higher than that of long vowels, though the two ranges overlap. Crucially, the low pitch vowels are shown to be truly exceptional in that their beginning F0 range is considerably lower than that of all other pitch ranges charted here. Further, the beginning and ending F0 ranges of low pitch vowels do not overlap at all.

3.1 An acoustic correlate of low pitch

The low pitch vowels seem impressionistically to involve glottalization or creaky voice. They are commonly compared by casual observers to a style of Tuvan throat signing that relies on creaky voice. To test this impression, we looked for irregularities in the waveform characteristic of such a glottal posture (phonation mode). The irregular pulses seen in the waveform for [àt] 'horse' between 30 and 50 ms are characteristic of the irregularities we found.

(9) Table B: Irregular glottal pulses indicative of creaky voice



We found such aberrations in the wave patterns for less than 15% of the tokens produced by three different speakers. Other speakers showed no such aberrations in their low pitch tokens. We thus interpret creaky voice as a possible but not essential correlate of contrastive low pitch. Evidence that phonation mode is acoustically salient for these vowels is unpersuasive. For example, we found no evidence that these vowels exhibit spectral tilt less than that of other vowels. Reduced spectral tilt is an expected acoustic correlate of creaky voice (Kirk, Ladefoged & Ladefoged 1993). Low F0, on the other hand, is the essential and invariant acoustic correlate of distinctive low pitch vowels.

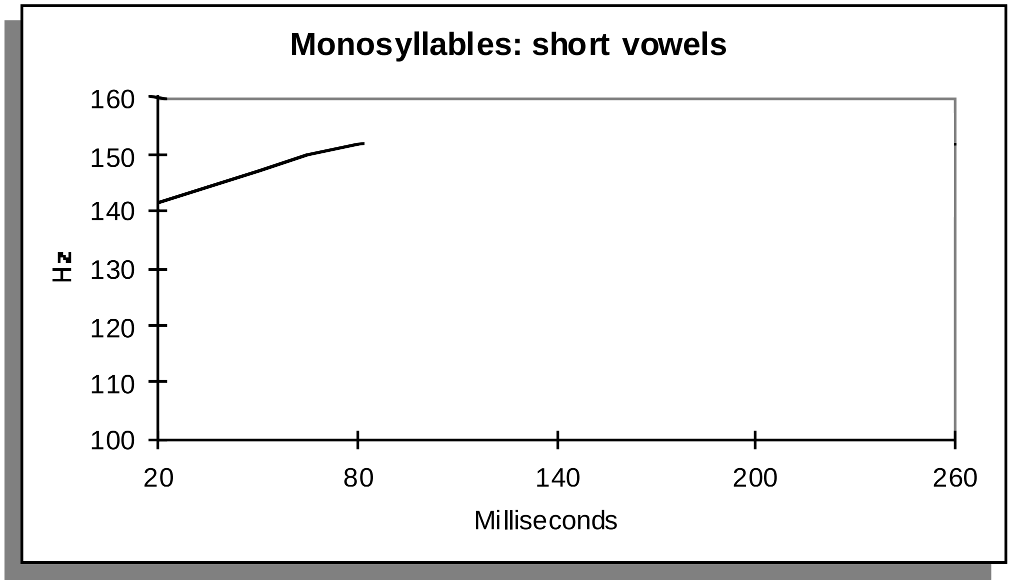
4.0 Pitch contour, duration and syllable structure

Based on the pitch values such as those shown in Table A, we proposed that the invariant acoustic property of these vowels is low pitch. Specifically, this entails pitch at the extreme low end of the modal range, below that which the speaker normally employs in speech. In this section, data is presented to demonstrate that low pitch is always produced in the context of a characteristic pitch contour. This contour results from the speaker's return to modal pitch after a period of low pitch. In monosyllables, low pitch typically is realized only on the first fifty to sixty percent of the vowel. The second half of the vowel contains a rapid, rising pitch transition. Acoustically, this sounds like a rising contour tone, perhaps somewhat like the falling-rising 3rd tone of Mandarin. The contour of Tuvan low pitch vowels begins with a stable plateau of low pitch. Towards the end of the segment's duration, pitch rises sharply, but never levels off to achieve the second plateau. Following a proposal by Yip (1999), we characterize this sharp return to modal pitch as 'updrift', because it appears to be phonologically targetless (cf. Table E below).

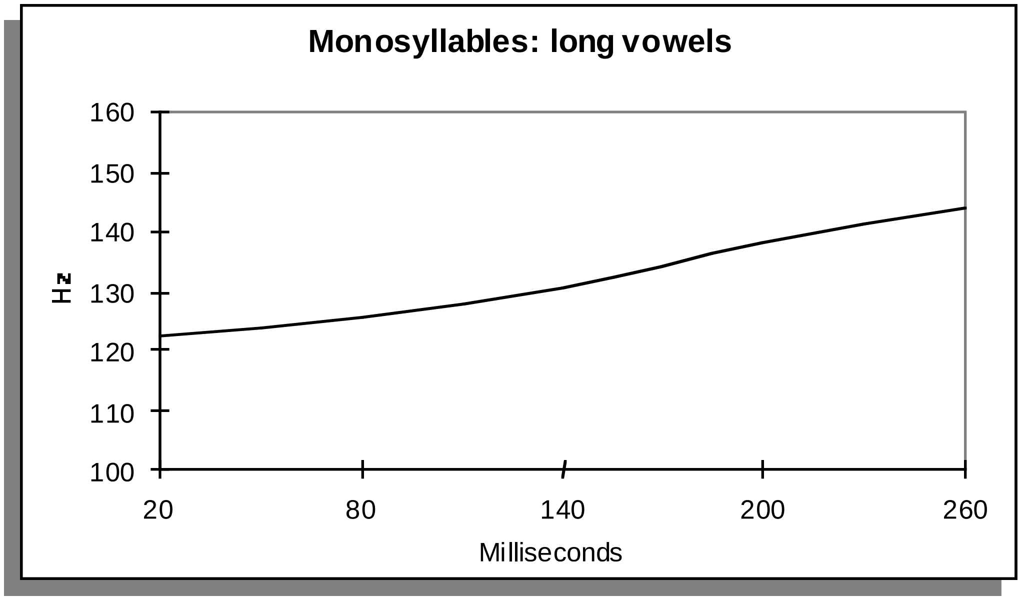
4.1 Pitch contrast in monosyllables

The following graphs illustrate the mean pitch contours for 36 tokens each of short, low pitch and long vowel in monosyllabic contexts. The tokens were read by a 26 year old male speaker who produced them in isolation as part of a wordlist. The graphs are uniform with respect to pitch range (from 100 Hz to 160 Hz) and duration (from 20 Ms to 260 Ms). Not all of the pitch contours depicted fill the entire range of the graph. However, to facilitate comparison across words and vowels of differing duration, all the graphs extend to a uniform length. Table C depicts short vowels in monosyllables, which have a duration of less than 100 Ms. They rise slightly in pitch, from about 140 Hz near the beginning to about 150 Hz near the end.

(11) Table C



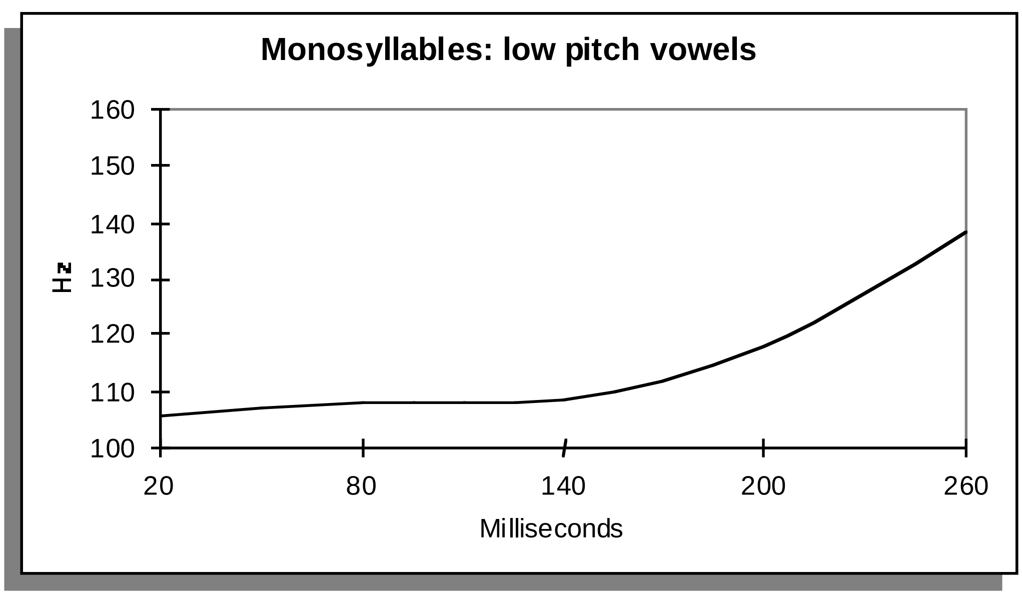
Long vowels in monosyllables average nearly 300 Ms in duration. Like short vowels, they also exhibit a gradual pitch rise, from about 122 Hz to 140 Hz.



(12) Table D

Low pitch vowels in monosyllables are equal or greater in duration to long vowels. They exhibit a steady plateau of very low pitch in the first half of the vowel, followed by a relatively rapid rise in the second half.

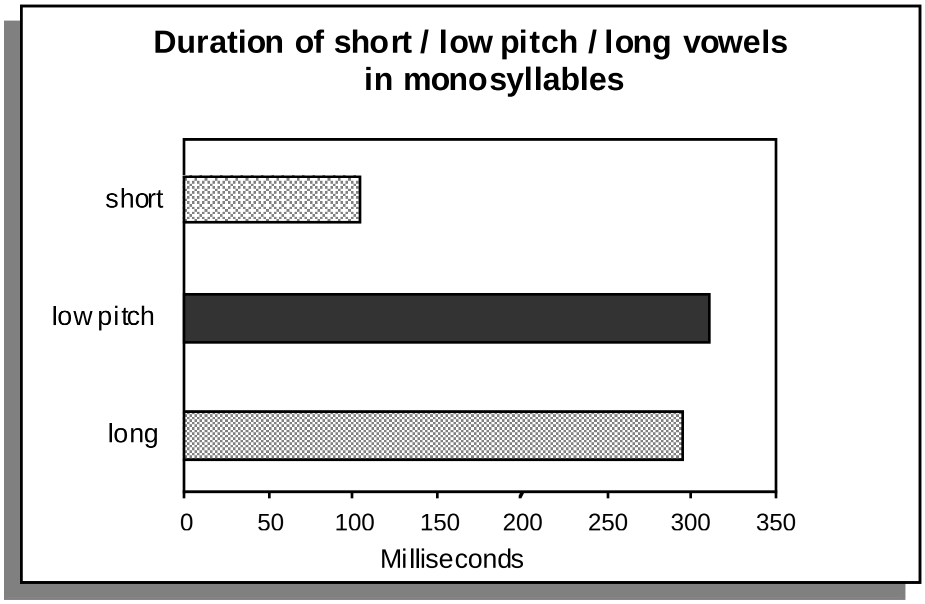
(13) Table E



4.2 Vowel duration in monosyllables

The length contrast in monosyllables is quite robust. Short vowels average only about one-third the duration of long vowels in monosyllabic contexts. Low pitch vowels also show extremely long duration, equal to or even slightly longer than that of long vowels.

(14) Table F

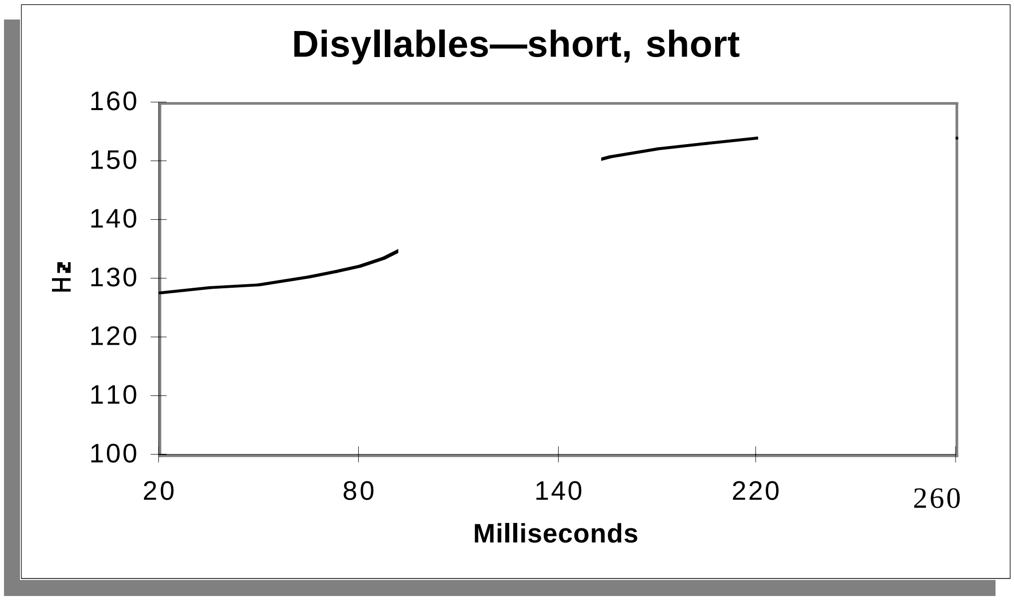


4.3 Pitch contrast in polysyllables

The pitch contours depicted above for monosyllables turn out to be quite similar in polysyllables spoken in an isolated context. The pitch contour differs in that it spreads out across two syllables. In the following graphs, the gaps in the pitch contours represent the period where the consonant closure intervenes and no pitch measurement was taken. Each pitch contour line is divided into two halves, representing the first and second syllable vowels, respectively. The time values along the X-axis run continuously through both syllables.

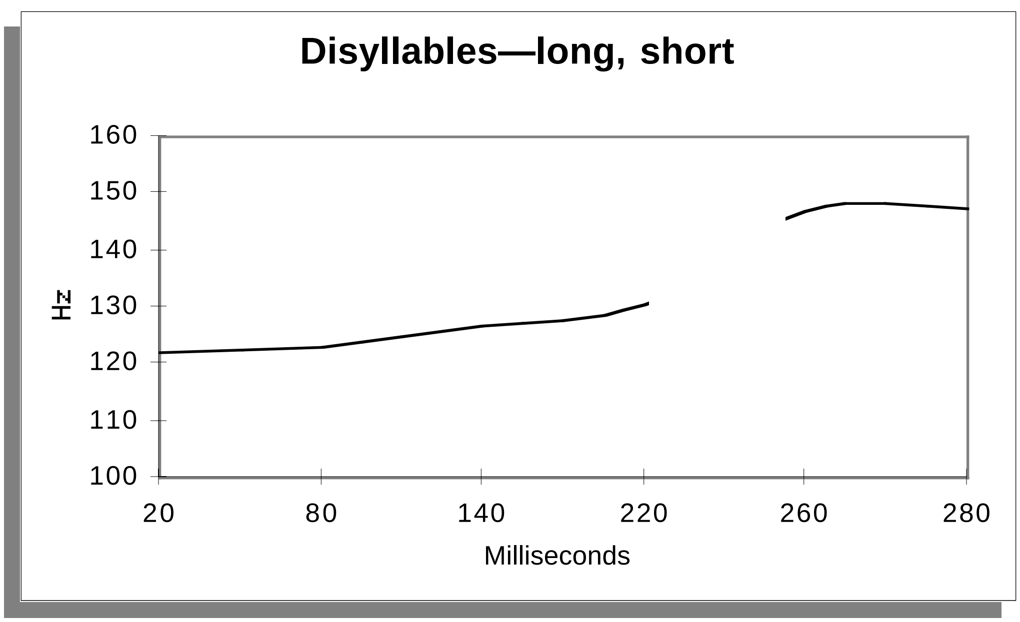
Short vowels in disyllables show a rise in pitch from about 130 Hz in the first syllable to 150 Hz in the second.

(15) Table G



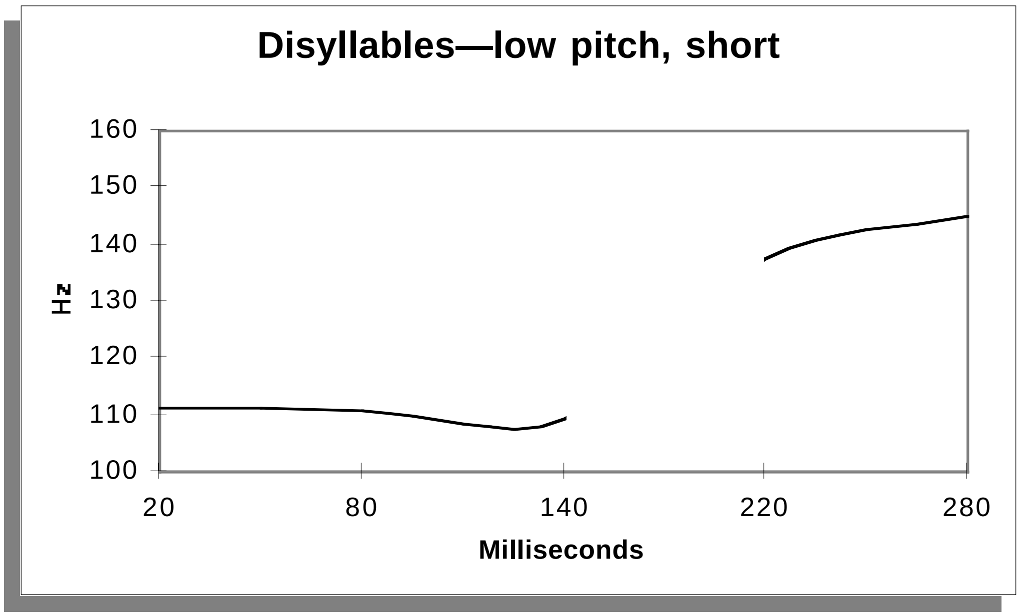
Long vowels also show a slight rise in the first syllable of a disyllable. Short vowels in the second syllable, as in Table G, rise from 130 to 150 Hz.

(16) Table H



Low pitch vowels in disyllables exhibit a similar pitch contour spread out across two syllables. The low pitch vowels remain uniformly low throughout the initial syllable. The rise in pitch is deferred to the second syllable. The updrift in pitch is thus somewhat obscured by the intervening consonant articulation.

(17) Table I

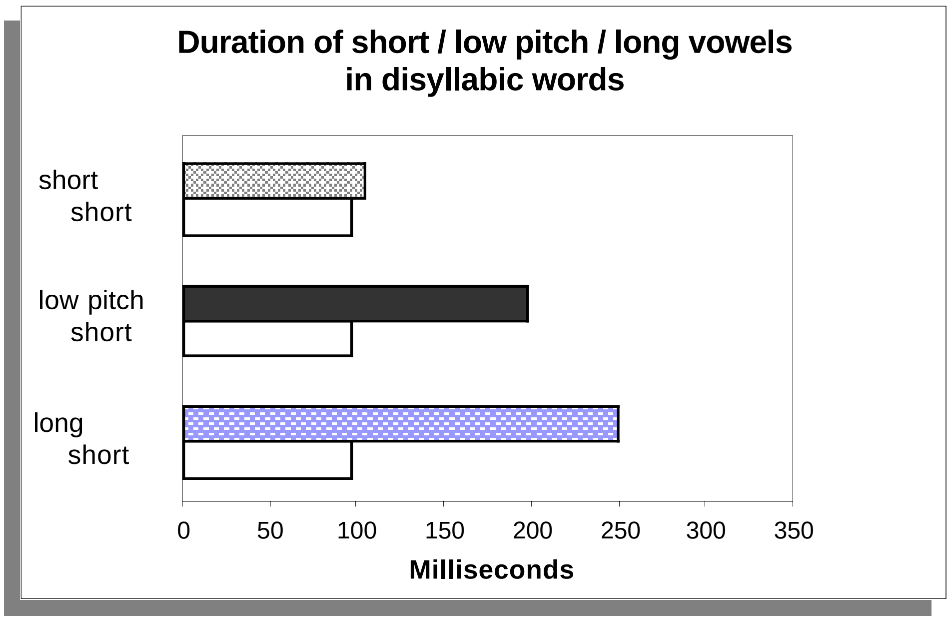


4.4 Duration in polysyllabic contexts

Low pitch vowels exhibit a chameleon-like character with respect to duration. Low pitch interacts with vowel length by essentially neutralizing the long vs. short distinction normally available to Tuvan vowels. The intermediate duration of these vowels in certain contexts was first noted by native Tuvan grammarian S. Lopsaŋ (1941). He described them as being ‘half-long’, within a system that has a long vs. short opposition. The term ‘half-long’ was adopted in subsequent descriptions (e.g. Iskhakov & Pal’mbakh 1972). These accounts fail, however, to note that the duration of low pitch vowels is not consistently of intermediate length, but rather is contextually variable. For the present analysis, the contextually determined length in mono- vs. polysyllabic contexts is important.

As noted above, low pitch vowels in monosyllables have a duration equal to or greater than that of long vowels. In polysyllables, their duration decreases significantly to a phonetic length mid-way between that of long and short vowels. We propose that duration of the vowel is determined by the intrinsic duration of the pitch event.

(18) Table J



As shown in Table J, short vowels in the first syllable of disyllables had a mean duration of 104 ms, while long vowels lasted 249 ms. Low pitch vowels patterned midway between long and short vowels, having a mean duration of 197 ms. In the disyllabic contexts depicted above, the second syllable always contains a short vowel, represented by the white bars. These short vowels exhibit a uniform duration of about 140 ms.

Phonological length contrast in Tuvan results in very large differences in phonetic length: short vowels average less than 40% of the duration of long vowels in monosyllabic contexts (Table F). In disyllables, short vowels again have less than 40% of the duration of long vowels. This robust length contrast appears to be neutralized in the presence of distinctive pitch.

In monosyllables, the phonetic target and subsequent return to modal pitch must be achieved within a single vowel; this vowel is therefore of extremely long duration (Table F). In polysyllables, only the phonetic target (low pitch) must be attained, while the updrift is deferred to the next syllable. The low pitch vowel thus requires less duration, and is about 20% shorter than long vowels in the same context. Still, the single target requires greater duration than for a true short vowel, such that low pitch vowels are still about 25% longer than short vowels in the same context.

4.5 Pitch and duration in phrasal contexts

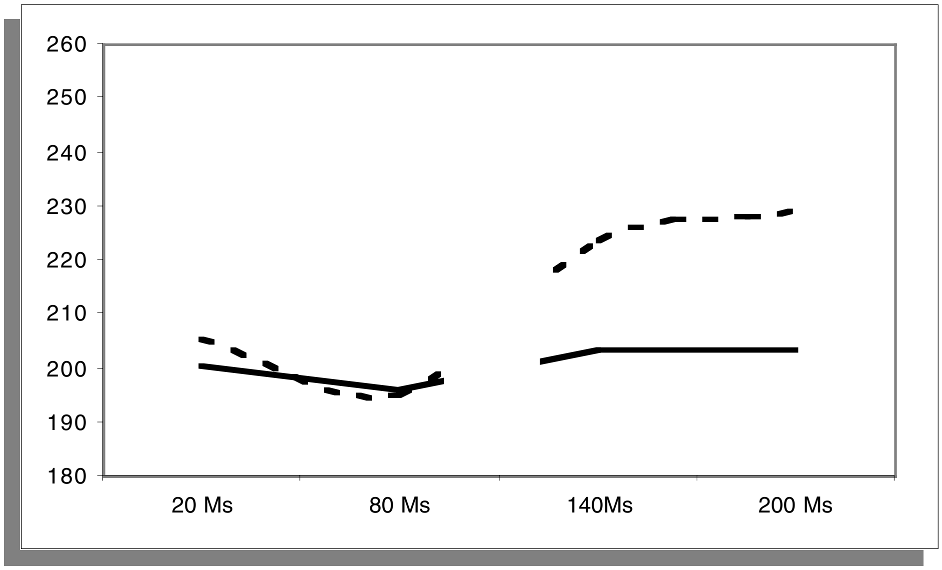
We have shown that disyllabic tokens produced in a controlled context (a wordlist) show stable low pitch throughout the first syllable and pitch updrift in the second (Table I). Low pitch vowels also undergo considerable shortening in a disyllablic word, approaching a duration midway between that of long and short vowels (Table K). Despite this shortening effect, we found that the period of low pitch remains about the same, e.g. between 160 and 200 ms. When disyllabic tokens containing low pitch are put in a sentence context, we find that the shortening effect is even more dramatic. The low pitch vowels may be reduced to a duration equaling that of the short vowels. Strikingly, the period of low pitch appears to retain its intrinsic duration, and the plateau reflecting the low pitch event carries over into the first part of the second syllable.

Table K (20) depicts pitch contours for a minimal pair of two words that differ only in pitch.

(19) qadɯː 'health'-3
qàdɯː 'porridge'-3

We recorded these tokens as produced by a 21-year old female speaker in both a wordlist and a sentence context. The tokens show a higher overall F0 range compared to the male speaker's pitch shown in tables C, D, E, G, H and I. The values shown in table K represent the average pitch for ten tokens each of [qadɯɯ] 'health-3' and [qàdɯɯ] 'porridge-3' produced in isolation by 21-year old female speaker.

(20) Table K: Pitch contour of [qadɯː] (solid line) and [qàdɯː].



Somewhat surprisingly, the putatively 'low pitch' tokens (dotted line) had an F0 not significantly lower than that of the modal pitch tokens (solid line) when produced in isolation in a wordlist. We found a similar sameness in pitch for tokens produced in a wordlist by the 26-year old male speaker. We attribute this to a list intonation effect, a relic of the highly emphatic speech a speaker employs as part of a wordlist reading task. We assume that the speaker attends more to producing a contour as the salient differentiating feature than to the absolute pitch under these experimental conditions. In actual, fluent speech, as reported above, we found both absolute pitch difference (low vs. modal) to be robust. [9] We consider the contour to be merely a by-product of the pitch event. In normal speech, low pitch is the salient contrastive feature.

The claim that low pitch alone provides the basis for contrast in normal speech is supported by the same female speaker's production of the two contrasting tokens (19) in a carrier sentence (21-22).

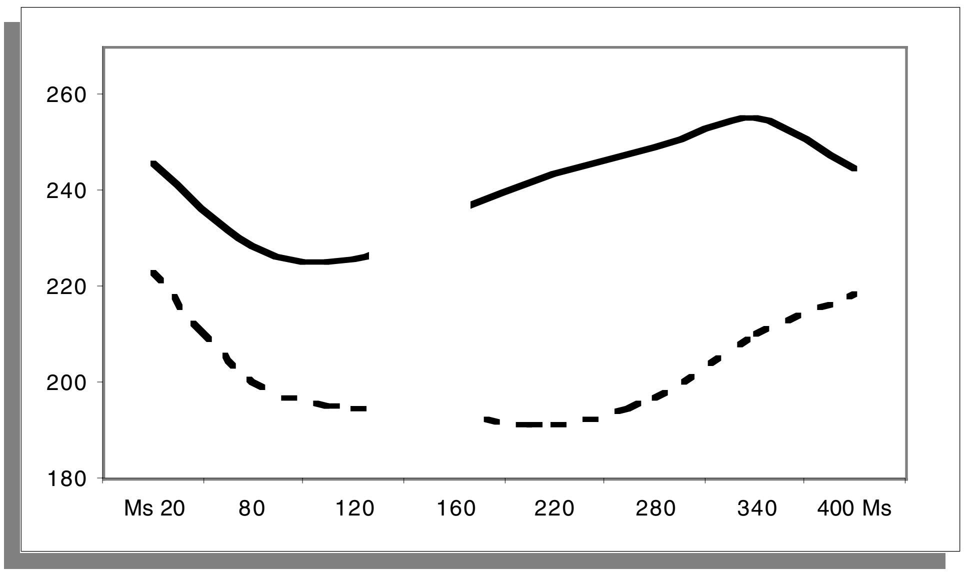
[9] We do not wish to exclude the possibility that some speakers (or even entire dialect regions) may re-analyze an absolute pitch difference as a pitch contour difference. Our pool of speakers does not include at present a diverse enough sample to tests this hypothesis.

(21) ooŋ **qadɯː** ekki dir be ?
s/he-GEN health-3 good=EVID QUES
Is his/her health good?

(22) ooŋ **qàdɯː** ekki dir be ?
s/he-GEN porridge-3 good=EVID QUES
(a.) Is his/her porridge good?

Table L shows the average pitch contours of six tokens of each word produced by the same 21-year old female speaker in the carrier sentence.

(23) Table L: Pitch contour of [qadɯː] (solid line) and [qàdɯː] (dotted line) produced in a sentence



In the twelve tokens graphed above, the duration of low pitch vowels is reduced significantly by the sentential context so that it nearly matches that of the short vowels. The long vowels in the second syllables are comparable in duration.

(24) Duration of vowels represented in Table L

	First syllable	Second syllable
[qàdɯː]	127 Ms	333 Ms
[qadɯː]	122 Ms	289 Ms

The significant finding here is that the low pitch plateau, which carries over into the second syllable, retains a stable duration of about 200 Ms. This reflected in Table L by the interval between 80Ms and 280Ms shown in the dotted line. We note that this duration is close to that of the pitch event in monosyllables, which is about 160 ms for a the male speaker we recorded (Table E).

5.0 Discussion

This chapter provides the first acoustic description of contrastive low pitch in Tuvan. As such, it attempts to lay the groundwork for future analyses, both phonetic and phonological. We have argued that the special 'pharyngealized' vowels contrast for low pitch alone, not for pitch contour, pharyngeal constriction or creaky (glottalized) phonation.

The essential distributional characteristic of low pitch is its positional neutralization. It appears in initial syllables only. We have as yet identified no clear historical or conditioning environment can be shown for the set of lexical items that bear low pitch. This will require a far more extensive historical and comparative study of related Turkic languages. We noted that new items can be added to the class, while different dialects of Tuvan are observed to augment or diminish the class.

We interpret low pitch as having a single articulatory target, with the subsequent pitch transition representing an updrift towards modal pitch. We have argued that low pitch appears to have an event-determined, intrinsic length which does not precisely match the duration of either long or short vowels in all contexts. As such, the low pitch plateau—which lasts about as long as a typical syllable—may carry over into the second syllable in order to achieve the necessary duration.

Supra-laryngeal gestures seem to be crucially coordinated with respect to the pitch event. For example, all monosyllabic tokens with low pitch happen to be closed syllables. They appear to critically coordinate the onset of the coda consonant gesture with respect to the low pitch gesture. The coda consonant gesture does not begin until the pitch gesture has achieved its target duration (Table E). Our preliminary data show that variability in pitch event duration across words with different coda segments is negligible. The consonant gesture does not appear to influence the duration of the pitch event. Rather, the consonant gesture appears to be controlled by the pitch gesture and not permitted to overlap with it.

Once it is no longer in coda position, the consonant need no longer be critically timed with respect to the pitch event. In disyllables, for example suffixed monosyllables, the consonant that was in coda position undergoes re-syllabification (25).

(25)	àt̲	‘horse’-1SG
	à.d̲ɯm	‘horse’-1SG

It is therefore no surprise that the consonant—now as an onset—succumbs to a typical pattern of gestural overlap which it seemed to resist in the monosyllabic context. In disyllables, the pitch event continues throughout the entire duration of the consonant closure and well beyond it into the second syllable. The pitch event appears to have no apparent effect on the duration of the closure (Tables I, L).

Although it is anchored in the initial syllable, low pitch will extend partially into the second syllable when one is available (Tables E, L). We found this effect to be robust, but to be strictly limited to the first part of the second syllable. It does not constitute an instance of ‘harmony’ by which a low pitch feature spreads in an unbounded manner across the word domain (contra Bicheldei 1980b). Again, this speaks to the putative intrinsic duration of the gesture, which we have proposed to be about 160 to 200 ms, or approximately the average length of a syllable.

After completing the research and analysis in this chapter, we obtained new field data from two peripheral dialects of Tuvan spoken in Mongolia and from the very closely related language Tuha, also spoken in Mongolia. The evidence suggests that the intrinsic timing hypothesis advanced herein may indeed be correct. In all three of these languages (or dialects), pitch *contour* is not the salient feature, and no extended duration is observed. All three lack both *contour* and *length* but employ some other distinctive quality. The Tuha language, which is nearly cognate and mutually comprehensible with Tuvan, employs *short* nasalized vowels where Tuvan employs pitch accent. In the examples below we use a length diacritic in the Standard Tuvan examples to emphasize that length is not present in the dialect examples.

(26) <u>Tuvan</u> <u>Tuha</u>
à:t a‡t ‘horse’-1SG

The Hovd and Tsengel dialects of Tuvan employ *short* low pitch vowels in these same environments.

(27) <u>Tuvan</u> <u>Tsengel and Hovd dialects</u>
à:t àt ‘horse’-1SG

We thus conclude that length (e.g. timing) may indeed be driven by the laryngeal event. When the nature of the laryngeal event is altered, timing follows suit.

Finally, we return to the question of whether the Tuvan case is properly construed a pitch accent system, a tone system, or something else. We tentatively interpret the data as resembling more closely a tone system than a pitch accent system. A true pitch accent, for example, the rising pitch accent of Serbian (formerly Serbo-Croatian, Gvozdanovic1980), might be expected to spread fully onto a second syllable. The Tuvan pitch event, by contrast, can spread precisely half-way onto the second syllable. Also, a contour pitch accent such as the gravis accent (falling-rising) of Swedish (Elert 1964) can be limited in its distribution to disyllabic words. The Tuvan pitch event, although clearly anchored in the first syllable, can occur in words of any syllable count. Such comparisons to other pitch accent systems are necessarily inconclusive and somewhat speculative. We expect that further acoustic and phonological research will shed light on the hypothesis that Tuvan has a tone-like pitch event, not a pitch accent.

Chapter Three: Hiatus resolution and velar deletion

0. Introduction

In this chapter we explore two related phenomena: hiatus resolution and velar deletion. Both processes reveal the effects of several constraints not otherwise visible in the grammar of Tuvan. These include constraints that enforce positional faithfulness, featural faithfulness, and recoverability. Hiatus resolution also reveals the operation of a low-ranking constraint that bans onsetless syllables thus ruling out vowel hiatus.

In section §1 of this chapter we explore vowel hiatus and the conditions that give rise to and resolve hiatus. Hiatus is typically resolved by coalescence that favors non-high vowels, exemplifying a kind of featural faithfulness. Vowels may also be protected by positional faithfulness, and these two types of faithfulness may come into conflict in hiatus resolution. Potential instances of hiatus are most frequently introduced by the deletion of intervocalic velars, such that the two phenomena are interrelated.

In section §2, we describe the unusual distribution and patterning of velar stops. Velars not only have a unique distribution, but are also (highly marked and) frequently subject to deletion. We describe the robust process of velar deletion and the seemingly diverse range of environments that can block deletion (§4). Blocking effects are driven by different types of faithfulness constraints; these all have the common effect of enhancing the recoverability of a distinct class of short stems and short suffixes (§5). We propose a unified account of velar deletion and varied blocking effects. In §6, we show that the interaction between velar deletion and hiatus may, under certain conditions, yield no optimal output. We consider whether this is a case of true *ungrammaticality*, in which unviolated constraints may force selection of a null parse as optimal. In §7, we apply two formal models of ungrammaticality to the Tuvan case.

1.0 Hiatus

Tuvan vowels are limited in their distribution by harmony and positional neutralization (cf. chapter 4 on Vowel Harmony). For example, back and front vowels never co-occur in a word, with the exception of a few non-alternating suffixal vowels (chapter 1). Further, low rounded vowels never appear post-initially, and a high unrounded vowel never appears after a rounded vowel. The sixteen possible sequences of two vowels are shown below. Shaded boxes represent impossible sequences.

(1) <u>Vowel co-occurrence patterns</u>

	2nd Vowel							
1st Vowel	i	ü	e	ö	ɨ	u	a	o
i	iCi		**iCe**					
ü		üCü	**üCe**					
e	**eCi**		eCe					
ö		**öCü**	öCe					
ɨ					ɨCɨ		**ìCa**	
u						uCu	**uCa**	
a					**aCì**		aCa	
o						**oCu**	oCa	

Six of the vowel pairs shown already agree in all features. Two pairs, [oa] and [öe], agree in all features except rounding. While these vowels may appear in successive syllables, they are unattested as potential hiatus pairs. (Consonant deletion that *would* cause [oa] and [öe] to become adjacent is blocked by markedness constraints banning low rounded vowels. We explore these in chapter 4.) The remaining eight potential hiatus pairs (boldface) differ solely in height; these are targeted by hiatus resolution.

1.1 Hiatus Resolution

Permitted sequences of two adjacent vowels may arise by affixation of a vowel-initial suffix to a vowel-final stem (2), by affixation of a vowel-initial suffix to a consonant-final stem, where a final consant is deleted (3), or by affixation of a consonant-initial suffix where the consonant undergoes deletion (4).

(2)	simple affixation			
	biži + ir [10]	→	bižiir	'write'-FUT
(3)	affixation and deletion			
	ug + u	→	uu	'direction'-3
(4)	ada + gan	→	adaan	'name'-PAST

In a sequence V_iV_j where the two vowels are identical, they simply coalesce into a long, tautosyllabic vowel (underlined) (5, 6).

(5)	H + H → HH			
	setki + ir	→	setkiir	'feel'-FUT
	öörü + ür	→	öörüür	'be joyful'-FUT
	xalɨ + ɨr	→	xalɨɨr	'run'-FUT
	udu + ur	→	uduur	'sleep'-FUT
(6)	L + L → LL			
	iyile + gen	→	iyileen	'go by two'-PAST
	kède + geš	→	kèdeeš	'sneak up on'-CV
	ada + gan	→	adaan	'shoot'-PAST
	ɨrla + gaš	→	ɨrlaaš	'sing'-CV

If the two adjacent vowels differ in height, they undergo coalescence by which the high vowel is assimilated fully to the non-high vowel regardless of their relative order (7, 8).

(7)	L + H → LL			
	kel + ir	→	keer	'come'-FUT
	ög + ü	→	öö	'yurt'-3
	dag + ɨ	→	daa	'mountain'-3
	čor + ur	→	čoor	'go'-FUT
	al + ɨr	→	aar	'take'-CV (Auxiliary)
(8)	H + L → LL			
	biži + geš	→	biželeš	'write'-CV
	öörü + geš	→	ööreeš	'be joyful'-CV
	udu + gan	→	udaan	'sleep'-PAST
	saktɨ + gan	→	saktaan	'remember'-PAST

The reverse pattern—assimilation of non-high vowels to high vowels—is unattested in Tuvan (and, as far as we are aware, in Turkic).

[10] In the data presented in this section, we adopt for expository purposes actual vowel symbols for suffix vowels. Elsewhere throughout this dissertation, we have used archiphonemic symbols for suffix vowels to emphasize the regular, predictable nature of their surface alternations. We note that not only suffix vowels, but *all* post-initial vowels in native words have fully predictable values for backness and rounding and are in some sense archiphonemic. In chapter 4, we will argue that post-initial vowels in harmonic words are, in fact, underspecified for backness and rounding.

1.2 Dominance of non-high vowels

The observed dominance of non-high vowels in assimilatory processes falls out from general phonological properties (Donegan 1985). These include the relative markedness of high vowels vs. non-high vowels, the greater sonority and perceptibility of non-high vowels, and the greater ability of non-high vowels to resist assimilation. Donegan (1985) observes that cross-linguistically, it is the more sonorant vowels that tend to dominate in many phonological processes. Surveying a broad range of synchronic and diachronic processes, she notes the following asymmetries between more sonorant and less sonorant vowels:

<u>Less sonorant vowels</u>:

- more likely to have non-syllabic alternants (glides)
- lose syllabicity earlier in historical change
- more often subject to deletion

<u>More sonorant vowels</u>:

- more likely to retain voicing and to spread voicing
- more likely to promote continuance in adjacent consonants
- more likely to retain syllabicity
- intrinsically longer

Donegan concludes that more sonorant vowels are more robust with respect to four key properties: voicing, syllabicity, continuance and sustainability. She notes that “the properties which are central to the nature of vowels seem to be present more strongly in the lower, more sonorant vowels.” Her conclusions implicitly predict that sonority can play a decisive role in various assimilatory processes that involve vowels.

1.2 A typology of hiatus resolution

Casali (1996) surveys vowel hiatus resolution in 92 languages to answer the question: "In hiatus resolution, which vowel goes?" He concludes that the most commonly attested patterns of coalescence (and deletion) are *positionally* determined; root vowels usually win out over affixal vowels. This observation finds support in the theory of positional faithfulness (Beckman 1997), which models the greater likelihood of root vowels to be protected by enhanced faithfulness.

When positional faithfulness plays no decisive role, the question of which vowel goes may be determined by another type of faithfulness. In what Casali (1996) terms 'symmetrical coalescence', sequences [e.i] and [i.e] would both yield [ee]. The relative ordering of segments plays no role. The question of which vowel stays is determined by the 'feature-sensitive' nature of this coalescence process. Vowels that possess a particular 'more highly valued' feature resist assimilation or deletion, while vowels that lack the feature undergo it. The valued feature, e.g., [-high], serves to elevate a vowel segment in a relative 'strength hierarchy' of vowels. Under symmetrical coalescence, a vowel segment is preserved in its entirety. This is attributed to a highly-ranked constraint on segment integrity (SEGINT). Thus, any vowel bearing the valued feature will be fully preserved and any vowel lacking it will be lost. A symmetrical coalescence pattern never blends the 'best' features from two adjacent vowels. Instead, it deletes one vowel and keeps the other, often lengthening the vowel to preserve mora count. The result, as Casali notes, is that "For any given pair of input vowels, the phonetic result will always be the stronger of the two vowels, regardless of their original input order."

Afar exemplifies this type of coalescence (Casali 1996:76-79). We note that although Afar has distinctive length, coalescence does not result in surface long vowels, as in Tuvan.

(9) Afar coalescence: L.H and H.L → L

kimmir**o-u**rte	→	kimmir**o**rte	'the bird got well'
urr**u-o**w	→	urr**o**w	'children'

Casali (1996:80) concludes that "Fully symmetric coalescence of the type found in Afar appears to be cross-linguistically rare." Among the 92 languages he surveys, only Afar, Yoruba[11] and modern Greek exhibit 'symmetrical, feature-sensitive coalescence'.

We have argued that symmetrical coalescence with dominance of non-high vowels is well-grounded in the general phonological properties of those vowels. Furthermore, it has been documented in numerous Turkic languages, e.g. Xakas (Anderson 1998), Gagauz

[11] Pulleyblank (1998)

(Pokrovskaya 1964), Dolgan (Ubryatova 1986), Chuvash (Kreuger 1961), and other Altaic languages, e.g. Buriat (Poppe 1967), Evenki (Nedjalkov 1999). Thus 'feature sensitive coalescence' based on the dominance of [-high] turns out to be the norm in Altaic and may be more common cross-linguistically than Casali's typology would suggest.

1.3 A formal model of hiatus and coalescence

We adopt an optimality theoretic (Prince and Smolensky 1993) model that reflects the attested (and expected) dominance of non-high vowels in symmetrical coalescence. The fact that many languages disprefer hiatus or ban it altogether is interpreted as reflecting a universal markedness constraint against onsetless syllables (McCarthy and Prince 1993).

ONSET *Syllables must have onsets.*

Tuvan bans onsetless syllables word-internally but allows them word-initially. Speakers are not observed to epenthesize a glottal stop as do, for example, German speakers. An additional contraint is thus needed to enforce alignment of the lexical (stem) material with the left edge of the prosodic word edge.

ALIGN (PRWD-L, STEM-L) *The edge of a prosodic word must coincide with phonological material that belongs to a stem*

Onsetless syllables that appear word-internally must either be filtered out or repaired. A typical repair for hiatus is vowel coalescence, and we will need a formal model of the dominance of non-high vowels in coalescence. Reviewing the work of Donegan (1985), we proposed that the observed dominance of non-high vowels falls out from general phonological properties. It is well-established, for example, that a more sonorous vowel makes for an intrinsically better syllable nucleus. Reference to vowel sonority is formalized as a universal well-formedness constraint (Prince & Smolensky 1993:16), called the Nuclear Harmony Constraint.

HNUC *A higher sonority nucleus is more harmonic than one of lower sonority.*

Given the better, or 'more harmonic' nature of more sonorous vowels, feature sensitive coalescence may be readily modeled with constraints that enforce enhanced faithfulness to more sonorous vowels. Casali (1996:80) adopts feature-sensitive parse constraints for this purpose.

PARSE(-HIGH) >> PARSE(+HIGH)

Pulleyblank (1998) describes a Yoruba pattern of vowel coalescence across word boundaries, in which the high vowel in a $V_i][V_j$ sequence assimilates fully to the non-high vowel. He adopts sonority-sensitive faithfulness constraints that render the loss of a more sonorous segment more costly than the loss of a less sonorous one. For our purposes, only a subset of Pulleyblank's constraints are needed:

MAXNONHI >> MAXHI

The full set of constraints we adopt for hiatus resolution is as follows:

ALIGN (PRWD-L, STEM-L)	*The edge of a prosodic word must coincide with phonological material that belongs to a stem*
FAITH.IO-V	*Vowels in the input must be faithfully rendered in output*
ONSET	*Syllables must have onsets.*
MAXNONHI	*Non-high vowels in the input must be represented in output*
MAXHI	*High vowels in the input must be represented in output*
DEP	*No segment lacking in the input may be present in the output*

The following sub-ranking ensures that onsetless syllables will surface in word-initial positions. They will not be repaired, for example, by epenthesizing an onset (candidate 10b).

ALIGN (PRWD-L, STEM-L) >> ONSET >> DEP

(10) Word-initial onsetless syllables will surface despite ONSET

	/#eC/	ALIGN (PRWD-L, STEM-L)	ONSET	DEP
☞	a. #eC		*	
	b. #CeC	*!		*

The following sub-ranking will repair word-internal onsetless syllables via coalescence that favors the non-high vowel.[12]

ONSET, MAXNONHI >> MAXHI

(11) Coalescence HL → LL

	/C i e/ (+hi -hi)	ONSET	MAXNONHI	MAXHI
	a. Ci.e	*!		
	b. Cii		* !	
☞	c. Cee			*

(12) Coalescence LH → LL

	/C e i/ (-hi +hi)	ONSET	MAXNONHI	MAXHI
	a. Ce.i	*!		
	b. Cii		*!	
☞	c. Cee			*

The Tuvan coalescence facts are significant in several respects. First, they provide a clear case of 'feature sensitive coalescence', where non-high vowels dominate regardless of their relative order. This pattern was previously thought to be typologically rare (Casali 1996). Second, Tuvan exemplifies a system in which post-initial stem vowels do not seem to enjoy any protected status vis-à-vis suffix vowels. This fact sets Tuvan apart from the majority of languages in Casali's typology, in which segment affiliation (e.g. root vs. affix) plays the decisive role. Thirdly, as we shall argue in §7, the coalescence patterns of one dialect provide evidence for the active status of two distinct constraints, ONSET and NOHIATUS.

[12] A hiatus sequence /Ci.e/ might also be repaired by turning the [i] into a glide, thus rendering it part of an onset cluster [Cje]. This is ruled out in Tuvan by a constraint banning complex onsets.

2.0 Distribution of Velars

Velars in Tuvan are highly marked, atypically distributed and regularly deleted. In this and the following sections, we explore the distribution of velars, deletion of velars, and the constraints that govern these. In section §2.2, we claim that the atypical distribution and deletion of /g/ can be attributed in part to its unusual moraic status. In section §2.3 we show that deletion of velars in intervocalic position applies robustly in Tuvan. However, in a range of environments, even though the conditions for deletion are present, deletion is blocked (§3). We will claim that blocking in such environments is driven by independent constraints on positional faithfulness to vowels (§4). These in turn are driven by the demands of lexemic/morphemic recoverability (§5). The coalescence patterns described in §1 will prove relevant here.

Velar stops in Tuvan pattern differently from other stops. They contrast for voicing where other stops do not; and they fail to contrast where other stops do. In onsets, alveolar (13) and bilabial (14) stops contrast for voicing, but velars do not (15).

(13)	bar	'go'-FUT
	par	'tiger'
(14)	dep	'say'-CV
	tey	'hill'
(15)	kas	'goose'
	*ga	

In codas, bilabial (16) and alveolar (17) stops do not contrast for voicing, but velars do (18).

(16)	ap	'take'-CV
	*ab	
(17)	at	'name'
	*ad	
(18)	uk	'lineage'
	ug	'direction'

Stops, when they are geminates, must be voiceless: [pp] [*bb] (19); [tt] [*dd] (20). Velars again prove exceptions to the rule: geminate velars do contrast for voicing [kk] [gg], (the voiced velar is generally fricated in this environment and surfaces as [ɣɣ])(21).

(19)	öppey	'baby'
	*abba	
(20)	attan	'name'-ABL
	*adda	
(21)	akkɨy	'elder brother'
	saɣɣan /saggan/	'milk'-PAST.I

Finally, there is one context in which velars pattern (at least partially) with other stops. All three stops appear intervocalically as singletons, but they must be voiced in this position (22-24). Bilabial and velar stops typically undergo lenition in this context, surfacing as [ʋ] and [ɣ] respectively (23, 24).

(22)	ada	'father'
	*ata	
(23)	aʋ al /ap al/	AUX-CV 'take'-CV
	*apa	
	*aba	
(24)	aɣaar /agaar/	'air'
	*aka	

We sum up the distribution of stops schematically below (25). Shaded areas indicate the distinct patterning of /g/ (26).

(25) <u>Voiceless stops</u>

	p	***t***	***k***
onset	+	+	+
coda	+	+	+
VCV	-	-	-
VCCV	+	+	+

(26) <u>Voiced stops</u>

	b	***d***	***g***
onset	+	+	-
coda	-	-	+
VCV	+ [13]	+	-
VCCV	-	-	+

Velars thus enjoy enhanced faithfulness to voicing when they appear in coda position or as geminates. Elsewhere, they undergo neutralization for voicing or deletion. Other stops enjoy enhanced faithfulness to voicing in onset position, and undergo

[13] The distinction between underlying [p] and [b] is neutralized in intervocalic environments, where both surface as the voiced labio-dental approximant [ʋ].

neutralization in codas and as geminates. Velars alone may contrast for voicing as geminates. Finally, velars undergo deletion as intervocalic singletons. The table below (27) depicts these contrasts schematically. The +/- symbols indicate presence/absence of voicing contrast. [14]

(27) Voicing contrasts of stops

velars are deleted here

	onset	coda	VCV	VCCV
Velar stops	-	+		+
Bilabial, alveolar stops	+	-	-	-

2.2 Syllable structure constraints

We have shown that velars, particularly /g/, pattern in a unique way in both their distribution and contrastiveness for voicing. This patterning is probably attributable in part to historical quirks of the language. In this section, we argue that the moraic status of /g/ also contributes to its unusual distribution. We propose that underlyingly voiced /g/ always projects a mora in Tuvan. This has several implications. First, it sets /g/ apart from other segments that are potentially moraic by virtue of voicing and position. Second, the moraic status of /g/ may be seen as licensing its unusual voicing contrast in codas. Third, the absence of /g/ in onsets, and its tendency to be deleted may be in part a consequence of constraints on syllable weight and moraicity.

In §2.1 we noted that /g/ is the only segment contrastively specified for [+voice] that may appear in a coda. Other than /g/, the only segments that appear as codas are sonorants [l, m, n, r, ŋ], all redundantly specified [+voice].

(28) mal 'livestock'
xar 'snow'
san 'number'

[14] It seems paradoxical that velars can be both relatively more marked (e.g. subject to deletion), yet at the same time enjoy a seemingly greater range of contrast than other stops. While we limit our analysis here to a formal, OT account of the facts, we suspect that functional, phonetic properties of velars which have not yet been investigated might yield a more satisfying explanation. For example, work by Kirchner (1998) suggests that the behavior of geminates—a difficult problem to model formally (see Hayes 1986)—is perhaps better understood in functional, articulatory terms.

xam	'shaman'
saŋ	'incense'

Evidence for the moraic status of contrastively voiced /g/ in codas is that it cannot appear in coda position when the syllable contains a long vowel. [15]

(29) * aag

We attribute this to a cross-linguistically motivated constraint banning super-heavy (tri-moraic) syllables. Redundantly voiced coda sonorants [l], [r], [n], [m] [ŋ], by contrast, appear freely as codas following long vowels. Hence, these must be non-moraic.

(30)	mool	'Mongolian'
	suur	'village'
	čaan	'elephant'
	ööm	'yurt'-1
	ööŋ	'yurt'-2

Contrastively voiced /g/ is also limited in onset position. We note that most suffix-initial velars have fully predictable, not contrastive, voicing. Velar onsets of suffixes are voiceless after voiceless segments (31), voiced after voiced ones (32), and usually deleted after vowels (33, cf. §3.0). We represent this predictably voiced/voiceless velar onset as archiphoneme /G/ below. We assume it to be unspecified for voicing.

(31)	at + Gan → at-kan	'shoot'-PAST
(32)	al + Gan → al-gan	'take'-PAST
(33)	ada + Gan → ada-an	'name'-PAST

Non-archiphonemic, contrastively voiced /g/ appears only in non-derived environments, where it may occupy both coda and onset position. However, cases of /g/ in onset position seem to be attested mostly in non-native loanwords of Mongolian origin (M).

(34)	i.gil	'horsehead fiddle' (M)
	a.ra.ga	'fermented milk alcohol' (M)
	ug	'direction'

[15] This is minimally different from Xakas, where both /k/ and /g/ surface as codas following long vowels: *soox* 'cold', *köög* 'music' (Greg Anderson, p.c.). Xakas thus differs minimally from Tuvan in that it either allows super-heavy syllables or or it does not count coda /g/ as moraic.

ug.ba	'lift'-NEG
ug.bay	'elder sister'
ɨg.laar	'cry'-FUT

Let us assume that contrastively voiced /g/ must project a mora in Tuvan. Given this assumption, several distributional facts about /g/ become clear. First, /g/ never appears as a word-initial onset. This reflects historical processes of the language that neutralized the voicing contrast for onset velars, thus precluding a moraic onset /g/. Second, /g/ appears as an onset word-internally only in non-derived environments, typically in loanwords of Mongolian origin, e.g. *i.gil, a.ra.ga* (34). We presume high-ranking faithfulness to a stem can cause onset /g/ to surface in such environments. Thirdly, in native lexemes, /g/ appears only in coda position, e.g. *ug.bay, ɨg.laar*. Fourthly, /g/ appears as part of a geminate /gg/ only when the first half of the geminate belongs to a root and the latter half to a suffix.

(35) sag + Gan → saɣɣan 'milk'-PAST

In this heteromorphemic environment, we propose that only the coda /g/ is contrastively voiced and moraic, while the onset is predictably voiced and non-moraic. Finally, when a contrastively voiced coda /g/ is resyllabified to become an onset, it undergoes deletion §3.0. Deletion can be better understood in light of the highly marked or even universally impossible status of a moraic element in onset position. Thus, any /g/ that resyllabifies from coda to onset position must either be deleted or must be rendered non-moraic in the surface structure. The data indicate that the constraint requiring that /g/ be moraic is more highly ranked than the I-O faithfulness constraint that calls for preservation of an input /g/.

To sum up, we have argued that contrastively voiced /g/ projects a mora. As a result, it is subject to a syllable weight constraints. It may surface as an onset only in non-derived environments due to high ranked root faithfulness. In §3.0 we will invoke the proposed moraicity of /g/ to account for a different deletion pattern of coda /g/ vs. coda /k/ in monosyllabic words.

3.0 Velar deletion

Velar deletion interacts in a complex way with the vowel coalescence patterns we have described herein (§1). This interaction sheds a unique light on the role of various constraints in the grammar. For example, deletion reveals the crucial difference between NOHIATUS and ONSET. Velar deletion also serves to reveal the effects of positional faithfulness constraints that are not otherwise apparent in the grammar. In §6 we explore potential ungrammaticality that arises as a consequence of velar deletion.

Deletion of intervocalic velars (and uvulars) is a widely attested process across Turkic languages, including the well-known case of Turkish (Sezer 1981).

(36)	<u>Turkish</u>	bebe**k** + i	→	bebe.i (*bebe**k**i)	'baby-3'
	<u>Gagauz</u>[16]	mopra**k** + ɨ	→	mopraa (*mopra**k**ɨ)	'bowl'-3
	<u>Kazak</u>[17]	qara**q** + ɨ	→	qaraa (*qara**q**ɨ)	'eye-3'
	<u>Tuvan</u>	ča© + ɨ	→	čaa (*ča©ɨ)	'fat'-3

A similar process is attested in unrelated languages, e.g. Faroese (Anderson 1992:247, H. Petersen, p.c.), and Kasem (Howard 1970).

Tuvan, along with closely related languages Tofa and Xakas, shows complex patterns of velar deletion in intervocalic environments. Deletion targets velars in both codas and onsets. Coda velars are found in nominal and verbal stems, both mono- and poly-syllabic. Onset velars are found in nine productive suffixes, both derivational and inflectional (see Appendix A for full listing). Deletion is blocked in a non-uniform class of environments where the structural conditions for it appear to be present. We show that deletion is constrained by various independently motivated effects that all relate to morphemic and lexemic recoverability.

The basic deletion pattern is seen in monosyllabic words. Coda velars are deleted when a vowel initial suffix renders them both intervocalic and in onset position (37). In the examples below, we use the 3Sg possessive suffix, shown archiphonemically as /-I/. Where

[16] Pokrovskaya 1964.
[17] Farrukh Suvankulov, p.c.

velar deletion creates potential vowel hiatus (37 c-d), this is resolved in favor of the non-high vowel (cf. §1.1).

		noun		noun + /-I/ (Possessive)	
(37)	a.	ug + u	→	uu (*uɣu)	‘direction’-3
	b.	sug + u	→	suu	‘water’-3
	c.	ög + u	→	öö	‘yurt’-3
	d.	čag + ɨ	→	čaa	‘fat’-3

Deletion of coda /k/ is blocked in this environment (38). This is an example of recoverability-driven blocking that will be examined further in §4.0.

(38)	uk + u	→	uɣu (*uu)	‘direction’-3
	ak + ɨ	→	aɣɨ (*aa, *ɨɨ, *aɨ)	‘white’-3

In polysylabic words, coda /k/ (39) and /g/ (40) pattern together; both undergo deletion.

(39)	idik + i	→	idii	‘boot(s)’-3
	inek + i	→	inee	‘cow’-3
	čuruq + u	→	čuruu	‘picture'-3
	ayaq + ɨ	→	ayaa	‘bowl’-3
(40)	bilig + i	→	bilii	‘knowledge’-3
	töleg + i	→	tölee	‘payment’-3
	urug + u	→	uruu	‘daughter’-3
	čadag + ɨ	→	čadaa	‘by foot’-3

Tuvan has nine velar–initial suffixes. These fall into two classes according to size. The ‘short’ suffixes have the form /-GV/. These generally block velar deletion and will be discussed further in §4.0. The ‘long’ suffixes have the form /-GVC/ or /-GVCV(C)/ (41). They uniformly allow velar deletion. Archiphonemic /G/ indicates that surface voicing is fully predictable for these segments.

(41) Long suffixes undergo velar deletion

/GAn/	Past
/GAš/	Converb
/GAy/	Conciliatory
/GAlA/	Converb
/GAlAk/	Unaccomplished

For long suffixs, we first show cases where velars are retained following a consonant, followed by cases where velars are deleted following a vowel.

(42)	Past Suffix /GAn/			
	bižit + ken	→	bižitken	'write'-CAUS-PAST
	at + kan	→	atkan	'shoot'-CAUS -PAST
	kel + gen	→	kelgen	'come'- PART
	biži + gen	→	bižeen	'write'-PAST
	ada + gan	→	adaan	'name'-PAST
(43)	biži + geš	→	bižeeš	'write'-PART
	at + kaš	→	atkaš	'shoot'-PART
	ègele + geš	→	ègeleeš	'begin'-PART
(44)	biži + geš	→	bižeeš	'write'-PART
	ada + gaš	→	adaaš	'name'-PART
(45)	Converb Suffix /-GAlA/			
	čorut + kala	→	čorutkala	'go'-PERF-CV
	ažɨlda + gala	→	ažɨldaala	'work'-PERF-CV
(46)	Unaccomplished Suffix /-GAlAk/			
	ün + gelek	→	ün-gelek	'come.out'-UNC
	ègeleve + gelek	→	ègele-ve-elek	'begin'-NEG-UNC

Stem-final velars, unlike affixal velars, are contrastively voiced. Their voicing is apparent in that they determine the voicing specification of suffixal segments that follow them. Stem-final velars do not undergo deletion when followed by a consonant-initial suffix (47, column 2).

(47)	Negative Suffix /-BA/		
	sag + Ba	sa**g**-ba (*saa.ba)	'milk'-NEG
	deg + Be	de**g**-be	'say'-NEG
	kag + Ba	ka**g**-ba	'leave'-NEG
	čɨɣ + Ba	čɨ**g**-ba	'gather'-NEG

Stem-final velars undergo deletion when suffixation renders them intervocalic and in onset position (48, column 2).

	Future Suffix /-Ir/		
(48)	sag + ɨr	saar (*sa.gar)	'milk'-FUT
	deg + ir	deer	'say'-FUT
	kag + ɨr	kaar	'leave'-FUT
	čɨɣ + ɨr	čɨɨr	'gather'-FUT

To sum up, velars in both stem-final and suffix-initial position are generally subject to deletion in derived, intervocalic environments. This applies to both singleton and geminate velars. We attribute velar deletion to three general mechanisms in the grammar. First, we have argued that that coda /g/ projects a mora. Constraints on syllable weight, we propose, limit the distribution of /g/ and perhaps trigger deletion. We have also assumed a general markedness constraint against intervocalic velar stops. A third factor, as we will argue in §5, is that the likelihood that markedness constraints will force deletion of /g/ is directly related to the size (and therefore, the inherent recoverability) of the stem or suffix to which /g/ belongs. We note that suffix-initial /g/ that regularly undergoes deletion belongs to 'long' suffixes (41), which have the form CVC(VC). These contrast crucially with the /g/ in 'short' suffixes, which can block deletion. An analogous division into 'short' and 'long' classes can be shown for stems that end in velars.

4.0 Blocking environments

Velars appear to enjoy immunity from deletion in various environments, both derived and non-derived. These blocking effects reveal clearly the activity of constraints whose role in the grammar is elsewhere obscured. These include positional and featural faithfulness, onset, and recoverability constraints. Each of these can be shown, in a small set of phonological environments, to outrank the velar markedness constraint and thus block velar deletion.

Velars are immune from deletion when they appear root-internally, in non-derived environments.

(49)	araga	'alcohol'
	čugaala-ar	'speak'-FUT
	igil	'horsehead fiddle'
	agaar	'air'
	è̀ge	'beginning'
	ègeleer	'begin'-FUT

There is a large literature devoted to non-derived environment blocking (NDEB) effects. For a recent analysis of NDEB as applied to Turkish velar deletion, we refer the reader to Inkelas (1998).[18] For our purposes, we simply assume that velars are preserved in such

[18] In Turkish, all deleted velars are underlyingly in coda position (*salak* → *sala-ɨm* 'stupid'-1sg.sbj). Turkish has no suffix-initial velars. With suffixation, word-final Turkish velars are resyllabified to become

environments to satisfy high-ranked constraints on root faithfulness. We merely mention these cases in passing, and do not offer any further analysis of NDEB effects herein.

We focus instead on blocking of velar deletion in derived environments. Blocking gives rise to apparent underapplication of velar deletion in environments like those in examples (50-53). Blocking usually involves either short /-gV/ affixes, or short /(C)Vg-/ roots. In the case of roots, deletion is blocked only when it would expose the initial vowel of a stem to assimilation through hiatus resolution (§1.3). We show minimally different patterns of blocking in Tuvan and two cognate languages, Tofa (Rassadin 1983) and Xakas (Anderson 1998). We will argue that all these blocking effects can be attributed to the demands of morphemic/lexemic recoverability.

In the Tuvan dative /-GA/, a short affix that begins with a velar, velar deletion is blocked. [19]

(50)	kiži-ge	‘person’-Dat’
	duza-ga	‘help’-Dat’
	araɣa-ga	‘alcohol’-Dat’
	ača-ga	‘father’-Dat’
	bɨzaa-ga	‘calf’-Dat’
	börü-ge	'wolf'-Dat’
	ivi-ge	‘deer’-Dat’

In Tofa (Rassadin 1993) and Xakas (Anderson 1998), by contrast, the cognate dative affix never constitutes a blocking environment. [20]

---

onsets. They then generally undergo deletion, in contrast to root-internal velars which do not. One proposed way to explain NDEB effects in Turkish is to locally conjoin *VgV with an anchoring constraint, such that only resyllabified velars undergo deletion (Lubowicz 1998, Inkelas 1998) Clearly, local constraint conjunction account offers no explanation of the Tuvan velars, which are often deleted without undergoing any resyllabification.

[19] There are some dialect exceptions to this rule, and the so-called ‘standard’ literary language encompasses both speech varieties that delete the dative velar onset and ones that do not.

<u>Two ‘standard’ variants</u>	
kiži.ge ~ kižee	‘person’-DAT’
duza.ga ~ duzaa	‘help’-DAT’

Dialect variation notwithstanding, the general pattern is that a given speaker will either uniformly block velar deletion for the dative affix or uniformly enforce it. We analyze herein those varieties that always block deletion for the dative marker.

[20] If the Dative affix follows a long vowel, deletion is blocked, thus preventing an ill-formed tri-moraic syllable: <u>Tofa</u> *čaa-ga* (**čaaa*)‘war’-DAT; <u>Xakas</u> xazaa-ɣa (*xazaaa)‘stockyard’-DAT

(51) Tuvan
xeme + Ge → xemeɡe (~xemee) 'boat-DAT'

(52) Xakas
kime + Ge → kimee (*kimeɣe) 'boat-Dat'

(53) Tofa
xeme + Ge → xemee (*xemeɣe) 'boat'-DAT

Similarly, a short derivational affix /-gI/ always blocks deletion in Tuvan (54) but allows it in Tofa (55)

(54) Standard Tuvan: no velar deletion
Kɨzɨl-da-ɣɨ 'the one in Kyzyl'
üstün-de-©i 'the one on top'
amdɨ-©ɨ [21] 'now'-ADJ

(55) Tofa: velar deletion
amdɨ + ©ɨ → amdɨɨ 'now'-ADJ

Short suffixes thus pattern differently than long suffixes with respect to velar deletion.[22]

(56) Short suffixes block velar deletion

/-GA/	dative case
/-GI=že/	limited mood [23]
/-GI/	adjective
/-GI deg/	modal

We note that long suffixes, which allow initial velar deletion, all have the form CVC or larger and contain a non-high vowel. Short suffixes are all CV, and often contain a high vowel. We interpret the different patterning of short affixes as indicative of differing thresholds of morphemic recoverability. This idea will be developed in §5.0.

Velar-final stems can also show blocking of velar deletion. As noted, (C)Vg- verb stems permit velar deletion if they contain a non-high vowel, regardless of the sonority of the affixal vowel. Given the fact that non-high vowels always win out in hiatus resolution, a non-high stem vowel is never threatened by assimilation.

---

[21] A dialect variant is [amdɨɨ], with velar deletion.

[22] A small class of unproductive derivational suffixes also blocks deletion, even though these suffixes are not "short": e.g. Adjectival /-GIr/, and Adjectival /-GAn/. We consider these to be fully lexicalized forms that do not undermine the basic short vs. long distinction we propose here.

[23] Historically, probably a compound of /GI/ + /že/ (allative marker).

(57) verb stem

sag + gan	→ saan [24]	'milk'-PAST
sag + ɨr	→ saar	'milk'-FUT
deg + gen	→ deen	'touch'-PAST
deg + ɨr	→ deer	'touch'-FUT

A CV(g) stem also allows velar deletion if both stem and affix contain a high vowel, that is, the stem and affix vowels agree in height. Again, in such cases, the stem vowel is not threatened by assimilation.

(58) verb stem

ug + ur	→ uur	'lift'-FUT
ug + up	→ uup	'lift'-CV
čɨg + ɨr	→ čɨɨr	'gather'-FUT
čɨg + ɨp	→ čɨɨp	'gather'-CV

However, if the stem contains a high vowel and the affix contains a non-high vowel, velar deletion is blocked. These are cases that would give rise to hiatus, which would in turn result in assimilation of the high (stem) vowel to the non-high (affix) vowel. The stem would lose its sole vowel, rendering it unrecoverable.

(59) verb stem

ug + gan	→ uggan	'lift'-PAST
ug + gaš	→ uggaš	'lift'-CV
čɨg + gan	→ čɨggan	'gather'-PAST
čɨg + gaš	→ čɨggaš	'gather'-CV

Deletion of an affixal velar is thus blocked if it follows a CV stem where V is [+high]. For this subset of cases, velar deletion, if it were allowed, would cause the only vowel of the stem to be assimilated via hiatus resolution. Deletion is therefore blocked to protect the stem vowel (59).

Short verb stems with high-vowel suffixes allow velar deletion under two circumstances: (i) stem and affix vowel *agree* in sonority (60), or (ii) stem vowel is non-high (61)

[24] Subject to dialect variation: saaš ~ saɣaš ~ saɣɣaš

	CV stem	Future /-Ir/	Converb /-Ip/	
(60)	či-	čiir	čip	‘eat’
(61)	de-	deer	dep	‘say’

Short verb stems with a non-high vowel suffix pattern differently, depending on the quality of the stem vowel. A short verb stem containing a high vowel blocks deletion (62), while a non-high vowel verb stem allows deletion (63).

	CV stem			
(62)	či + gen	→	či**g**en (*čeen)	‘eat’-PAST
	či + **g**eš	→	či**g**eš (*čeeš)	‘eat’-PART
(63)	de + gen	→	deen	‘say’-PAST
	de + **g**eš	→	deeš	‘say’-PART

We now turn to the closely related language Tofa for a minimally different example. Tofa, like Tuvan, allows deletion of stem-final velars in (C)Vg stems when the stem vowel is non-high (64) or if the stem and affix vowels are both high (65):

(64)	Tofa			
	eg + ir	→	eer	‘crush’-P/F
	čžag + ɨr	→	čžaar	‘to precipitate’ (rain)
(65)	ug + ur	→	uur	‘lift’-P/F
	čžɨg + ɨr	→	čžɨɨr	‘gather’-P/F

Unlike Tuvan, Tofa, also allows velar deletion when the stem vowel is high and the suffix vowel is non-high (66). The initial syllable vowel of the stem subsequently undergoes assimilation.

(66)	či + gan	→	čeen (*čigen)	‘eat’-PAST
	či + gaš	→	čeeš (*čigeš)	‘eat’-CV

In the /-A(y)/ converb, Tuvan and Tofa again show minimally different patterning. Both allow deletion where a CVg- stem contains a non-high vowel (67). Tuvan blocks deletion where the (C)Vg- stem contains a high vowel (68, column 2), while Tofa allows deletion here (68, column 3).

	Tuvan & Tofa		Tuvan	Tofa	
(67)	čag + ay	→	ča(a)y	čaa(y)	‘rain’
	deg + ey	→	de(e)y	dee(y)	‘touch’
(68)	čɨg + ay	→	čɨ**g**ay	čžɨ(ɨ)(y)	‘gather’

We thus have two distinct environments in which blocking may occur. One type of blocking protects the sole consonant of short suffix; the other the sole vowel of a short stem. In §5.0, we propose that these seemingly disparate blocking effects are really quite similar and related to recoverability.

One additional blocking environment may be found in monosyllabic words. In contrast to coda /g/ (cf. 37), coda /k/ is never deleted in suffixed monosyllables (cf. 38).

	noun + 3			
(69)	tuk + u	→	tuɣu (*tuu)	'flag'-3
	ök + ü	→	öɣü (*öö)	'button'-3
	šak + ɨ	→	šaɣɨ (*šaa)	'time'-3
(70)	seek + i	→	seeɣi (*seee)	'insect'-3
	sook + u	→	sooɣu (*sooo)	'cold'-3
	öök + ü	→	ööɣü (*ööö)	'glottis'-3

We note that the cases in (70) can be explained by the ban on tri-moraic syllables. In the cases in (69) though, the failure of /k/ to undergo deletion must be driven by some other constraints. We suggest that the loss of coda /k/ in monosyllables would render them unrecoverable. The threshold of recoverability thus protects coda /k/ (but not coda /g/) in Tuvan and Xakas monosyllables. This differs minimally from Turkish, where both coda /k/ and /g/ in monosyllables are protected from deletion (Inkelas 1998):

(71)	Turkish monosyllables			
	ek + i	→	eki (*e.i)	'addition'-3
	lig + i	→	ligi (*lii)	'league'-3
	Turkish disyllables			
	ekmek + i	→	ekme.i (*ekmeki)	'bread'-3

Inkelas attributes blocking of velar deletion here to a disyllabic minimality constraint for Turkish. No morphologically complex word of Turkish may be monosyllabic, thus the expected form [e.i] 'his bread' would be ill-formed. Tuvan has no such constraint, as illustrated by forms like *öö* 'his yurt'. Rather, deletion of final voiceless velars in monosyllables is disallowed to satisfy recoverability, which dictates that underlying coda [k] be retained for short stems. Underlying coda [g] will be deleted perhaps because it is less essential from the point of view of recoverability, and also because it is more highly marked as a moraic onset (§2.2). We discuss the effects of recoverability further in §5.0.

4.0 A constraint-based model of deletion and blocking

We propose a straightforward application of Optimality Theory to the facts at hand. An important goal is to unify the entire range of velar deletion and blocking facts under a uniform set of constraints. We assume that deletion of all velars in intervocalic position is the normal state of affairs, governed by a general markedness constraint that may be observed in varying degrees in most Turkic languages. The full set of markedness and faithfulness constraints we adopt is as follows:

A.	*VGV	*No intervocalic velars* [25]
B.	*VGGV	*No intervocalic geminate velars (some dialects)*
C.	F.V1	*Faithfulness to 1st syllable vowel* [26]
D.	ONSET	*Syllables must have onsets*
E.	MAXNONHI	*Faithfulness to non-high vowels* [27]
F.	ALIGN (PRWD-L, STEM-L)	*The edge of a prosodic word must coincide with phonological material that belongs to a stem*

The factorial typology for constraints A through E (constraint F will be applied later) yields four minimally different rankings of the proposed faithfulness and markedness constraints. In each of the three rankings, one constraint is crucially dominated (boldface), while the relative ranking of the remaining three has no effect on the output. Each ranking selects an attested output that is unique to the language (or dialect) indicated in the rightmost column.

(72)

	Possible rankings			Output for /či + gen/	Dialect, language where attested
a.	*VGV, ONSET, MAXNONHI	>>	**F.V1**	čeen	C, Tofa, Xakas
b.	ONSET, F.V1, MAXNONHI	>>	***VGV**	či.gen	A
c.	*VGV, F.V1, MAXNONHI	>>	**ONSET**	či.en	B
d.	*VGV, F.V1, ONSET	>>	**MAXNONHI**	čiin	UNATTESTED

[25] See Orgun 1996, Inkelas 1998.
[26] For effects related to positional faithfulness, see Steriade 1993, Beckman 1998.
[27] On sonority sensitive faithfulness constraints such as MAXNONHI, see Pulleyblank 1998.

Three of the four forms predicted by the factorial typology are, in fact, attested. This suggests that we have identified the right set of constraints. In the tableaus below, we show outputs for the three attested rankings. Standard Tuvan places *VGV at the bottom: intervocalic velars will thus surface in a candidate that satisfies the other three constraints (73d).

(73) Standard Tuvan / č i + Gen / →[či.gen]

	/ č i + Gen / +hi -hi	MAXNONHI	F.V1	ONSET	*VGV
	a. čiin	*!			
	b. či.en			*!	
	c. čeen		*!		
☞	d. **cËgen**				*

Xakas by contrast places F.V1 at the bottom and will sacrifice the sole stem vowel to avoid hiatus.

(74) Xakas / č i + gen / →[čeen]

	/ č i + Gen / +hi -hi	MAXNONHI	*VGV	ONSET	F.V1
	a. čiin	*!			
	b. či.en			*!	
☞	c. **cËen**				*
	d. či.gen		*!		

Lastly, the Süt-xöl [28] dialect subordinates ONSET, and will select a candidate with hiatus. For now, we shall assume that candidate (75b) is actually the optimal output. In §7, we show that the *actual* output for this dialect is a Null Parse. The winning candidate shown here, (b.) [či.en], is the output speakers *would choose if there were an output*. We shall refer to this as the 'pseudo-optimal' output.

[28] Süt-xöl means 'Milk Lake', a region in western Tuva where a distinctive dialect is spoken. Much of the data in chapters 2 and 5 and elsewhere in this dissertation was contributed by speakers from this region.

(75) Süt-xöl dialect / č i + gen / →[či.en]

/ č i + Gen / +hi -hi	MaxNonHi	*VGV	F.V1	Onset
a. čiin	*!			
☞ b. cËen				*
c. čeen			*!	
d. či.gen		*!		

No dialect is observed to subordinate MaxNonHi, yielding an output such as (a.) [čiin]. Given the observed meta-ranking of MaxNonHi >> MaxHi across Altaic languages (cf. §1), such a ranking, though theoretically possible, would be unexpected in the typology of Altaic.

The preceding three tableaus demonstrate that deletion and blocking effects differ minimally across Tuvan dialects and closely-related, cognate languages. Each language/dialect may establish a different threshold of recoverability by ranking F.V1 higher or lower. In consequence of this threshold, short and long suffixes will pattern differently with respect to deletion. A full inventory of the thirty-nine morphologically cognate environments that allow or block velar deletion in Tuvan, Tofa and Xakas is given in Appendix A at the end of this chapter. Here, we illustrate minimally different deletion/blocking effects for one long suffix, /GAn/, and one short suffix, /GA/. Shaded boxes (76) indicate those contexts where velar deletion is blocked.

(76) Tuvan Dialects

	<u>long suffix</u> či + gen 'eat-Past'	***<u>short suffix</u>*** xeme+ ge 'boat-Dat'
A Standard	či.gen	xeme.ge
B Süt-xöl	či.en	xeme.ge
C Toža	čeen	xemee

(77) <u>Related languages</u>

Tofa (Rassadin 1983)	čeen	xemee
Xakas (Anderson 1998)	čeen	kimee

5.0 Recoverability

The data in §4.0 reveal two basic scenarios for blocking of velar deletion:

- when the sole vowel of a <u>short</u> stem would be assimilated .
- when the sole consonant of a <u>short</u> suffix would be deleted.

We propose that recoverability is the common motivating factor that unifies these two seemingly different blocking effects. Dialects of Tuvan, as well as cognate languages, exhibit varying thresholds of recoverability. Specifically, recoverability sets a threshold of tolerance for small stems or affixes. Differing thresholds are apparent with the patterning of short suffixes and short stems in Standard Tuvan on the one hand vs. Tofa and Xakas on the other. Short stems and short suffixes thus serve as indices of the threshold of recoverability. They either allow or block deletion, depending on the threshold setting in a particular dialect/language. Long stems and suffixes uniformly allow velar deletion in Tofa, Xakas and all Tuvan dialects, never blocking it. We have shown that grammars must be able to effectively protect small elements of the input from assimilation by demoting markedness constraints or promote faithfulness constraints. In other words, constraint rankings that enhance recoverability can effectively block an otherwise robust phonological processes (e.g. velar deletion) to enhance recoverability of small stems and suffixes.

The patterning of short suffixes such as the Dative marker /-GA/ allows us to establish empirically an exact 'threshold' of recoverability of the short stems and suffixes in these languages. In support of this, we also cite several diachronic and synchronic facts that point towards the functional role such a threshold plays in the grammar. We have suggested that the dative marker serves as a kind of barometer that registers the ongoing conflict between the speaker's need to delete a velar and the hearer's need to successfully recover a small morphological element. Diachronic changes in the dative marker in Tuvan and Xakas provide further evidence that thresholds of recoverability—to which the Dative marker is so sensitive—may differ between closely-related language and also within individual languages over time. Anderson (1998) reports that in Xakas, speakers of some (non-standard) dialects employ 'double marking' of the dative case.

(78) Standard Tuvan: Single marking with **no** velar deletion

kiži + ge → kiži.ge 'person'-DAT

(79) Standard Xakas: Single marking with velar deletion

kiži + ge → kižee 'person'-DAT

(80) Dialect Xakas: Double marking (once with velar deletion, once without)

kiži + ge → kižee.ge 'person'-DAT

We suggest that 'double marking' phenomena represent speakers' reanalysis of the dative marker as having a non-deleting velar onset. In cases where deletion applies, the only recoverable remnant of the suffix is vowel length. For some speakers, this has apparently become unrecoverable. They may come to view such forms as having no suffix at all. They thus add the seemingly redundant full dative suffix /-ga/ to such stems.

Historical evidence suggests Tuvan passed through a similar stage (cf. 79) in which speakers deleted the onset /g/ of the Dative marker. Such 'archaic' dative suffixes may be found in oral folklore, in forms such as /xovu + ga/ → *xovaa* 'steppe'-DAT (Darɨmaa 1976). In contemporary standard Tuvan, the velar onset of the dative marker is generally protected from deletion: /xovu + ga/ → *xovuga* (**xovaa*) 'steppe'-DAT, although some dialect variation persists.

We maintain that grammars are formal representations of communicative needs. The need for lexeme or morpheme recoverability can thus be encoded directly in grammatical statements. Recoverability is simply another word for faithfulness to input, and requires no additional theoretical apparatus. We have shown that recoverability can be enhanced by satisfying a faithfulness constraint (e.g. F.V1), or by allowing violations of a markedness constraint (e.g. *VGV). The data presented herein speak to the active role of faithfulness constraints in promoting 'recoverability' in the grammar.

6.0 Avoidance, Null Parse and Ungrammaticality

The interaction of velar deletion and hiatus resolution give rise to an apparent case of ungrammaticality in the Süt-xöl dialect of Tuvan. Ungrammaticality manifests itself as a paradigm gap, in which a form that is logically possible within a given paradigm does not surface. Instead, speakers select as optimal a form from another paradigm that does not exactly match the expected form in semantic content.

Before turning to the facts of Süt-xöl, we review the possible outputs yielded by the factorial typology.

(81) Possible rankings

	Possible rankings	Output for /či + gen/	Dialect, Language where attested
a.	*VGV, ONSET, MAXNONHI >> **F.V1**	če.en	C Tofa, Xakas
b.	ONSET, F.V1, MAXNONHI >> ***VGV**	či.gen	A
c.	*VGV, F.V1, MAXNONHI >> **ONSET**	či.en	B
d.	*VGV, F.V1, ONSET >> **MAXNONHI**	čiin	unattested

In the Süt-xöl dialect, represented by the ranking in (81c), intervocalic /g/ must always be deleted in derived environments where recoverability is not at stake (e.g. the past suffix /-GAn/). Yet velar deletion in this context gives rise to a situation in which speakers avoid the output dictated by the putative constraint ranking. In many hours of field observation and elicitation, we never found Süt-xöl speakers to spontaneously produce instances of hiatus as in *či.en* 'eat'-PAST.

Rather than produce forms containing an internal onsetless syllable, speakers employed morphological avoidance. They did so by skipping to an alternative paradigm that is semantically nearly (though not exactly) synonymous. Three related past tense paradigms provide the context for paradigm skipping.

(82) <u>Three past tense verbal paradigms</u>

DIALECTS	I. VERB-PAST	II. VERB-CV-PAST	III. VERB-CV AUX-PAST	
A, B, C, D	iš-**k**en	iž-ip-ken	iž-ip aldɨm	'drink'

DIALECTS				
A, C	či-GEN	či-p-ken	či-p aldɨm	'eat'
D[29]	če-en	či-p-ken	či-p aldɨm	
B	**cËen**	či-p-ken	či-p aldɨm	

In experimental field conditions, we found speakers could be 'forced' to produce forms like *či.en* 'eat-PAST', if we required them to stay within paradigm I. We refer to these forced outputs as 'pseudo-optimal' outputs (83).

(83)

<u>UR</u>	<u>pseudo-optimal output</u>	
čig + gen	či.en	'eat-PAST'
čɨg + gan	čɨ.an	'gather- PAST'
ug + gan	u.an	'lift'-PAST

Under such conditions, speakers were very clear in their judgments that if there were paradigm I outputs for inputs such as /či + gen/, then the *optimal and only possible* outputs would be those containing hiatus (i.e. an internal, onsetless syllable). We emphasize that hiatus is found nowhere else in this dialect, nor in any other dialect of Tuvan. Though generally aware of forms used in dialects A and C, where hiatus is avoided (or repaired), speakers rejected these.

The gap between competence and performance is striking. Speakers clearly know the relevant ranking of their constraints and can be induced to produce pseudo-optimal, non-null output(s) dictated by the constraint ranking they employ. However, they systematically avoid such outputs via paradigm skipping. Forced to stay within paradigm I, speakers prefer

[29] Dialect D represents the Tuvan spoken in the far north of Tuva in close proximity to the Tofa and Xakas speech communites.

as optimal a null parse—no output for a given input. We propose that this constitutes an example of *ungrammaticality*, as defined in Prince and Smolensky 1993.

7.1 Deriving ungrammaticality

In this section we review a proposal for modeling *ungrammaticality*: a situation in which the observed constraint hierarchy applied to the plausible candidate set yields no output. We explore the ability of each model to handle the facts of Tuvan.

Prince and Smolensky (1993) propose a special constraint MPARSE to account for such cases.

MPARSE *Morphemes in the input must be parsed in the output.*

MPARSE is universally high-ranked and in most cases, undominated. It is uniquely violated by a Null Parse (no output). The Null Parse is claimed to fully satisfy all faithfulness and markedness constraints. Since it accrues only a single violation (for MPARSE), Null Parse emerges as the winning candidate when all other candidates violate some constraint that outranks MPARSE.

(84)

		CONSTRAINT A	MPARSE
	Candidate (a)	*!	
☞	Null Parse		*

Any constraint that outranks MPARSE is unviolated in any output structure, because the null output would be more optimal in such cases. If two constraints are in a relation such that a violation of the lower ranked one would satisfy the higher ranked one, their individual ranking vis-à-vis MPARSE becomes crucial. For example, given the following ranking, the optimal output will be a non-null one if it violates B to satisfy A:

CONSTRAINT A >> MPARSE >> CONSTRAINT B

(85)

	CONSTRAINT A	MPARSE	CONSTRAINT B
Candidate (a)	*!		
☞ Candidate (b)			*
Null Parse		*!	

But given the ranking:

CONSTRAINT A >> CONSTRAINT B >> MPARSE

the optimal output will be the Null Parse. In this case, even if B is violated to satisfy A, the violation of B is still worse than a null parse, which emerges as optimal.

(86)

	CONSTRAINT A	CONSTRAINT B	MPARSE
Candidate (a)	*!		
Candidate (b)		*!	
☞ Null Parse			*

In Standard Tuvan, the optimal output for input /či + gen/ is *či-gen*. The constraint hierarchy avoids onsetless syllables and hiatus in one fell swoop. It satisfies higher-ranked FAITH-IO.V1 and ONSET by forcing a violation of the lower-ranked *VGV. Since there is no instance of ungrammaticality in the Standard dialect, we may assume that MPARSE is undominated, and that its high ranking rules out the null candidate.

(87) Standard Tuvan

/či + gen/	MPARSE	MAXNONHI	FAITH-IO.V1	ONSET	*VGV
☞ a. či.gen					*
b. či.en				*!	
c. čeen			!*		
d. čiin		*!			
e. Null Parse	*!				

This result also satisfies recoverability, as it protects the sole vowel of a short stem from assimilation. But *VGV is violated only to satisfy higher-ranked FAITH-IO.V1. When the vowel is not threatened by assimilation, FAITH-IO.V1 is fully satisfied and cannot force a violation of *VGV. This is observed in disyllabic words (88).

(88)

	/biži + gen/	MPARSE	MAXNONHI	FAITH-IO.V1	ONSET	*VGV
	a. biži.gen					*!
	b. biži.en				*!	
☞	c. bižeen					
	d. bižiin		*!			
	e. Null Parse	*!				

‘write’-PAST

We now turn to the Süt-xöl dialect. We shall first determine how the ‘pseudo-optimal’ candidate [či.en] is selected. A minimally different ranking from that of Standard Tuvan will suffice.

MPARSE >> FAITH-IO.V1, MAXNONHI, VGV >> ONSET

This ranking ensures that the velar will always be deleted, as is indeed the case in this dialect. The pseudo-optimal output may violate ONSET in order to satisfy *VGV.

(89) Süt-xöl *pseudo-optimal* output for /či + gen/

	/či + gen/	MPARSE	MAXNONHI	FAITH-IO.V1	*VGV	ONSET
	a. či.gen				*!	
☞	b. či.en					*
	c. čeen			!*		
	d. čiin		*!			
	e. Null Parse	*!				

'eat'-PAST

For inputs where the vowel of the initial syllable is not threatened, the output is identical to that of Standard dialect. No ONSET violation will surface.

(90) Süt-xöl *optimal* output for /biži + gen/

	/biži + gen/	MPARSE	FAITH-IO.V1	MAXNONHI	*VGV	ONSET
	a. biži.gan				*!	
	b. biži.an					*!
☞	c. bižeen					
	d. bižiin			*!		
	e. Null Parse	*!				

'write'-PAST

We found that Süt-xöl speakers always deleted /g/ in an intervocalic context in derived environments. Clearly, *VGV is highly ranked in this dialect, and cannot be violated to repair a sub-minimal output. The resulting hiatus also demonstrates that speakers will not resolve hiatus at the expense of the sole stem vowel. We assume therefore that the *recoverability* constraint, FAITH-IO-V1, is also quite highly-ranked.

7.2 Ungrammaticality and paradigm skipping

We now turn to the *actual* optimal out put for /či + gen/ in Süt-xöl. The attested winning candidate is the Null Parse. This forces us to revise our ranking by demoting MPARSE.

(91) Süt-xöl optimal output for /či + gen/

	/či + gen/	ONSET	*VGV	FAITH-IO.V1	MAXNONHI	MPARSE
	a. či.gen		*!			
	b. či.en	*!				
	c. čeen			!*		
	d. čiin				*!	
☞	e. Null Parse					*

'eat'-PAST

In order for Null Parse to win legitimately, any available phonological repair strategy must be ruled out. This can be accomplished only if the constraint that drives the repair ranks above MPARSE. Output *či.gen,* in which an ONSET violation is averted, can only be ruled out by the ranking *VGV >> MPARSE. But ample evidence from other output forms shows that *VGV is frequently violated. It cannot therefore outrank MPARSE in the grammar of this dialect. On the other hand, if we rank *VGV below MPARSE, we are at a loss to explain why the onsetless syllable is not avoided to yield [či.gen]. Evidence from numerous other outputs in this dialect suggest the sub-optimal output *could be* repaired by parsing the velar. Since intervocalic velars do surface in this dialect to satisfy recoverability, we have no explanation for why they should not do so to satisfy MPARSE. This leaves us with a ranking paradox. The ranking needed to generate the attested Null Parse is (*VGV >> MPARSE); this is contradicted by numerous non-null outputs that do manage to violate *VGV.

We proposed that ONSET outranks MPARSE. This ranking was needed in (91) to rule out [či.en] and select as optimal the Null Parse. But, this ranking appears to have the undesirable result of eliminating *all* onsetless words from the grammar (e.g. *er* 'male'). However, we note that Onset is violated *only* in the special case of word-initial onsetless syllables. Thus, we appeal to the alignment constraint to rescue these forms:

ALIGN (PRWD-L, STEM-L) *The edge of a prosodic word must coincide with phonological material that belongs to a stem*

(92)

	/er/	ALIGN	MPARSE	ONSET
☞	a. er			*
	b. Null Parse	*!	*	

'male'

7.3 Semantic faithfulness

Paradigm skipping entails unfaithfulness to the semantic content of the input. We assume the existence of a constraint requiring speakers to faithfully parse the semantic content of the input:

F-SEMANTIC *The semantic content of the input must be faithfully represented in the output.*

When F-SEMANTIC ranks below a phonological constraint, speakers may violate it to avoid the phonologically ill-formed output. This trade-off between phonological form and semantic content is exemplified by paradigm skipping in Süt-xöl.

(93) Süt-xöl optimal output for /či + gen/ entails paradigm skipping

	/či + gen/	ONSET	*VGV	FAITH-IO.V1	MAXNONHI	MPARSE	F-SEMANTIC
	a. či.gen		*!				
	b. či.en	*!					
	c. čeen			!*			
	d. čiin				*!		
	e. Null Parse					*!	
☞	f. čip.ken						*
☞	g. čip aldɨm						*
	h. Tuva						* *!

'eat'-PAST

We note that by ranking F-SEMANTIC below MPARSE, speakers are able to chose a form from a different paradigm, indicated by the lower portion of the tableau. They chose one that still comes as close as possible semantically. Thus, candidates (f) and (g) are evaluated by F-SEMANTIC as being equally good. This bears out our observation that speakers are likely to produce either form in response to input /či + gen/ . We further note that F-SEMANTIC must be gradiently evaluated. Speakers are not free to chose any word, so candidate (h) 'Tuva' must fail because it more flagrantly violates F-SEMANTIC by preserving *none* of the semantic content of the input.

Appendix A

Velar deletion and morphological blocking environments in Tuvan, Tofa and Xakas.

	MORPHOLOGICAL ENVIRONMENT	TUVAN	XAKAS	TOFA	VELAR DELETION? TUVAN	XAKAS	TOFA
1	Root-final & root-internal velars	VG VGV	VG VGV	VG VGV	no	no	no
2	Long affixes	GV(C)	GV(C)	GV(C)	no	no	no
3	Short stem + velar-initial long affix	CAg +/GAC/	CAg +/GAC/	CAg +/GAC/	yes	yes	yes
4	Short stem + velar initial long affix	CIg +/GAC/	CIg +/GAC/	CIg +/GAC/	no	yes	yes
5	Short stem + velar initial long affix	CI +/GVC/	CI +/GVC/	CI +/GVC/	no	yes	yes
	LONG AFFIXES						
6	PAST	/-GAn/	/-GAn/	/-GAn/	yes	yes	yes
7	CONVERB	/-GAš/	/-GAš/	/-GAš/	yes	yes	yes
8	CONCILIATORY / OPTATIVE	/-GAy/	/-GAy/	/-GAy/	yes	yes	yes
9	UNACCOMPLISHED / ITERATIVE	/-GAlAk/	/-GAlAk/	—	yes	yes	—
10	HYPOTHETICAL	—	/-GAdAg/	—	—	yes	—
11	INTENSIFIER	—	/-GInA/	—	—	yes	—
12	PRETEND	—	/-GAčAk/	—	—	yes	—
	SHORT AFFIXES						
13	DATIVE	/-GA/	/-GA/	/-GA/	no	yes	yes
14	LIMITED MOOD	/-GI=že/	—	/-GIšA/	no	—	yes
15	MODAL	/-GI deg/	—	/-GI deg/	no	—	yes
16	LOC + DERIVATIONAL	/-DA-gI/	/-DA-gI/	/-DA-gI/	no	no	yes
17	DERIVATIONAL	/-GI/	/-GI/	/-GI/	no	yes	yes

Notes to Appendix A

Environments 1-5 consist in nouns and verb stems and fully lexicalized (unproductive) affixal morphology containing velars. In non-derived environments (1, 2) none of the languages allow deletion. In derived environments, deletion of a stem-final (3) or affix-initial (4) velar will usually obtain, but may be blocked to protect the initial syllable vowel. This blocking effect applies only in Tuvan, not in Xakas and Tofa.

Environments 6-12 consist in productive **long** verbal affixes having the form /GVC/ or /GVCV(C)/. These pattern alike in all three languages, uniformly allowing velar deletion.

Environments 13-17 consist in nominal and verbal **short** affixes. These are either simplex affixes of the form /GV/ or morphologically more complex affixes containing a short /GV/ affix. These affixes reveal the minimally different recoverability threshold of Tuvan, which blocks velar deletion in such cases. Tofa, as well as Xakas, with a single exception (16), allow velar deletion here.

Chapter Four: Tuvan Vowel Harmony

0. Introduction

In this chapter and chapter 5 (Reduplication and Harmony), we present an empirical and theoretical account of Tuvan vowel harmony (VH). The Altaic family has a high concentration of the world's rounding harmony systems. It also has the typologically less common backness harmony (Anderson 1980). We examine in detail the operation of backness harmony (BH) and rounding harmony (RH) in Tuvan. We also show the interaction between the two systems.

Adopting an Optimality Theoretic framework, we employ constraints that govern faithfulness, markedness, and feature alignment. For each class of constraints, we consider their role in providing an adequate explanatory framework for modeling vowel harmony. We conclude that much of Tuvan VH is readily handled within standard OT models of VH (Smolensky 1993, Kaun 1995, 2000, Cole & Kisseberth 1994a, 1995b). We also examine whether the Tuvan data call for a separate class of constraints on feature licensing (Zoll 1998, Piggott 1999) and conclude that they do not. Apparent licensing effects are shown to be epiphenomenal and to fall out from the application of the alignment constraints.

In addition to the basic patterns of VH, we introduce a number of secondary VH phenomena in Tuvan that have not been widely documented in vowel harmony languages (e.g. Altaic). Further, these effects are not attested in contexts generally encountered in Tuvan, but emerge only when VH applies in novel and/or disharmonic environments such as those provided by loanwords and reduplication. These secondary, emergent VH phenomena supply us with crucial information regarding the following issues:

1. Directionality
2. Active role of the feature [-round]
3. Length conditioning

This chapter explores the basic VH patterns in detail and sets the stage for discussion in chapter 5 of more complex harmony patterns that emerge in novel contexts.

1.0 Vowel harmony

Vowel harmony may be most simply described as a requirement that certain vowels agree with other vowels in terms of a particular phonological feature. Harmony may be based on, for example, backness (Finnish), ATR (Even), height (Shona) or rounding (Turkish). Tuvan exhibits both backness harmony (BH) and rounding harmony (RH).

The symmetry of the Tuvan vowel system allows segments to pattern neatly together in natural classes according to height, backness and rounding. These classes provide the bases for vowel harmony. As the harmony classes contain equal numbers of phonemes (four each), each vowel has its counterpart bearing an opposing value for the features [back], [high] and [round], respectively. No vowels are left out of the harmony system, that is no vowels ever function as neutral or transparent.

(1)

		front		*back*	
		unround	*round*	*unround*	*round*
short	*high*	i	ü	ɨ	u
	non-high	e	ö	a	o
long	*high*	ii	üü	ɨɨ	uu
	non-high	ee	öö	aa	oo

2.0 Backness Harmony

Under backness harmony (BH), all vowels in a word must belong to either the front class [i ü e ö] or the back class [ɨ u a o]. Following Chomsky and Halle (1968), we assume [back] to be an equipollent feature. Tuvan provides no clear evidence that either [+back] or [-back] is more active or more marked, and we will assume both values are present in lexical representations. BH applies robustly, both root-internally and in affixal morphology (i.e., suffixes). Post-initial root vowels and suffix vowels take their cue from the vowel to their left, whether it is in a root or another suffix. The backness value is thus fully predictable for all post-initial vowels in a word. Virtually no native Tuvan words contain vowels from both the front and back classes. The only exceptions are a small number of compound words (§2.2). The following lexemes and suffixes exemplify harmony.

(2) <u>[-back] harmonic roots</u>

idegel	'hope'
xülümzüreer	'smile'-Fut
eeren	'totem'
xöömey	'throat singing'

(3) <u>[+back] harmonic roots</u>

ɨrak	'far'
ulu	'dragon'
ayɨɨl	'danger'
oruk	'road'

(4) <u>harmonic roots with harmonizing suffixes</u>

is-ter-im-den	'footprint'-PL-1-ABL
esker-be-di-m	'notice'-NEG-PAST.II-1
udu-va-dɨ-m	'sleep'-NEG-PAST.II-1
at-tar-ɨm-dan	'name'-PL-1-ABL

The active enforcement of backness harmony is most clearly visible in suffixes, most of which have at minimum two vowel allomorphs, a front one and a back one. [30]

(5)

underlying	surface	
/-LAr/	-lar ~ -ler	PLURAL
/-DA/	-da ~ -de	ACCUSATIVE
/-DI/	-di ~ -dɨ	PAST-II

2.1 Epenthesis driven backness harmony

Loanwords containing consonant clusters in the source language are subject to cluster simplification by means of consonant deletion or vowel epenthesis. The latter provides an opportunity to observe the active application of harmony constraints in novel contexts. The fact that harmony can and does apply actively to epenthetic vowels has potential theoretical significance for modeling speakers' awareness of harmony patterns. Epenthesis-driven harmony also provides evidence that BH may operate in a directional manner.

Epenthetic vowels, except under special conditions discussed in chapter 5, are always high in Tuvan. Epenthetic vowels (underlined in the data below) appearing in a post-initial syllable always agree in backness with the immediately preceding vowel (6). In the

[30] Additional rounded-vowel allomorphs exist for suffixes containing high vowels, as will be discussed in §3.3.

second column, we give the Russian source words. Recall that Russian stress is rendered as vowel length (ch. 1), and vowels in monosyllabic loanwords are usually also rendered as long.

	Tuvan word	Russian source word	
(6)	seeki̲s	sʲeks	‘sex’
	giipi̲s	gips	‘cast’
	faarɨ̲š	farš	‘ground meat’

We refer to this as progressive BH. Prothetic vowels (7) or epenthetic vowels appearing inside a word- initial cluster (8) always agree in backness with the vowel to their right in backness.

	Tuvan word	Russian source word	
(7)	ɨ̲šqoola	ˈškola	‘school’
	ɨ̲staqaan	staˈkanʲ	‘drinking glass’
	i̲spiirt	spirt	‘grain alcohol’
(8)	pɨ̲laan	pɭan	‘plan’
	pɨ̲loop	pɭof	‘rice pilaf’

We refer to the above as regressive BH. If an epenthetic vowel sits between two vowels that disagree in backness, it always assumes the backness value of the vowel to its left (9). This indicates that when progressive and regressive BH conflict, it is progressive BH that wins out.

	Tuvan word	Russian source word	
(9)	ačɨ̲ki, *ači̲ki	ačˈki	‘eye-glasses’
	texi̲naar, *texɨ̲naar	texˈnar	‘grain alcohol’
	partɨ̲fel, *parti̲fel	partˈfʲel	‘wallet’

Before introducing a formal model of BH, we briefly discuss the contexts for backness disharmony.

2.2 Backness Disharmony

Turkic languages show varying degrees of tolerance for disharmony. Uzbek, which enforces BH only very weakly if at all, might be situated at the least harmonic extreme along a continuum of Turkic languages. Standard Uzbek no longer shows suffixal alternations of the type *-ler* ~ *-lar*, but simply uses *-lar* (PLURAL) everywhere (Farrukh Suvankulov, p.c.).

BH in Uzbek may be interpreted as merely an historical relic in roots, not as an actively-enforced well-formedness constraint. Tuvan, by contrast, is at the more harmonic end of the continuum, enforcing BH robustly. As do other Turkic harmony systems (e.g. Uighur, Turkish), Tuvan has a few suffixes containing non-alternating vowels (underlined below). These four non-harmonizing elements constitute the entire class of morphological exceptions to harmony. They may be classified respectively as invariant (non-alternating) elements (10), borrowed suffixal elements (11), or fused elements (12, 13).

(10)	ALLATIVE	/ =Je /	
	aal=ǯe		'yurt'-ALL
	aas=če		'mouth'- ALL
	diis=če		'cat'- ALL
(11)	DIMINUTIVE	/ -maa /	
	belek-maa		'gift'-DIM
	sedeŋ-maa		male given name- DIM
	oyu-maa		female given name- DIM
(12)	DURATIVE	/ -BIšaan /	
	ɨrla-vɨšaan		'still singing'
	čemnen-mišaan		'still eating'
(13)	Sequential	/ -BIšaan /	
	piǯi-višaan		'while writing'
	ɨrla-vɨšaan		'while singing'

Tuvan also has a small number of compound words (most include the negative marker *čok*) that are back-disharmonic (14). When they take suffixes, the suffix agrees in backness with the final syllable of the compound.

(14)	epčok	(*ep* 'comfort' + *čok* NEG)	'discomfort'
	eŋmežok		'very many'
	epčoɣum		'discomfort'-1

Besides compounding, which is marginally productive, backness disharmony may arise from a morphologically productive process of ablaut. In Tofa (Rassadin 1978) and the Tozha dialect of Tuvan (Z. Chadamba 1972, p.c.), the intensive degree is formed by applying ablaut to the second syllable vowel of any disyllabic adjective: *biče* 'small' *bičii* 'very small'. The ablauted vowel is always front, long and unrounded (either [ee] or [ii], lexically determined. For [+back] adjectives (15 a-d), this productive morphological process yields disharmonic segments (underlined below).

	adjective	intensified form	
(15) a.	čɨlɨg	čɨleeg	'warm'
b.	kɨrgan	kɨrgeen	'old'
c.	bàhay	bàheey	'bad'
d.	čaraš	čariiš	'little'
e.	biče	bičii	'small'
f.	ninge	ningii	'thin'
g.	ulu	ulee	'big'

Co-articulatory exceptions to BH may be found in fluent speech. For example, a palatal glide in a syllable coda may exert a fronting effect on [+back] vowels. Compare standard Tuvan *ačay* 'father' vs. dialectal Tuvan *ačey* 'father', where the palatal glide triggers vowel fronting. Palatal fronting yields disharmonic forms in a number of Turkic languages, e.g. Dolgan (John Ziker, p.c.) and Tuha.

Similarly, the [+back] value of the enclitic auxiliary verb [aar] may exert an assimilatory effect on the final vowel of a [-back] converb, thus yielding a disharmonic form (16). Note that harmony does not extend leftward across the entire word span. Only the final stem vowel becomes [+back] in its articulation. We interpret his effect as purely articulatory in nature, not phonological.

	harmonic (careful speech)	disharmonic (rapid speech)	
(16)	piʒip # aar	piʒɨvaar	'write'-CV -SBEN-P/F
	čemnenip # aar	čemnenɨvaar	'eat'- CV -SBEN-P/F

Lexical exceptions to back harmony may be found in numerous (recent) Russian loanwords. Speakers occasionally alter the shape of borrowings to conform to back harmony (cf. 'television set'), but they generally allow disharmonic vowels (underlined) to remain (17).

	Tuvan word	Russian source word	
(17)	teleVizer ~ teleVizar	teleˈvizar	'television set'
	kinoo	kiˈno	'cinema'
	sizo	siˈzo	'jail'
	piiva	ˈpiva	'beer'
	sudija	suˈdja	'judge'
	padeš	paˈdʲeš	'case' (grammatical)
	qambeet	kanˈfʲet	'candy'
	kaačɨstfa ~ kaačistfa	ˈkačistfa	'quality'

Loanwords (Mongolian or Russian) belonging to an earlier stratum of the lexicon (e.g. words borrowed prior to the advent of widespread bilingualism in the 1950's), have nearly all been subjected to BH (Baskakov 1958, Tatarintsev 1976). Disharmonic vowels in the source words are underlined below.

	Tuvan	Mongolian (disharmonic)	
(18)	čurumal	žirumal	'pattern'
	araga	araki	'alcohol'

	Tuvan	Russian (disharmonic)	
(19)	xapajaq	kaˈpʲejek	'kopeck'
	mašɨna	maˈšina	'automobile'

Many loanwords, regardless of their source language, happen to be accidentally harmonic, even though they may diverge from native phonotactics in other ways (20). It is not therefore the case that the native/non-native distinction coincides with harmonic/disharmonic classes in any regular way. This fact will become relevent in our discussion of re-harmonization patterns in chapter 5.

(20)	fiidik[31]	ˈvidik (Rus.)	'video cassette'
	pleeyer	ˈpleyer (Rus.)	'video player'
	lama	lama (Sanskrit)	'lama'

2.3 A formal model of BH

In our formal model, we adopt standard alignment constraints that enforce harmony. In addition, we assume in this chapter underspecified underlying representations for vowels that are harmonic and thus predictable. The argument for underspecification, which we develop in chapter 5, rests largely on patterns of vowel re-harmonization. We show that under a productive process of reduplication, harmonic and disharmonic segments pattern differently. Disharmonic segments are shown to resist re-harmonization, while harmonic segments universally undergo it. In this chapter, in anticipation of the arguments advanced in chapter 5, we adopt underspecified representations for all post-initial, harmonic vowels.

2.4 Harmony as Alignment

Following the theory of Generalized Alignment (McCarthy and Prince 1993) we model BH as a high-ranking alignment constraint. This requires the feature [+/-back], present underlyingly in the initial syllable, to align with the edges of the word domain.

ALIGN[αBACK]-PRWD *Align [αback] with edges of the prosodic word*.

BH does not typically extend beyond the domain of the word. Some enclitics, for example, fail to undergo BH.

(21)	on=daa	'I'=EMPH
	men=daa (*=dee)	'he'=EMPH

The alignment constraint determines the backness value of affixal and epenthetic vowels, and also serves to rule out vowel sequences that disagree in backness. Special faithfulness constraints that enforce enhanced faithfulness to suffixes or roots will allow the occasional disharmonic form to surface:

FAITH.IO-AFFIX *An underlying affix must be faithfully rendered in the output*
FAITH.IO.ROOT *An underlying root must be faithfully rendered in the output*

If these rank above the alignment constraint, an affix or root that is underlyingly disharmonic will remain so (cf. Chapter 5). The re-ranking of harmony constraints vis-à-vis root faithfulness may be observed as an historical change in Tuvan. At an earlier stage, back disharmonic segments in loanwords were consistently subjected to harmony. In contemporary Tuvan, back disharmonic segments are usually tolerated. In (§3.4), we demonstrate that a similar process has occurred for rounding harmony.

[31] Tuvan has no native phoneme [f]

2.5 Directionality

In the native lexicon, speakers have no evidence that BH applies in any particular direction, i.e. right-to-left or left-to-right. It makes no difference for the underlying representations whether [+back] is assigned to the first syllable, another syllable, or the entire word domain (Cole & Kisseberth 1994a). The feature can still be viewed as aligning with both edges of a word domain, yielding a fully harmonic output. However, in the context of epenthesis in consonant clusters, we have seen that [+/-back] must align in a directional manner (e.g. leftwards) to target an epenthetic vowel (underlined).

(22)

Tuvan	Russian source word	
ɨ̱šqoola	škola	'school'
ɨ̱raqeeta	rakjeta	'rocket'
pɨ̱lan	pɭan	'plan'
i̱spirt	spirt	'grain alcohol'

But epenthetic vowels are only targeted by regressive BH when they appear to the left of the first underlying vowel of a word. Elsewhere, they are targeted by progressive BH. In disharmonic loanwords, a post-initial epenthetic vowel takes on the backness value of the vowel to its left rather than that of the vowel to its right.

(23)

Tuvan	Russian source word	
ačɨ̱kii, *ači̱kii	ačˈki	'eye-glasses'
texi̱naar, *texɨ̱naar	texˈ nar	'grain alcohol'
pɨ̱razdɨ̱nik, * pɨ̱razdi̱nik	ˈpraz(d)nik	'holiday'

To capture these directional effects, we adopt two distinct alignment constraints.

ALIGN[αBACK]-L-WD	*Align [αback] with left edge of word*
ALIGN[αBACK]-R-WD	*Align [αback] with right edge of word*

The forms given in (23) demonstrate the potential conflict between the alignment constraints. The attested outputs will surface only if rightward alignment outranks leftward alignment.

ALIGN[αBACK]-R-WD >> ALIGN[αBACK]-L-WD

We note that other factors (e.g. stress, vowel quantity) do not exert any effect on the backness value of the epenthetic vowel. Epenthetic vowel backness correlates strictly with the backness of the relevant neighboring vowel.

2.6 Directionality as a 'reserve' constraint

Regressive BH appears only in novel contexts such as epenthesis. This raises the question of how speakers could be aware of a constraint ALIGN-LEFT that plays no observable role in the native phonology. Further, how do they know that ALIGN-RIGHT dominates ALIGN-LEFT? In experimental field conditions, we found that monolingual Tuvan speakers, when confronted with the novel context of impermissible clusters in loanwords, seemed to show no hesitation in harmonizing the epenthetic vowel as described above. We found this to be true of both very young speakers, who had probably never previously encountered any loanwords, and elderly speakers who were asked to repeat novel loanwords.

We propose that ALIGN-LEFT may be an example of an 'inactive' or 'reserve' constraint (Ross 1996, Kaun 1999a, b, Harrison 1999). Reserve constraints are those that never exert any observable effects in the native phonology, due to configurational properties of the language. Tuvan, for example, is exclusively suffixing, so constraints governing regressive harmony (e.g. harmony targeting prefix vowels) simply never do any work in the native grammar. Novel contexts such as loanword phonology and reduplication provide an opportunity for reserve constraints like ALIGN-LEFT to emerge and show effects. Reserve constraints are somewhat different from a class of phonological effects known as the emergence of the unmarked (TETU) (McCarthy and Prince 1994). Under TETU, it is only markedness constraints (which are normally outranked by faithfulness constraints) that emerge and show effects. Reserve constraints give rise to a class of effects that do not exactly overlap with TETU effects. We propose one way that reserve constraints differ is that they are not limited to the markedness family, but may include alignment constraints as well.

A further possible difference between TETU and reserve constraints is that the latter may give rise to considerable variation among or within individual speakers. Kaun (1999a) demonstrates that Turkish speakers, when confronted with impermissible clusters in loanwords, seem to know that, as a general strategy, they should not only epenthesize a vowel but should also apply regressive VH to the epenthetic vowel. However, speakers exhibit significant differences among themselves in the actual application of VH to the epenthetic vowel, and also variation at the level of individual grammars. This result has been confirmed by a different set of experimental results from Turkish reported in Harrison & Kaun (2000). The presence of variation may reflect properties of both the grammar itself and the speech community. While speakers may, in fact, have access to reserve constraints, they remain uncertain as to the exact ranking of such within the hierarchy because they lack direct evidence. Further, as reserve constraints are most readily detected in the context of novel configurations (e.g., prefixes in an exclusively suffixing language and loanwords) the speech community has not formed a consensus on how such cases are to be handled.

We note that a markedness constraint *alone* cannot account for the systematic way in which speakers assign a backness value to epenthetic vowels in disharmonic environments. Relying on markedness alone, speakers should consistently prefer to make the epenthetic vowel either front or back, depending on whichever class is less marked. They do not do so. Nor is it the case that speakers automatically adopt the backness value of vowel of the first (and most prominent) or final (and stressed) syllable. Instead of markedness or positional faithfulness, then, speakers seem to rely on strict directionality. If speakers do possess the constraint sub-ranking ALIGN[αBACK]-R-WD >> ALIGN[αBACK]-L-WD in their grammar, as we have claimed, then it is no surprise that they are able to apply it consistently in disharmonic loanwords. Though speakers never make use of this sub-ranking in the native grammar, they can call upon it to enforce harmony in completely novel contexts. If this ranking were not fixed, we might expect speakers to be uncertain or produce variation in epenthesis. In fact, we found that speakers never hesitated in enforcing progressive harmony on the epenthetic vowel, even in situation where regressive harmony could have been applied to yield a different result (see Tableau D below).

In the following tableaus, we model standard backness alignment, disharmony, and epenthesis-driven BH. We assume suffix vowels (shown as archiphonemes) to be unspecified for backness in the input. For a detailed argument on the necessity and distribution of underspecification, we refer the reader to chapter 5. Tableau (A) illustrates standard rightward alignment. As Tuvan is exclusively suffixing, the proposed ALIGN-LEFT constraint normally has no effect, so we omit it in this tableau.

(24) Tableau A

	/is-tAr-Im/ | -B	FAITH.IO-[BACK]	ALIGN[αBACK]-R-WD
	is-tar-ɨm | -B		*! *
	is-ter-ɨm | / -B		*!
☞	is-ter-im | / / -B		

‘footprint’-PL-1

Tableau B shows progressive alignment of backness to an epenthetic vowel. We assume that an epenthetic vowel bears no underlying backness specification. As such, this vowel does not incur violations for failing to align or instantiate its own backness feature.

(25) Tableau B

	/gips/ | -BK	*COMPLEX	ALIGN[αBACK]-R-WD
	gips | -BK	*!	
	gipɨs | -BK		*!
☞	gipis | / -BK		

‘plaster cast’

Tableau C illustrates faithfulness to underlying backness specification in disharmonic forms, e.g. compounds or loanwords. We assume these to be fully specified. The two alignment constraints conflict here, but neither has a decisive role in selecting the output.

(26) Tableau C

	/ep + čok/ | | -B +B	FAITH.IO-[BK]	ALIGN[αBACK]-R-WD	ALIGN[αBACK]-L-WD
	epček | / -B	*!		
	apčok \ | +B	*!		
☞	epčok | | -B +B		*	*

Tableau D illustrates leftward alignment of backness to an epenthetic vowel.

(27) Tableau D

	/plan/ | +B	*COMPLEX	ALIGN[αBACK]-R-WD	ALIGN[αBACK]-L-WD
	plan | +B	*!		
	pi̲lan | +B			*!
☞	pɨ̲lan \ | +B			

Tableau D shows the dominance of the rightward over leftward alignment in cases of conflicting right/left directionality

(28) Tableau E

/texnaar/ |　| -BK +BK	*CODA[x][32]	FAITH.IO-[BK]	ALIGN[αBK] -R-WD	ALIGN[αBK] -L-WD
texnaar |　| -B +B	*!		*	
texịnaar |　\\ | B +B			* *!	*
☞ texịnaar |/　| B +B			*	**

2.7 Consequences of BH for underspecification

We adopt the view that in vowel harmony languages, speakers' phonological competence includes knowledge of harmony as a pervasive pattern and of disharmony as a marked departure from that pattern. In this chapter and the next one, we will demonstrate how speakers deploy their knowledge of harmonic patterns to construct underlying representations. These representations are crucially underspecified, we claim, for the harmonic feature(s). This gives rise to a basic distinction between harmonic (contextually predictable) and disharmonic (contextually unpredictable) vowel segments.

In chapter 5 we demonstrate that underspecification is more than just a useful formal device, and that it has observable consequences for speakers' ability to manipulate harmonic and disharmonic strings. Harmonic and disharmonic segments are shown to pattern differently in the context of reduplication that gives rise to novel vowel alternations within roots. The distinct patterning is best understood, we argue, not as a result of surface constraints alone, but as the consequence of differing underlying representations. In chapter 5 and in Harrison & Kaun (2000), we introduce a preliminary proposal for a model of grammar (more precisely, a model of *lexicon optimization*). The model allows underlying representations to be unspecified for predictable (i.e. harmonic) features. It also attempts to define the range of contexts in which speakers might be driven to posit underspecification.

2.8 The status of underlying representations

Generative linguistics has long assumed the principle of lexical minimality (Chomsky & Halle 1968), which requires redundancy in lexical representations to be minimized. Representations may be underspecified for certain features that are filled in at a later stage of derivation by phonological rules (Kiparsky 1993). Various diagnostics have been proposed to identify the class of features which could be unspecified (Archangeli 1984, Steriade 1995). These include features that are predictable, unmarked, or redundant. In harmony systems, features governed by harmony constraints are fully predictable based on adjacent segments. In a typical Turkic harmony system such as that found in Tuvan, this holds true for all harmonic features in post-initial syllables.

In contrast to lexical minimality, Optimality Theory (Prince and Smolensky 1993) has provided a framework in which underlying representations have less theoretical and practical utility. Instead, grammars are defined almost exclusively in terms of output well-formedness. The "richness of the base" hypothesis (Prince and Smolensky 1993) has been widely interpreted as a requirement that the grammar be able to generate a well-formed output for any possible input. As a consequence, the set of inputs to the grammar is construed as infinite. No restrictions may be placed on input forms. The task of filtering out ill-formed outputs is to be done by output constraints alone.

In this chapter, we introduce an argument that harmony systems can give rise to underspecified inputs like those employed in the preceding tableaus. We consider an explicit model that predicts the distribution of underspecification (Inkelas 1995). Further, we find that the model does not adequately account for the operation of harmony systems. We argue that if harmony is to be understood as an active process targeting all post-initial vowels, then those vowels may not be pre-specified for the harmonic feature (e.g. backness) in input representations. This claim about the nature of underlying representations in harmony languages is presented in greater detail in chapter 5 and other recent work

---

[32] Vowel epenthesis here is driven by a ban on [x] in coda position.

(Harrison and Kaun 2000a, b). The claim that inputs are crucially underspecified is controversial, both in some pre-OT, and most OT models of grammar. However, we will present empirical evidence that speaks to the active status of root-internal harmony that targets non-alternating vowels.

A longstanding debate over the psychological reality of harmony centers around the question of whether speakers notice that non-alternating segments are harmonic at all (Zimmer 1969, Clements & Sezer 1982, Inkelas 1995, Yip 1998, Kirchner 1993, Inkelas, Orgun & Zoll 1997). Zimmer (1969) asked Turkish speakers to evaluate harmonic and disharmonic pseudo-words, and concluded that speakers were 'only imperfectly aware' of regularities in vowel co-occurrence. Root vowels in Turkish, which never alternate, may appear trivially to obey harmony, but cannot be thought of as undergoing harmony. Clements and Sezer (1982:226) take a similar position, stating that:

> …the burden of proof is on the linguist who wishes to demonstrate that roots [in Turkish] are governed by vowel harmony at all.

On this view, harmony within roots is merely a philological or historical fact, not a phonological one. An opposing claim was advanced by Campbell (1980), who asserts the 'psychological reality' of root-internal backness harmony in Finnish. On the view which we develop in chapter 5, harmony within roots is a 'linguistic' fact and does not depend on the presence of surface alternations. We hypothesize that speakers do in fact notice harmony patterns, even in the absence of surface alternations, and that this has observable consequences for the grammar.

In an OT framework, a commitment to enriched inputs requires that the grammar yield a well-formed output for any reasonable input. Thus, underspecification of features is viewed as an unnecessary and undesirable limitation on inputs. Itô, Mester and Padgett (1995) articulate this position:

> ...there is no need for a separate theory of feature minimization: The constraint hierarchy itself forces the correct output, irrespective of specification in the input.

This means that there is no grammatical imperative against even a redundantly specified input form... (1995:28)

Yip (1998) concurs:

> ...there is no evidence that the child learns abstract underlying vowel phonemes; the question becomes almost irrelevant from an OT perspective, because any reasonable input will give the right output, and the abstract input enjoys no special advantage.

Further, she notes that "...the easiest strategy may be to stick close to the phonetic surface form, absent clear guidance to the contrary." This formulation clearly leaves open the possibility that presented with 'clear guidance to the contrary,' the child *could* posit abstract representations.

This possible need for abstract representations is one that Prince and Smolensky anticipated (but did not explicitly predict) in their original model. While early OT work allows for underspecified inputs, it provides no clear diagnostic as to when such would be posited by speakers. Current interpretations of Lexicon Optimization (Inkelas 1995, Yip 1998) predict speakers will make use of underspecification in lexical entries only under highly restricted circumstances. Inkelas' (1995) model, *Archiphonemic Underspecification*, posits that speakers only underspecify segments that predictably alternate.

Inkelas' (1995) model of Archiphonemic Underspecification attempts to predict the distribution of underspecification. On her view, underspecification may arise only in instances where there are fully regular and predictable surface alternations. Turkish backness harmony within roots is subject to numerous exceptions. In such a harmony system, where the static pattern of vowel co-occurrence is highly irregular, Inkelas, Orgun & Zoll (1997) claim that:

> …empirical, theoretical and methodological reasons make it ill-advised to capture partial static distributional patterns directly in the grammar. Partial regularities should be accorded status in the linguist's grammar of a language only if there is *evidence from observable alternations* to support the generalization. (1997:1 emphasis mine)

On this view, only alternating segments will be construed by the learner as undergoing harmony.

(29) Archiphonemic underspecification

	Alternating	***Non-alternating***
Predictable	Underspecified	Fully Specified
Unpredictable	Fully Specified	Fully Specified

Applied to Tuvan, Archiphonemic underspecification predicts fully specified representations for harmonic forms:

(30) /idik-tig/ /arat-tɨg/
| | | |
-B -B +B+B
'boot'-ADJ 'peasant'-ADJ

For disharmonic forms, it likewise predicts fully specified representations:

(31) /ali-lig/
| |
+B-B
'Ali'-ADJ

On this model, all root vowels and any disharmonic affix vowels are required to be fully specified. Harmonic suffix vowels alone will be underspecified for the harmonic feature(s), because they undergo predictable surface alternations. We introduce an opposing claim here, which we develop further in chapter 5. Namely, BH in Tuvan, as in Turkish and Finnish, is a part of speakers' phonological knowledge. Even where there are no surface alternations, as in Turkish or Tuvan roots, speakers will notice a pervasive co-occurrence (harmony) pattern. Based on this pattern, they will infer the active status of harmony (alignment constraints). This may in turn guide them, in the sense of 'clear guidance' called for by Yip (1998), to posit abstract underlying representations that are underspecified for the harmonic feature [+/- back]. We thus posit underspecified underlying representations for harmonic forms:

(32) /idik-tig/ /arat-tɨg/
| |
-B +B
'boot'-ADJ 'peasant'-ADJ

But fully specified representations for disharmonic forms:

(33) /ali-lig/

+B -B

'Ali'-ADJ

The harmony patterns presented in the next chapter will provide a basis for our hypothesis of the necessity of underspecification in harmony systems.

3.0 Rounding Harmony

In this section, we provide an empirical description and theoretical discussion of the basic pattern of RH in Tuvan. This pattern accords well with RH systems commonly found in Turkic. An important theoretical goal of this section is to employ only constraints of the type that have been previously established in the OT literature and that are cross-linguistically motivated for a range of RH systems. In the following chapter, we introduce a number of exceptional environments in which RH appears to overapply, underapply, or apply in an optional or variable manner. These represent the kind of 'messy' data typically swept under the rug in standard descriptions of harmony systems (Anderson & Harrison 2000). We view these data as essential to formulating an adequate theoretical model to accommodate the full range of RH processes in Tuvan.

Most simply stated, RH requires a vowel to be [+round] when it appears in the vicinity of another rounded vowel. In an autosegmental framework (Clements and Sezer 1982), RH may be construed as spreading the feature [round] rightward from a round vowel to the vowel of the following syllable.

(34) V C_0 V

[+rd]

As Kaun (1995, 2000) notes, however, an autosegmental framework fails to capture the broad, cross-linguistic generalization that the above harmony pattern—though superficially the easiest one to describe—is only very marginally attested. In Kaun's (1995) survey, this type of harmony was in fact attested only in a single dialect of Kirghiz (Comrie 1981).

More complex RH patterns that require more complex rules are far more commonly attested. The autosegmental approach therefore offers no predictive power as to what types of systems will be attested. Nor does it offer any explanation as to why the system that is putatively simplest from a rule-based viewpoint turns out to be very rare.

3.1 A schematic typology of RH

Tuvan RH, stated in a maximally simple manner, requires a high vowel to be rounded if it follows a rounded vowel. Thus, only high vowels undergo RH but any rounded vowel triggers RH. Schematically, we represent this system as follows:

(35)

<u>Trigger(s)</u>	<u>Target(s)</u>
U, O	U

This system is quite typical within the Turkic family and is found, for example, in Turkish and Gagauz (Pokrovskaya 1964). However, RH systems that appear superficially to be identical (e.g. Tuvan and Turkish) turn out to differ greatly in their details. In addition to overall, systemic differences among RH systems, any given harmony language may be observed to possess a great deal of fine-grained variation across dialects, across speakers and even within individual speakers. In sections §3.5 through §3.7, we introduce patterns of variation at each of these levels.

When viewed in a broad, cross-linguistic typology (e.g. Adam 1898, Boguroditskij 1953, Korn 1966, Kaun 1995) RH systems are seen typically to impose certain conditions that refer to phonological dimensions other than lip-rounding. These conditions determine what segments or combinations of segments constitute the environment for the application of RH. Most commonly, they refer to the height and/or backness of trigger or target vowels. For example, in Kuu Kiži, a southern dialect of Altai (Baskakov 1958), high vowels do not trigger RH (36), nor do they undergo it (37). Only low vowels trigger and undergo RH (38).[33]

(36)	d^{j}ürgen, (*d^{j}ürgön)	'fast'
(37)	bol-dɨr, (*-dur)	'be'-FUT
(38)	bol-gon, (*-gan)	'be'-PAST

Kaun (2000) enumerates the full set of conditions that appear to (dis)favor RH. For example, RH is favored if the target is high and/or back. RH is also favored if the trigger is front and/or non-high. A condition may also be imposed on height agreement between trigger and target. Individual languages or dialects that employ RH typically select a subset of the full range of possible conditioning factors. Tuvan RH—in its most basic form—imposes a condition only on the height of target vowels. RH targets must be high. The output of RH is thus only the high rounded vowels [ü] and [u]. Standard Tuvan imposes no conditions on triggers: any rounded vowel is a potential RH trigger without regard to its height. No condition is imposed on height agreement between the trigger and target, as is the case in Kuu Kiži Altai. Finally, no condition is imposed on the backness of the trigger/target, nor on backness agreement between trigger and target. We note however, that due to pervasive BH, trigger and target vowels almost always agree in backness. There is some evidence that when trigger and target *do not* agree in backness, as in a few disharmonic loanwords, RH *does not* obtain. We attribute this not to an intrinsic limitation on RH itself, but to the active status of the feature [-round] in the underlying representation of disharmonic forms. This issue is addressed in the context of reduplication in chapter 5.

Finally, we note that the *length* of the trigger vowel is not generally a conditioning factor in Tuvan RH. This contrasts with RH in the Tungus languages Bayinna Oroch (Li 1996) and Evenki (Nedjalkov 1997), and in Daur Mongolian (Wu 1996) where long vowels fail to trigger RH, although they consistently undergo it. Shortness of trigger vowels thus

[33] 'Altai' is a cover term for a very diverse range of dialects spoken in the 'Altai' region of southern Siberia. What little documentation exists (e.g. Baskakov 1958) shows considerable dialect variation, morphological conditioning of RH, and variability within individual doalects. The examples given here are intended to illustrate one of many *attested* RH patterns of Altai, and are not necessarily representative of Altai as a whole.

seems to be a factor favoring harmony. We know of no languages where long vowels trigger RH but short vowels do not. The relationship between triggers and length is implicational: if long vowels are good RH triggers in a given language then short vowels are too. The reverse does not seem to be true. We show in chapter 5 that although length plays no role in normal Tuvan RH, it may emerge as a conditioning factor under special conditions arising from reduplication.

3.2 The status of [round] in Tuvan

From a cross-linguistic perspective (Maddieson 1984), front rounded vowels—of which Tuvan has two—are rare. In the majority of vowel inventories, the non-low vowels are unrounded if they are front and rounded if they are back (with rounding being the norm, particularly among back non-low vowels). Among the non-low vowels, this mutual relationship between backness and rounding is referred to as *enhancement* (Stevens, Kaiser & Kawasaki 1986)[34]. Enhancement in this context refers to the second formant (F2), which is lower for back vowels than for front vowels. This acoustic hallmark of backness (i.e. low F2) can be further enhanced by lip-rounding, which lowers F2 even more. Mutual enhancement of rounding and backness may serve to make contrasts among vowel qualities more salient. In a vowel inventory like that of Tuvan, which has both front rounded vowels and back unrounded vowels, the feature [round] is not enhanced by the feature [back] as it is in English. Following Kaun (1995), we propose that distinctive rounding without backness enhancement is a potentially difficult feature, not only in Tuvan but in languages in general.

Languages may employ various strategies for dealing with difficult contrasts. First, they may employ positional neutralization (Steriade 1993) to limit such contrasts to prominent syllables. Tuvan, like most Turkic languages, restricts contrastive rounding to initial syllables. Low rounded vowels [ö] and [o] never appear post-initially. High rounded vowels [ü] and [u] appear post-initially only when generated by harmony, making rounding a fully predictable feature in post-initial syllables.

[34] Enhancement is discussed for other features as well, e.g., voicing, nasality, etc.

A second strategy for difficult features is to extend their domain by spreading them. This can be done with vowel harmony, which aligns a feature across the word span. It is probably not a coincidence that languages that extend [+round] via rounding harmony tend to have both front rounded vowels and back unrounded vowels in their inventories. The spreading strategy may thus be advantageous in vowel systems where rounding is unenhanced by backness. Still, no direct, causal relationship between positional neutralization of [round] and rounding harmony has been definitively established. This remains a subject for further research.

3.3 Tuvan RH patterns

The basic operation of rounding harmony is most apparent in affixes that contain a high vowel, e.g. the 3-POSS suffix /-(z)I/. In the data below, suffix vowels targeted by RH are underlined.

(39)	lexeme	suffix		
	xöl	-ü	* -i	'lake'-3
	xöör	-ü	* -i	'cemetary'-3
	bürü	-zü	* -zi	'wolf'-3
	tool	-u	* -ɨ	'story'-3
	xol	-u	* -ɨ	'hand'-3
	mool	-u	* -ɨ	'Mongolian'-3
	ulu	-zu	* -zɨ	'dragon'-3
	xuu	-zu	* -zɨ	'percent'-3
	suur	-u	* -ɨ	'village'-3

Rounding of suffixal vowels is not permitted in the absence of RH.

(40)	lexeme	suffix		
	is	-i	* -ü	'footprint'-3
	diis	-i	* -ü	'cat'-3
	xep	-i	* -ü	'clothing'-3
	san	-ɨ	* -u	'number'-3
	ača	-zɨ	* -zu	'father'-3
	aas	-ɨ	* -u	'mouth'-3
	ɨr	-ɨ	* -u	'song-3

Were RH to under- or over-apply, outputs such as **iz-ü* or **xol-ɨ* would be possible. In fact, such forms are never attested in the lexicon and morphology of Tuvan and are judged by speakers to be ill-formed.

RH also exhibits the property of unboundedness (Anderson 1980). The feature [+round] spreads any distance within a word span, so long as potential target vowels (underlined) are available.

(41) xögʒüm- č<u>ü</u>-n<u>ü</u> 'music'-AGENT.ACC
tool-dʒ<u>u</u>-l<u>u</u>ɣ 'story'-AGENT.COMIT

RH in roots is apparent as a co-occurrence pattern among vowels. For example, sequences of [round vowel] / [high unrounded vowel] are never attested in the native lexicon, as the second vowel in such a sequence would be targeted by round harmony.

(42) *CöCi
*CüCi
*CoCɨ
*CuCɨ

Native Tuvan words never contain the sequence [unrounded vowel] / [rounded vowel], because a rounded vowel may appear only when preceded by another rounded vowel.

(43) *CiCü
*CeCü
*CɨCu
*CaCu

Finally, high round vowels [ü] [u] appear in post-initial syllables only when a RH trigger (e.g. any rounded vowel) is present in the preceding syllable.

(44) bör<u>ü</u> 'wolf'
bür<u>ü</u> 'leaf'
or<u>u</u>q 'road'
ud<u>u</u> 'sleep'-IMPER

Since RH targets only high vowels, the mid-rounded vowels [ö] [o] *never* occur post-initially in native words.[35] This is an example of positional neutralization of the feature [round] (cf. §4.0).

(45) *CüCö
*CeCö
*CɨCo

[35] The only exceptions to this are a few words that appear to be derived from compounds: *xorʒoq* 'impossible'.

3.4 RH in diachronic perspective

The diachronic operation of round harmony is apparent in the phonology of loanwords (Tatarintsev 1976, Sat 1987). Three assimilation patterns are clear in Mongolian loanwords (46, 47). Here, high vowels that follow (or in some cases, precede) a round vowel have undergone RH (vowels targeted by RH are underlined):

	Tuvan word	(Old) Mongolian source	
(46)	šügümčülel	šigümži.lel	'criticism'
	mölčükčü	mölžigči	'exploiter'

Second, regressive RH appears to have occurred in some borrowings (45).

	Tuvan word	(Old) Mongolian source	gloss
(47)	čurumal	žirumal	'pattern'
	büdügüülük	bidegüülig	'primitive'

Finally, rounded vowels that are not motivated by RH (underlined) underwent de-rounding:

	Tuvan word	(Old) Mongolian source	
(48)	tergiilegči	tergüülegči	'director'
	seriin	serigün	'cool'
	temir	temür	'iron'
	ayɨɨl	ayuul	'danger'
	alžɨɨr	alčuur	'napkin'
	oray	öröö	Mong. 'dusk', Tuv. 'late'
	ovaa	oboo	'shrine'

The diachronic facts reveal two important RH patterns. First, speakers applied RH both progressively and regressively wherever an eligible (high) target vowel was available. Second, speakers filtered out unmotivated rounded vowels in post-initial syllables. Although contemporary Tuvan differs somewhat in its enforcement of RH, we propose a formal model that handles both stages of the language.

3.5 Epenthesis-driven RH

Due to configurational properties, the normal lexicon and morphology provide no evidence as to whether RH exhibits any directional asymmetries. We claimed in §2.3 that BH does apply directionally, and that both progressive and regressive BH are attested. When the two conflict, as we have shown, the former wins out. Epenthesis also provides a unique environment in which RH might potentially apply in a directional manner.

As discussed above, consonant clusters in loanwords undergo epenthesis of a high vowel (underlined in the data below). Progressive RH obtains, as shown in (49).

	Tuvan	Russian source word	
(49)	booqus, *booqɨs	boks	‘boxing’
	sunuupqa, *sunɨɨpqa	ˈsumka	‘bag’
	gooluf, *goolɨf	golf	‘golf’
	qoovuš, *qoovɨš	kofš	‘ladle’

However, round harmony does not operate regressively to target epenthetic vowels in initial or medial clusters.

	Tuvan	Russian source word	
(50)	qɨluup, *quluup	klup	‘club’
	pɨloop, *puloop	pɭof	‘rice pilaf’
	pɨloomba, *puloomba	ˈpɭomba	‘lead seal’
	kɨrušaaf, *kurušaaf	kruˈš:of	‘Krushchev’
	abɨzoor, *abuzoor	abˈzor	‘overview’

3.6 Directionality

RH effects apply in a directional manner in several Turkic languages. We make use of these facts in our formal model (§3.0) to propose that directionality constitutes a primitive of harmony systems.

It has been argued that directionality is not a primitive of harmony systems (Beckman 1998, Lombardi 1999, Padgett 1995, Steriade 1995). Bakovic (2000) proposes a model of ‘symmetrical’ harmony driven by agreement constraints rather than alignment constraints. Agreement constraints are indifferent to the direction of feature spreading that yields harmonic forms. This model thus excludes the possibility of directionality as a primitive of harmony systems. Bakovic concludes that any apparent directional application of harmony is epiphenomenal, arising out of the morphological structures of a language (e.g. suffixing or prefixing).

A primary argument against directionality relies on the notion of morphological impoverishment (Beckman 1998, Lombardi 1999). For example, in a language with no prefixes, harmony cannot usefully be thought of as operating leftward, it is claimed, because no context for regressive application would ever arise. The same holds true for an exclusively prefixing language such as Yoruba (Akinkugbe 1978). Likewise, in languages that have both prefixes and suffixes, such as Akan (Clements 1977, 1981) the root vowel(s) clearly trigger harmony in affixal vowels. The simpler alternative is thus to describe harmony as being controlled by roots and undergone by affixes. This insight does not however, preclude the possibility of directionality. If directional effects can be found solely within roots, we can avoid the paucity of data problem arising from morphological impoverishment.

In Altaic languages, which are exclusively suffixing, epenthetic vowels may appear at both the left and right edges of the root. They thus provide ample opportunity to observe the directional nature of harmony. Altaic epenthesis-driven harmony provides clear evidence for the directional nature of RH constraints within roots. Furthermore, regressive and progressive RH are shown to pattern differently. Regressive RH, which is invisible in the native grammar, patterns differently with respect to epenthetic vowels than does progressive RH. This patterning is at odds with the normal pattern of RH, but it does accord well with harmony patterns attested in other languages.

We present a partial typology of Altaic directional effects below. We emphasize that this typology is quite preliminary: there are a great number of undocumented harmony systems (e.g. the Siberian Turkic languages Tofa, Dolgan, and Kyzyl Xakas). Few of those systems have been well-documented. Epenthesis harmony has been little studied even in well-known harmony systems. One notable exception is Turkish epenthesis harmony (Kaun1999), which has been shown to exhibit a great variety of patterns across speakers. It is thus unlikely that the full range of RH systems has been documented. It would be premature to argue against directionality, as does Bakovic (2000), on the grounds that it opens the door to a number of logically possible but unattested systems.

In order to properly situate the Tuvan facts, we have surveyed the secondary literature to assess epenthesis-driven harmony in a dozen Altaic languages, including

Evenki, Daur, Buriat (various dialects) Xakas, Gagauz, Karaim, Dolgan, Uzbek and Hungarian. Results of this research are preliminary at present, and will only be summarized briefly herein. Based primarily on facts from Turkish, Uighur (Hahn 1991), Tuvan, Buriat and Evenki (Vasilevich 1934, Nedjalkov 1999), and to a lesser extent on other Altaic languages, we make three tentative observations regarding empirical differences between *regressive rounding harmony* (RRH) and *progressive rounding harmony* (PRH).

(A) Progressive RH applies in a robust and nearly exceptionless manner, (modulo the conditioning factors enumerated herein), while regressive RH may be less robust and subject to greater variation in its application

(B) Regressive RH may be subject to different (e.g. more stringent) conditioning factor(s) than is progressive RH.

(C) In cases where regressive RH patterns in a manner not attested in the normal lexicon and morphology of a given language, it will at least pattern in accord with RH patterns found in other languages.

We note that all Altaic language are almost exclusively suffixing, such that RRH may be observed solely in novel contexts like epenthesis. Often, speakers may not all share the same intuitions about how RRH should apply to epenthetic vowels. This may be due to its limitation to novel contexts. The observations we make about RRH thus represent *emergent* properties of the grammar.

There are of course exceptions to these general tendencies. For example, In Evenki (Vasilevich 1934, Nedjalkov 1999, Budaeva 1972) RRH appears to apply more robustly than PRH to epenthetic vowels. Low epenthetic vowels that are word-initial or word-internal almost always undergo RH (51), but word-final ones usually do not (52) (but note counter-example *ʃoolk<u>o</u>* ‘silk’ below).

(51)	<u>Evenki</u>	<u>Russian source word</u>	
	<u>o</u>ʃtoora	ˈʃtora	‘curtain’
	<u>o</u>ʃkoola	ˈʃkola	‘school’
	p<u>o</u>roon	front	‘front (military)’
	g<u>o</u>wosna	gvost	‘nail’
(52)	kioosk<u>a</u>	ˈki.osk	‘kiosk’
	soort<u>a</u>	sort	‘type’
	ʃoolk<u>o</u>	ʃolk	‘silk’

Tuvan provides evidence in support of all three observations. For example, speakers *always* apply progressive RH to an epenthetic vowel (underlined).

(53)	Tuvan	Russian source word	
	gooluf, *goolɨf	golf	‘golf’
	noolu, *noolɨ	nol^j	‘zero’
	booqus, *booqɨs	boks	‘boxing’
	ruuvul, *ruuvɨl	rubl	‘ruble’

By contrast, they never apply regressive RH to a high epenthetic vowel:

(54)	Tuvan	Russian source word	
	sɨnop *sunop	snop	‘haystack’
	qɨnoopqa, *qunoopka	ˈknopka	‘button’
	ʃɨlom, *ʃulom	ʃlom	‘helmet’
	ɨroolik, *uroolik	ˈroolik	‘rollerskates’
	ʃɨlus, *ʃulus	ʃlus	‘lock (dam)’

The Tuvan facts show that epenthetic vowels fail to undergo RH uniquely in environments that may be characterized by a parameter of directionality. No other variable (e.g. vowel height) can be shown that would motivate the opacity of certain epenthetic vowels to RRH. Further, the role of directionality in RH is precisely analogous to the role we have shown it to play in BH (§2.2). PRH is both higher ranked and distinct from RRH.

Turkish also shows differing effects of PRH and RRH. All speakers apply RH regressively to an epenthetic vowel, modulo basic conditioning factors (the target vowel must be high) (Kaun 1999).

(55)	fülüt, *filüt	‘flute’
	buluz, bɨluz	‘blouse’

Some speakers (Kaun’s ‘Type F’ speakers) apply RRH when the trigger vowel is low and the target vowel is high.

(56)	Type F speakers apply RH	
	büröve, *biröve	‘diploma’
	bulok, *bɨlok	‘block’

Other speakers (Kaun’s ‘Type A’ speakers) do not apply RRH in this environment.

(57)	Type A speakers do not apply RRH	
	biröve, *büröve	‘diploma’
	bɨlok, *bulok	‘block’

Kaun notes that the height of the trigger vowel plays no role in normal Turkish PRH. Only in the unique context of RRH does trigger height emerge as a conditioning factor. This illustrates observation B above, that RRH may be subject to more stringent conditioning.

In Uighur epenthesis-driven harmony, speakers apply RH progressively (Hahn 1991).

(58) konturul, *kontirol — 'control'
gilobus, *gulobus — 'globe'

Speakers do not generally apply RH regressively, even if an appropriate trigger and target are present.

(59) tiropik, *turopik — 'tropic'

However, if a labial consonant is *also* present, regressive RH *does* apply:

(60) purotestant, *pirotestant — 'protestant'
purofessor, *pirofessor — 'professor'
furont — 'front' (war)

We note that a labial consonant alone is not sufficient to trigger rounding of the epenthetic vowel:

(61) pilan, *pulan — 'plan'
filim, *fulim — 'film'

Thus, in looking at just three Turkic languages, we find a range of possible ways in which progressive and regressive RH may pattern differently. We thus have a preliminary typology of directionality effects. Though it would be premature to construct a formal model based on such a small typology, we propose that a universal meta-ranking may be responsible for the fact that progressive RH appears to be generally more robust that regressive RH. No Turkic language, as far as we know, exhibits regressive RH to the exclusion of progressive RH. Based on ample evidence of this type from Altaic, we adopt a formal model (§4.0) in which alignment constraints include directionality as a parameter.

3.7 Dialect variation and RH typology of Altai-Sayan

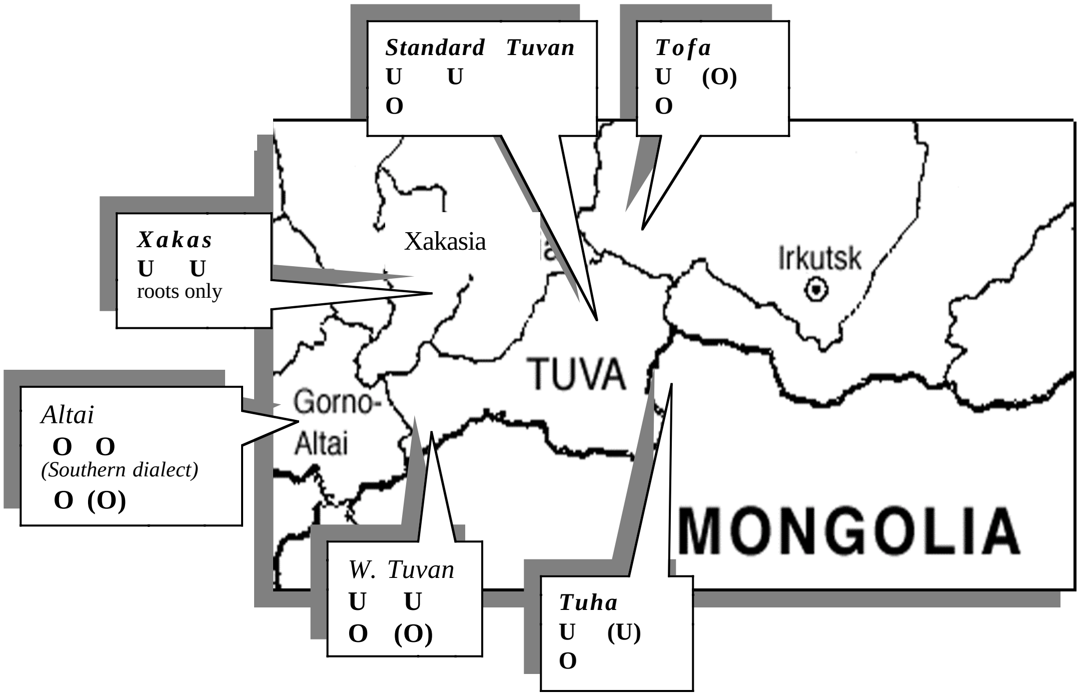
RH exhibits considerable dialect variation throughout the Tuvan language area. Typically, harmony is observed to apply in peripheral dialects in a manner that is less robust, more optional or more variable. However, most peripheral dialects are only very scantily documented. Seren (1993) notes that RH fails to target ‘certain’ high vowels in the Tuha (or Dukha) dialect of Tuvan (spoken in the Hövsgöl region of north-western Mongolia). [36] The relevant harmonic and disharmonic vowels are underlined in the data below, showing different RH in Standard Tuvan vs. Tuha.

(62)	Tuha Dialect	Standard Tuvan	
	qudɨs	qudus	‘mattress’
	on-ɨŋ	onuŋ	‘his, hers’
	ulus-tɨŋ	ulus-tuŋ	‘people’-GEN
	poq-sɨʊɨ̀s	poq-suʊus	‘trash’-1PL
	xol-dɨŋ	xol-duŋ	‘hand’-GEN
	ok-tɨɣ	ok-tuɣ	‘bullet’-ADJ

The breakdown of the harmony system, as attested in peripheral Tuvan dialects, can be accommodated in our constraint-based model of RH (§4.0).

A useful framework for Standard Tuvan RH is provided by RH systems attested in neighboring languages and peripheral dialects of Tuvan. Most of these systems are only minimally documented. On the map below, we depict RH systems of the Altai-Sayan area, to the extent that they are known, in order to provide a broader context for Tuvan. RH systems of the region differ minimally by their possible combinations of harmony triggers and targets (Anderson & Harrison 2000). These are represented schematically with RH triggers on the left side of the box and RH targets on the right side. Systems can also differ, though to a lesser extent, in whether they enforce RH optionality or obligatorily for a given trigger/target combination. Optionality of RH for a given trigger/target pair is indicated by placing the target vowel(s) in parentheses. Another parameter of variation in Altai-Sayan is that RH may apply robustly within roots but less systematically (or not at all) within suffixes (e.g. Xakas).

[36] Unfortunately, no front vowel tokens are included in the very small data set presented in Seren (1993). The Tuha (Dukha) dialect, spoken by a tiny, remote group of reindeer-herding pastoralists, is virtually undocumented.



Variation in rounding harmony along the Altai-Sayan language/dialect continuum[37]

The map shows schematic harmony systems across part of the Altai-Sayan language continuum. In western Tuvan, for example, harmony is consistently observed when trigger and target are both high [U U] *or* when the trigger is non-high and the target is high [O U]. By contrast, harmony is only optionally or less robustly observed when the trigger and target are both non-high [O (O)] *or* when the trigger is high and the target is non-high [U (O)].

[37] In the western Tuvan pattern, we note that a low vowel is only a possible target when a low rounded vowel [o] is the trigger. The high trigger vowel [u] does not induce RH in a low target vowel.

4.0 A formal model of RH

Our formal model will need to account for harmony in both static, distributional patterns and in surface alternations. Distributional patterns include positional neutralization of rounded vowels, static patterns of vowel co-occurrence (e.g. root-internal harmony), and the presence of disharmonic strings (e.g. in loanwords). Assimilation (harmony) patterns appear in suffixes, epenthetic vowels and reduplicants (ch. 5). We address the question of whether static, distributional patterns that appear to be trivially harmonic should have the same status in the grammar as active patterns of harmonic assimilation. This question is important to our claim that harmony remains an active phonological process even in the absence of surface alternations.

In our basic analysis, we assume that the feature [round] is privative. In the next chapter, when we introduce certain novel contexts for harmony, we will consider the possibility that [round] may also function as a binary feature (Chomsky and Halle 1968). We present evidence that [-round] plays an active role in harmony processes and at certain stages of historical development of the langauge. We then incorporate the feature into our revised model.

Rounding in Tuvan is positionally neutralized (Steriade 1995). Vowels of post-initial syllables in the native lexicon never contrast for rounding. Native words never contain instances of post-initial rounded vowels that are cannot be viewed as being the output of harmony. All Altaic languages that have RH also have, to a greater or lesser degree, positional neutralization of the feature [round]. There is clearly an implicational relationship between positional neutralization and RH. It is not so clear whether there is a causal relationship linking the two phenomena, or which one (if either) must precede the other in the development of RH systems. This remains an area for further research.

The Tungus language Evenki, for example, (Nedjalkov 1998, Grenoble & Bulatova 1998) enforces RH in which only low rounded vowels participate as triggers and targets. High rounded vowels—which neither trigger nor undergo RH—appear freely in post-initial syllables in Evenki. Tuvan thus differs from Evenki in that it enforces strict positional neutralization of the feature [round].

Analagous to our treatment of backness harmony, we follow Smolensky (1993), in adopting an alignment constraint as proposed in the theory of *Generalized Alignment* (McCarthy and Prince1993). This constraint is responsible for aligning the feature [+round] to the edge(s) of the relevant domain (here, the prosodic word) to any available targets.

ALIGN([RD], PRWD) *The autosegment [round] is aligned with the edge of the prosodic word*

As we argued above for backness harmony, directionality in the application of harmony indicates the presence of two distinct constraints.

ALIGN-L([RD], PRWD) *The autosegment [round] is aligned with the left edge of the prosodic word*

ALIGN-R([RD], PRWD) *The autosegment [round] is aligned with the right edge of the prosodic word*

Following Kaun (forthcoming), we adopt an additional constraint that militates against RH. This constraint bans the linking of a feature to a segment, unless that link is present in the input

DEP(LINK) *The output may contain no association line absent in the input.*

If DEP(LINK) outranks *both* alignment constraints there will be *no* RH in a given language. If it outranks only ALIGN[RD]/L, there will be only progressive harmony, and so on.

Thirdly, following Kaun (1995) we assume the presence of markedness constraints that ban the feature [round] in combination with other features. Feature combinations that are difficult for acoustic or articulatory reasons may be banned (cf. §3.2). In a full model of RH, these would ban both front rounded vowels and low rounded vowels. For Tuvan, we shall only need a constraint disallowing low rounded vowels. Low rounded vowels are difficult from an articulatory standpoint because they require mutually antagonistic gestures.

Increased sonority calls for jaw lowering but lip-rounding calls for jaw raising. We encode this phonetic fact as a formal constraint:

*RoLo *Low rounded vowels are banned.*

The effects of *RoLo are tempered by a constraint that enforces positional faithfulness (Beckman 1998). It protects the vowel of the initial syllable.

Faith-IO.V1 *Input must be faithful to output with respect to the vowel of the first syllable.*

These five constraints suffice to describe the basic RH pattern of Tuvan. The following ranking is apparent in the basic RH pattern:

Faith-IO.V1 >> *RoLo >> Align-R[Rd] >> Dep(Link) >> Align-L[Rd]

In all the tableau below, we represent suffixes as archiphonemic, e.g. specified only for height and unspecified for backness and rounding. We also assume underspecification for for all harmonic, post-initial vowels. As we have argued in this chapter (§2.8) regarding backness, and will argue in chapter 5 regarding rounding, harmony systems seem not only to allow underspecification of harmonic features, but also to require it.

We now apply the above constraint hierarchy in tableau form to both harmonic and disharmonic words. Tableau F illustrates how the basic pattern of RH targets to a high vowel in a suffix.

(63) Tableau F

/xöl + I/ | +R	Faith-IO.V1	Align-R[Rd]	Dep(Link)
xöl-i | +R		*!	
☞ xöl-ü | / +R			!
xel-i	*!		

'lake'-3

Tableau G illustrates the same basic pattern of alignment of round to a high vowel, this time within a root. We assume an underspecified (archiphonemic) representation for this vowel, as its values for backness and rounding are fully predictable.

(64) Tableau G

	/orIq/ | +R	Faith-IO.V1	*RoLo	Align-R[Rd]
	orɨq | +R		*	*!
☞	oruq | / +R		*	
	arɨq	*!		

‘road’

Tableau H illustrates that progressive RH does not target a low vowel, because the alignment constraint is outranked by *RoLo.

(65) Tableau H

	/oray/ | +R	Faith-IO.V1	*RoLo	Align-R[Rd]
	oroy | / +R		*! *	
☞	oray | +R		*	*
	aray	*!		

‘late’

Tableau I shows that Richness of the Base (Prince & Smolensky 1993) can be maintained. An unspecified input vowel will not surface as arbitrarily rounded in the absence of a RH trigger (candidate b). Nor will it align its rounding to the vowel of the first syllable (candidate c).

(66) Tableau I

/xar + I/	FAITH.IO/V1	*ROLO	ALIGN-L[RD]
☞ a. xarɨ			

b. xaru
      |
     +R
			*!

c. xoru
     \ |
     +R
	*!	*	

'snow'-3

To sum up the basic pattern, in the above tableaus we have shown the following. First, that rounding on the vowel of the initial syllable will be faithfully parsed in the output (Tableaus F, G, H). Second, that rounding will align from the initial syllable to a following high vowel (Tableaus F, G) but will not align to a low vowel (Tableau H). Third, that unmotivated rounding will not surface in post-initial syllables (Tableau I).

4.1 Disharmony

At this point, having accounted for the basic pattern, we will need to consider disharmonic forms. From a theoretical standpoint, it is highly desirable that both harmonic and disharmonic forms should fall out from a single constraint ranking. The alternative would be to posit distinct grammars, constraint hierarchies or co-phonologies for the two classes of words. We maintain that this approach is unnecessarily complicated and fails to reflect the fact that loanwords can be well-integrated into the native phonology.

Allowing native/harmonic words and foreign/disharmonic words to belong to the same grammar not only allows for economy in the grammar but is also plausible from a learnability perspective. Historically, loanwords may seep into a language gradually, and only some loanwords violate harmony. Disharmonic loanwords may start out by being fully assimilated at one stage of the language, but may end up resisting assimilation as has been the case in Tuvan. However, it does not follow from this that the best or only solution for speakers to accommodate loans is to posit a separate constraint hierarchy for them.

Loanwords are clearly divisible into two historical periods: an early period when all disharmonic forms were forced to become harmonic, and the present period when disharmony is tolerated. Early Mongolian loanwords were subjected to full assimilation to the native patterns of VH. Later, as Russian loanwords become widespread and bilingualism more common, disharmonic words were tolerated and not subjected to assimilation.

The difference between native, harmonic words, and disharmonic loanwords resides, we propose, not in the constraint hierarchy (or co-phonology) that governs them, but in their input representations. We thus treat loanwords as a practical problem for Lexicon Optimization (Prince & Smolensky 1993), not as a purely formal problem in constraint (re)ranking. For speakers, the task is as follows: how should a loanword that violates VH be mapped onto a plausible underlying representation? The answer suggested by Lexicon Optimization is that speakers make the 'most harmonic' mapping, assuming fully concrete representations with no abstract (underspecified) features. Archiphonemic Underspecification (Inkelas 1995) also predicts that disharmonic (and thus unpredictable) segments are fully specified.

What strategy do speakers adopt for dealing with an influx of disharmonic loanwords in a (formerly) fully harmonic language? They could assume that loanwords have the same kinds of UR as do harmonic words, but that they are governed by a different constraint hierarchy or a co-phonology (Inkelas, Orgun & Zoll 1996). Or, speakers could apply lexicon optimization differently to loanwords. For example, speakers might underspecifiy predictable segments, but fully specify unpredictable ones.

We propose that a relatively minor adjustment in the constraint ranking could allow the grammar instantly to accommodate disharmonic forms, without adopting a new constraint hierarchy (e.g. a co-phonology). The necessary adjustment would be to promote a constraint governing faithfulness to [-round]. Recall that we have up to now adopted the more orthodox view that rounding is privative. We propose that although [-round] is not needed in Tuvan phonology to account for native words, it is still potentially available as feature. To represent the unusual distribution of rounding in loanwords, speakers might need to employ [-round] to mark segments that resist RH.

Of course, when segments of this type first entered the language, speakers were unfaithful to disharmonic occurences of [-round] and rendered them as harmonic. As more words entered, speakers would have been confronted with more pervasive irregularities. In response, they could have promoted FAITH[-ROUND] to preserve instances where RH appeared to underapply to a particular segment.

There is a clear historical precedent for the role of [-round] as a distinctive feature in Turkic. In the earliest attested written form of Turkic, Orkhon Turkic (Tekin 1968), which is presumed to be close to reconstructed proto-Turkic (von Gabain 1938), rounding harmony was morphologically conditioned. There were affixes that systematically underwent RH, affixes that were always [+round] and affixes that were always [-round] and resisted RH. As Anderson (1998) notes, this ternary opposition among suffixal vowels for rounding would seem to require [-round] as a distinctive feature. This situation is mirrored in modern Uighur (Hahn 1991), where some suffix vowels are always rounded, others are subject to RH, and still others never undergo RH even though they apparently meet the structural environment to do so.

In both Orkhon-Turkic and contemporary Tuvan the active role of the feature [-round] (and, by extension, faithfulness to [-round]) seems necessary. However, these also seems to have been an intermediate stage—represented by pre-20th century Tuvan—where an active role for [-round] may have been unnecessary. This was due to the fact that the distribution of rounding was *fully* predictable at that time. It was during this time that early Mongolian loanwords came into the language. Segments that were unexpectedly round or unrounded in these loanwords were subjected to native RH patterns. The constraint mandating faithfulness to [-round] was ranked so low as to be inactive in the grammar.

We propose the active or inactive status of a constraint FAITH[-ROUND] correlates to historical stages of the language and the degree to which the distribution of rounding is fully or less than fully predictable at a given stage.

Language	Status of harmony	Status of [-ROUND]
Early (Orkhon) Turkic (8th century)	RH is morphologically conditioned, unpredictable in native suffixes	[-round] must be specified for invariant suffixal vowels
Tuvan (to 19th century)	RH is fully regular and predictable in roots and suffixes	[-round] has no use
Tuvan (20th century)	RH is unpredictable in some loanwords	[-round] must be specified for certain disharmonic root vowels

We noted that loanwords acquired prior to the 20th century were uniformly assimilated to the native RH patterns. This suggests that Faith[-round] was low ranked, and could not serve to protect disharmonic vowels in the input. Loanwords acquired from the mid-20th century up to the present resist assimilation to the native RH pattern. We suggest that disharmonic vowels marked with the feature [-round] are protected by a high-ranking FAITH[-ROUND], allowing them to remain disharmonic.

For the pre-20th century Mongolian loanwords, we shall want our model to account for four distinct effects of RH, as follows:

(i) High vowels in initial syllables are targeted by regressive RH (67)
(ii) High vowels are targeted by progressive RH (68)
(iii) Unmotivated high rounded vowels in post-initial syllables are derounded (69)
(iv) Low rounded vowels are eliminated in post-initial syllables but retained in initial syllables (70)

	Tuvan word	Mongolian source word	
(67)	čurumal	žirumal	'pattern'
	büdügüülük	bidegüülig	'primitive'
(68)	šügümčülel	šigümžilel	'criticism'
	mölčükču	mölžigči	'exploiter'
	temir	temür	'iron'
(69)	tergiilegči	tergüülegči	'director'
	ayɨɨl	ayuul	'danger'
	alžɨɨr	alčuur	'napkin'
(70)	oray	öröö	Mong. 'dusk', Tuv. 'late'

oʋ<u>aa</u>	ob<u>oo</u>	'shrine'

4.2 Licensing

The forms in 67, vis-à-vis the forms in 69, seem to show what could be interpreted as a classic licensing effect. In 67, when a post-initial rounded vowel has potential RH target(s) to its left, it can align its rounding. Rounding thus comes to be associated with the initial syllable, fulfilling the standard licensing requirement (Steriade 1995). If there is no eligible target in the initial syllable, as in 69, then the rounded vowel cannot become licensed by association to the initial syllable, and must be de-rounded.

Zoll (1997), although arguing for licensing on independent grounds, observes that licensing constraints can in many cases be adequately re-interpreted as 'context-specific' or *positional faithfulness* constraints (Steriade 1995a, 1995b, Beckman 1997, Casali 1996). We have already adopted a constraint that enforces positional faithfulness to the vowel of the first syllable. We now augment our model by adopting a more specific one that enforces enhanced faithfulness to rounding in the initial syllable.

FAITH.IO[RD](V1) *The autosegment [round] affiliated with the vowel of the initial syllable must be faithfully rendered in the output.*

By invoking positional faithfulness and the Left-Align constraint, we obviate the need for a separate licensing constraint. Following Zoll's suggestion, we will find that licensing becomes epiphenomenal when we adopt a standard alignment approach.

For the disharmonic loanwords, we adopt the following ranking representing the earlier stage of Tuvan when Mongolian words were assimilated:

FAITH.IO[RD](V1) >> ALIGN-L[RD] >> FAITH.IO[RD]

Tableau J shows that under this ranking, a post-initial rounded vowel will trigger rounding in the vowel of the initial syllable (b). Further, a vowel that violates left alignment (a) is not allowed to surface. Note that in this case of assimilation the vowel is also subjected to BH.

(71) Tableau J

/ži̲rumal / | +R	FAITH.IO[RD](V1)	ALIGN-L[RD]	FAITH.IO[RD]
a. či̲rumal | +R		*!	
☞ b. ču̲rumal \ | +R			
c. čirimal			*!

The same ranking—with the addition of the *RoLo constraint—accounts for data in 67. Here, a high rounded vowel in a post-initial syllable is prevented from aligning leftward due to the absence of a legitimate RH target (i.e. a high vowel). In this situation, as shown in Tableau K, the rounded vowel is eliminated in favor of an unrounded vowel. [38]

FAITH.IO(V1)[RD] >> *RoLo >> ALIGN-L[RD] >> FAITH.IO[RD]

(72) Tableau K

/aldʒuur/ | +R	FAITH.IO[RD](V1)	*RoLo	ALIGN-L[RD]	FAITH.IO[RD]
a. aldʒuur | +R			*!	
☞ b. aldʒɨɨr				*
c. oldʒuur \ | +R		*!		

A distinct licensing constraint proves unnecessary in accounting for the fact that rounded vowels only surface when they affiliate with the initial syllable. However, since leftward alignment accomplishes the same end result (i.e. affiliation of rounding to the initial syllable), this may be a purely formal distinction. A more thorough investigation of

[38] In tableau K, we assume that faithfulness to input height specification is high-ranking, thus ruling out a candidate such as *u̲ldʒuur.

loanwords in harmony languages is needed to discern a possible empirical distinction between leftward alignment and quasi-licensing effects.

We conclude that licensing offers no insight into the behavior of low rounded vowels in loanwords (70). Post-initial [o] can never surface, even when it would appear superficially to be licensed by association to an [o] in the initial syllable. As shown in tableau L, the standard ranking disallows rightward alignment of [round] to a low vowel and accounts for the adaptation of this loanword. Likewise, the ranking of positional faithfulness above markedness ensures that a candidate containing no rounded vowels (candidate d) will not surface. We can thus thus maintain Richness of the Base, allowing for the full range of possible inputs, and still select at the attested output.

Faith.IO(V1)[Rd] >> *RoLo >> Align-R[Rd] >> Faith.IO[Rd]

(73) Tableau L

	/oboo/ | / +R	Faith.IO[Rd](V1)	*RoLo	Align-R[Rd]	Faith.IO[Rd]
	a. oʋoo | / +R		**!		
☞	b. oʋaa | +R		*	*	*
	c. avoo	*!	*		*
	d. avaa	*!			

We stated above that the Tuvan data provide no clear evidence for licensing. Licensing is commonly invoked to prevent the occurrence of a feature in a post-initial syllable (Walker 1999). However, the ban on rounded vowels in post-initial syllables can be enforced using only constraints we have already introduced. Tableau M shows that an unmotivated rounded vowel in a post-initial syllable cannot surface.

(74) Tableau M

	/xep + I/	*RoLo	Align-L[Rd]
☞	xep-i		
	xöp-ü \ / +R	*!	
	xep-ü | +R		*!

‘clothing’-3

Moving on to contemporary Tuvan, we must account for the fact that loanwords may contain segments that fail to undergo rounding or may be unexpectedly rounded. This is the case in recent Russian loanwords such as *kino* ‘cinema’ or *arbus* ‘watermelon’. For contemporary Tuvan, we employ an additional constraint. Namely, Faith-IO[Rd] is replaced by Faith-IO[αRd], reflecting the potential role of the feature [-round] we introduced in §4.1. The ranking for contemporary Tuvan is as follows:

Faith.I-O/V1 >> *RoLo >> Faith-IO[αRd] >> Align-L/[Rd]

The presence of high-ranking Faith-IO[Rd] allows these segments to surface. Tableaus N and O show that a high rounded vowel that is unmotivated (i.e. not motivated by harmony, but motivated by the input) can surface in post-initial syllables in loanwords. The vowel does not, however, trigger RH in the vowel of the initial syllable.

(75) Tableau N

/arbus/ | +R	FAITH.I-O/V1	*ROLO	FAITH.I-O[αRD]	ALIGN-L/[RD]
☞ /arbus/ | +R				*
/orbus/ \ / +R	*!	*		
/arbɨs/			*!	

(76) Tableau O

/kino/ | +R	FAITH.I-O/V1	FAITH.I-O[αRD]	*ROLO
☞ kino | +R			
küno \ | +R	*!		
kina		*!	

We also noted that in contemporary Tuvan, initial vowels in loanwords are not targeted by RRH. We ensure this by ranking Align[Rd]/L below FAITH.I-O.V1 (Tableaus P, Q).

(77) Tableau P

/xirurg/ | +R	FAITH-IO(V1)	FAITH-IO[αRD]
☞ xirurg | +R		
xürurg \ | +R	*!	
xirirg		*!

It does not matter here if we regard the vowel of the initial syllable as being specified [-round] or not. If it is not specified, it will still be protected by positional faithfulness. If it is specified, it will be protected by faithfulness to [-round]. The vowel in the final syllable is crucially specified as [-round] denoting that it fails to undergo RH. Faithfulness to [-round] will ensure that it surfaces as unrounded, i.e., *žiguli*, not *žɨguli*.

(78) Tableau Q

	/ žiguli / | | +R –R	FAITH.I-O/V1	FAITH.I-O[αRD]
☞	žiguli | | +R -R		
	žüguli \ | | +R –R	*!	
	žigɨli		*!
	žigulü | / +R		*!

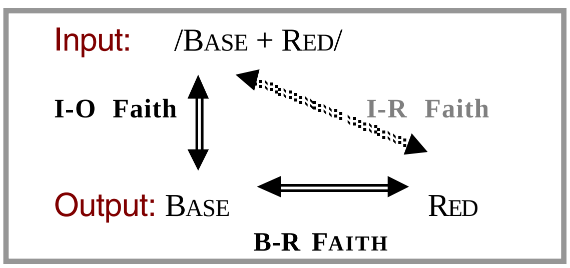
To sum up, we have developed a model of RH using standard constraints. We have also argued for both the possibility and necessity of underspecification. This allows us to generate both harmonic and disharmonic forms using a single constraint ranking, obviating the need for separate grammars or co-phonologies. In chapter 5, we present patterns of re-harmonization that demonstrate speakers' ability to actively manipulate harmonic and disharmonic strings. These patterns, we will claim, demonstrate both the psychological reality of harmony and the role of underspecification in harmony languages.

Chapter Five: Morphophonology of Reduplication

1.0 Introduction

In this chapter, we explore a wide range of reduplication phenomena, ranging from fully lexicalized to fully productive. Reduplication has been the subject of a number of empirical and theoretical studies (e.g. Wilbur 1973, Downing 1994, 1995, Alderete et al. 1996, Ross 1996, etc.). Our understanding of reduplication has been fundamentally revised by recent work in Optimality Theory. Correspondence Theory (McCarthy & Prince 1995) construes reduplication as governed by Correspondence relations between the base and reduplicant. The reduplicant does not ‘copy’ the base. Instead, the two stand in a relation of mutual correspondence. The so-called full model of Correspondence Theory includes each the following possible correspondence relations among the various compentents—input, base, and reduplicant—shown schematically here.

(94) Correspondence: The Full Model



The existence of I-O Faith and B-R Faith is uncontroversial. The possible effects of I-R Faith, by contrast, have not been so thoroughly explored. Some analyses (e.g. Kager 1998) adopt only the ‘basic model’ in which there is no I-R Faith (McCarthy & Prince 1995:11). The absence of IR-Faith renders the reduplicant free of the burden of being faithful to the input. Reduplicants—unencumbered by any faithfulness obligations—may thus emerge as closer to the ideal phonotactics of the language. The emergence of phonotactic constraints in reduplicants that are not visible in bases gives rise to a class of effects known as TETU—the emergence of the unmarked.

In this chapter, we explore Tuvan reduplication phenomena, ranging from fully lexicalized (non-productive) to fully productive. The most productive type—Full Reduplication—appears to give rise to a ranking paradox. Namely, all vowel segments in the reduplicant may be unfaithful tot he base in order to conform to vowel harmony. They thus behave as if free from any obligation to be faithful to either the base or the input. This can be handled under the basic model which lacks any correspondence relation between input and reduplicant. But disharmonic vowel segments pose a problematic exception here. They must remain faithful to the base (and, by extension the input) and may not conform to harmony. They thus behave as if constrained by either the base or the input, or both. Reduplicants thus arise both from the nature of the correspondence relations (represented by the arrows in (1)) and from the nature of the input (e.g. specified or underspecified).

2.0 Tuvan reduplication

Tuvan has a range of reduplication phenomena. We begin with the simplest, least productive ones. We then introduce the more complex types, along with a discussions of their consequences for current modesl of reduplication and correspondence.

2.1 Emphatic reduplication

Diminutive 'S'reduplication 'copies' the entire base and replaces the initial C with [s] in the reduplicant. For bases that are vowel-initial, an onset [s] is added in the reduplicant.

(95)	*aar*	'heavy'
	aar-saar	'heavy'-DIM
	uuruk-suuruk	'simultaneously'

This reduplication is marginally productive, the following forms were elicited from several speakers but were not commonly known.

(96)	*pelek*	'gift'
	pelek-selek	'gift'-DIM
	porbaq	'round'
	porbaq-sarbaq	'round'-DIM,

A second type, emphatic reduplication takes the initial $(C_1)V_1(C)$ of the base and adds [p] in the reduplicant to yield a prefix $[(C_1)V_1p]$. This process is common across Altaic. It tends to be semantically limited to certain color terms and descriptive modifiers (see **2.2** below).

(97)	*qɨzɨl*	'red'	*qɨp-qɨzɨl*	'completely red'
	uzun	'long'	*up-uzun*	'very long'
	türgen	'quickly'	*tüp-türgen*	'very quickly'
	čiŋge	'thin'	*čip-čiŋge*	'very thin'.

(98) Mongolian(Baatar Tsend, p.c.)

	xap-xar	'black'
	nov- nogoon	'green'
	tsav-tsagaan	'white'
	uv-ulaan	'red'
	shuv-shur	'yellow'
	bov-bor	'grey'

This type of reduplication (usually with fixed segment [p]) is reported in Amenian (Vaux 1998), Finnish (Heli Harrikari, p.c.), Uzbek (Farrukh Suvankulov, p.c.), Kazak (Askar Abilov, p.c.) and elsewhere in Turkic. The most complex case is that of Turkish, where the fixed element [p] shows dissimilatory effects, depending on the phonological content of the base. For discussion of the Turkish pattern we refer the reader to Yu (1999) and Kelepir (1999)

In Tuvan, unlike in any other Turkic language, CVp- reduplication has become fully productive for verbs as an aspectual marker. This reduplication may apply to various tenses to add an aspectual connotation of rapid or intense action (6, 7). It may also indicate unexpected or sudden action (8). In other instances, it simply adds a meaning of strong assertion to the verb (9-10)

(99) *xap-xalaan men*
REDUP-'drive'-PAST.I 1
'I was driving fast'

(100) *xap-xaladɨm*

REDUP-'drive'-PAST.II-1
'I drove fast'

(101) *köp-kördüm*
REDUP-'see'-PAST.I I
'I saw!' (unexpectedly/suddenly)

(102) *köp-körbeen men*
REDUP-'see'-NEG-PAST.I 1
'I have never seen at all'

(103) *inekti sap-saar men*
'cow'-ACC REDUP-'milk'-P/F 1
'I will definitely milk the cow'

(104) *inekti sap-saɣbas men*
'cow'-ACC REDUP-'milk'-NEG-P/F 1
'I will definitely <u>not</u> milk the cow'

2.2 Full Reduplication

Tuvan has a highly productive morphological process of expressive reduplication. Speakers apply this process to produce novel allomorphs of virtually any noun or verb plus any attendant morphology.

Full reduplication takes the entire base (consisting of virtually any lexeme plus its attendant morphology) while replacing the vowel of the initial syllable with a pre-specified vowel, usually either [a] or [u] (hereafter referred to as "replacement vowels"). In its simplest form, the rule changes [a] to [u], and changes all other vowels to [a].

Monosyllabic words are reduplicated as follows (12) . Polysyllablic forms will be introduced later (17):

(105)

base	base + reduplicant	
nom	*nom-nam*	'book'
er	*er-ar*	'male'
eet	*eet-aat*	'river delta'
is	*is-as*	'footprint'
ög	*ög-ag*	'yurt'
süt	*süt-sat*	'milk'
qɨs	*qɨs-qas*	'girl'
xol	*xol-xal*	'hand'
at	*at-ut*	'name'
aar	*aar-uur*	'heavy'

As we will show, successful production of allomorphs—both harmonic and disharmonic—requires speakers to be able to manipulate underlying representations. These underlying representations, we argue, must contain certain segments that are unspecified for the harmonic features [+/- back] and [+round].

The semantics of reduplication signals intentional vagueness or a special, casual informal, jocular register. For example, the reduplication of *horse* might appear in a sentence meaning ‘You haven't seen *anything like a horse* have you?’ or ‘There were *horses and stuff* around the yurt’. Semantically, this type of reduplication contributes a sense of vagueness and/or jocularity to almost any lexeme in the language. Its use is restricted, however, to a subset of speakers and a special register, such that not all native speakers know how to do reduplication. Nonetheless, the process is sufficiently transparent that we were able to teach it to both adults and young children in a matter of seconds. Speakers who had just learned the rule were able to produce novel reduplicants and their output matched that of speakers who use reduplication regularly.

This simplified data set (12) shows that all vowels except [a] are replaced by [a], while [a] is replaced by [u]. Regional variations allow greater flexibility in the output. For many speakers, [e] and [ö] in the base may be changed to either [a] or [u] in the reduplicant: *er* ‘male’ becomes *er-ar* or *er-ur.* The choice of the replacement vowel varies widely in different (regional) versions of the rule, as we show in §3.0

Speakers may also differ as to whether they treat the [base+reduplicant] as a single morphological word for the purpose of affixation, or as two distinct words.

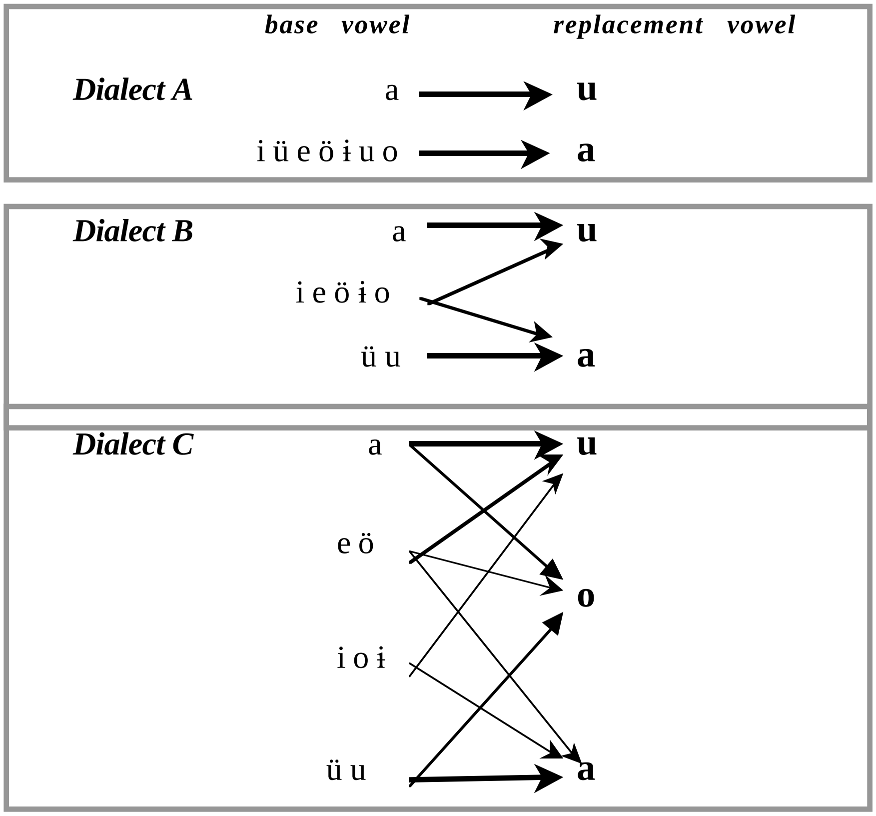
(106)	*idik* ‘boot’,	*idiim-adɨɨm*	‘boot’-1-REDUP-1
		idik-adɨɨm	‘boot’-REDUP-1

Output forms generated by this morphological process reveal the operation of phonological constraints that are not apparent in the regular lexicon and morphology of Tuvan. For example, the status of markedness constraints governing individual vowel segments is revealed is revealed by speakers' variation in the choice of the replacement vowel in reduplicants. Further, reduplication yields output forms in which rounding harmony fails to apply even though all conditions for its application are present. We argue that a sub-set of the constraints responsible for rounding harmony emerges in the novel context of reduplication. Crucially, these constraints exert no observable effects anywhere else in the language. We claim that such constraints, which may remain dormant or inactive in the standard lexicon and morphology, can emerge and apply robustly in more marginal areas of the language such as reduplicants and loanwords. We shall refer to such constraints as being ‘inactive’, or ‘in reserve’. Our investigation of Tuvan reduplication provides evidence for a rounding harmony constraint which is attested cross-linguistically, but which remains inactive in Tuvan and surfaces only under special conditions.

3.0 Variability in output

The true complexity of Tuvan reduplication becomes apparent in the many dialect and individual versions of the rule. The author documented reduplication patterns for thirty-five individual speakers, all of whom already knew and used the reduplication rule. Three types of data were collected: (i) spontaneous instances of reduplication from speech; (ii) elicited reduplicated forms; (iii) speakers’ judgments of well-formedness of reduplicants proposed by the author or produced by other speakers. Speakers’ production of reduplicants may be classified into one of three basic patterns (which correspond to dialect groups) according to the number of possible replacement vowels they employ (either two or three), and the degree of flexibility in mapping inputs to output. The htree patterns correspond fairly simply to dialects of the language. In the schema below, bold lines denote more robust mapping patterns, while plain and dotted lines denote less robust ones.

(107) Mapping of base (input) vowel to replacement (output) vowel



Dialect A (spoken in Kyzyl, the capital city of Tuva), has only two replacement vowels, [a] and [u], and invariable mapping of inputs to outputs. Speakers of this dialect do not accept any alternative outputs as well-formed. Dialect B (spoken in central Tuva) also has two output vowels, but speakers show flexibility in mapping mid vowels and high unrounded vowels to either [a] or [u]. Speakers usually have an opinion about whether [a] or [u] is better, with an overall preference for [a].

Dialect C, (of the Süt-xöl region of Tuva), includes [o] as a possible replacement vowel, and shows even greater flexibility in mapping input to output. Dialect A is maximally simple and therefore less interesting, while dialect C has not yet been fully documented and is left for further research. The remainder of this paper will thus be primarily concerned with dialect B, which includes a number of regional variants of spoken Tuvan. For a speaker of dialect B or C, more than one well-formed reduplicant may be derived from a single base. Under (15) we see some of the various reduplicants produced by a single speaker of dialect B.

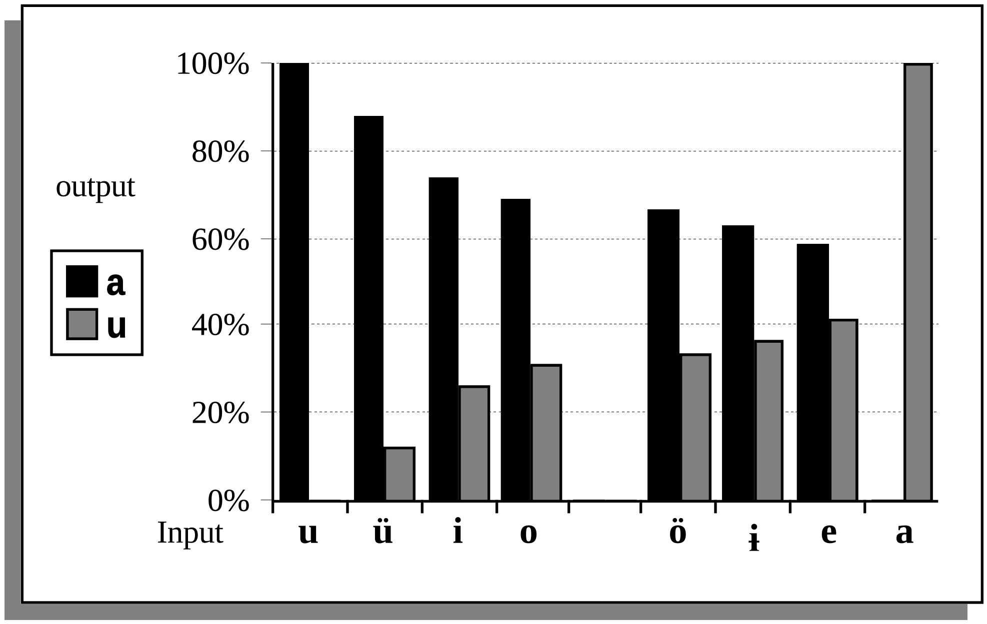
(108)

base	base+reduplicant I	base+reduplicant II	gloss
xol	xol-xul	xol-xal	‘hand’
ök	ök-ak	ök-uk	‘button’
is	is-as	is-us	‘footprint’
er	er-ar	er-ur	‘male’
öörenip	öörenip-aaranɨp	öörenip-uuranɨp	‘study’-CV
ekkep	ekkep-ukkap	ekkep-akkap	‘bring’-CV
eʒim	eʒim-uʒum	eʒim-aʒɨm	‘friend’-1
mal	mal-mul	mal-mol	'livestock'

The relation of input (base) vowels to output (replacement) vowels are shown in the bar chart (16), which represents the composite output of ten speakers of dialect B (421 tokens). The horizontal axis shows the base vowels which serve as input to the rule. The vertical bars show outputs (replacement vowels) for each given input. Dark colored bars show the percentage of tokens for which [a] emerged as replacement vowel, and gray colored bars show how often [u] emerged.

When the base contained [u], the reduplicant had [a] one hundred percent of the time. The same proportion held true for input [a]. All other input vowels varied in their output. Speakers showed an overall preference for [a], but [u] also appeared robustly as an alternative. I argue that the ratio of [a] to [u] in the output reflects speakers’ uncertainty as to the relative markedness of these vowels. It also reflects the fact that the reduplication rule requires at minimum two replacement vowels, so both [a] and [u] are valid choices, even though [u] is more marked.

(109) Choice of replacement vowel [a] or [u] for ten speakers (dialect B)



3.1 Variability and markedness

The reduplication process as we have formulated it does not specify the quality of the replacement vowel. We argue that this falls out from general markedness principles. A markedness hierarchy among vowels might best account for the proportional occurrence of [a] and [u]. Cross-linguistically, [a] is the least marked vowel, and thus it is not surprising that [a] emerges most frequently as output. But a single replacement vowel (e.g., [a]) is insufficient to satisfy the minimal condition of reduplication, namely that the first vowel of the reduplicant always be different from the first vowel of the base. At least two replacement vowels are needed for this. Since [a] is the least marked in the inventory, it makes an ideal replacement vowel. The next best choice appears to be [u], given that the output is always a back vowel. Since [o] and [ɨ] are both highly marked, the proposed markedness hierarchy [a] > [u] > [o] > [ɨ] for back vowels is at least a plausible one.

Establishing a markedness hierarchy for the vowels of a language is a difficult theoretical and empirical problem. In proposing a tentative hierarchy for Tuvan vowels, we considered the following factors: (i) frequency of segments across languages, (ii) frequency of segments in Tuvan, (iii) alternations in productive morphology, (iv) positional neutralization, (v) acoustic values (central vs. peripheral), and (vi) the facts of reduplication. The exact realization of markedness principles in Tuvan is left for further research. For now, we represent Tuvan markedness as a single constraint, *MARKED, requiring speakers to avoid marked segments.

4.0 A Correspondence Model of Full Reduplication

Except for the replacement vowel, the above reduplicants are identical to the bases from which they derive (we omit minor consonant alternations herein). We assume the base and reduplicant to be related by a faithfulness constraint (McCarthy and Prince 1995). This ensures that—modulo other constraints—they will resemble each other as closely as possible:

IDENT-BR *The reduplicant must be identical to the base.*

Apart from their replacement vowels, monosyllabic reduplicants are identical to bases. The reduplication rule must introduce a higher-ranked, 'anti-faithfulness' constraint (Kelepir 1999). This constraint, which must be specific to Tuvan, not universal, is akin to other language-specific constraints that govern word games or speech disguises. It dictates that the first vowel of the base and replacement vowel must differ from one another:

*IDENT-BR(V1) *The replacement vowel must differ from the vowel of the initial syllable of the base.*

These two constraints alone account for the data presented so far. But the reduplicant becomes considerably less faithful when the base is polysyllabic:

(110)	base	base + reduplicant	gloss
	idik	idik-adɨk	‘boot(s)’
	inek	inek-anak	‘cow’
	ulu	ulu-alɨ	‘dragon’
	arɨ	arɨ-uru	‘bee’
	nomdʒuur	nomdʒuur-namdʒɨɨr	‘read’
	oktaan	oktaan-uktaan	‘throw’-PAST
	uduur	uduur-adɨɨr	‘sleep’-FUT

4.1 Backness harmony in reduplication

Backness harmony operates on the natural vowel classes defined by the feature [+/-back], and requires all the vowels in a word to agree in backness. Although backness harmony is a pervasive pattern in Tuvan, violations are fairly well tolerated in the form of disharmonic loanwords from Russian:

(111) a. *ispan* ‘Spanish’
b. *tiraž* ‘print run’
c. *kazino* ‘casino’
d. *ali* ‘Ali’ (given name)
e. *(ɨ)rakeeta* ‘rocket’

As noted in chapter 4, many loanwords happen to be accidentally harmonic, even though they may violate other aspects of Tuvan phonotactics. There is thus no necessary correlation between the perceived “foreignness” of a word and its harmonic status. The following examples are harmonic, but they contain non-native phonemes [f], onsets [pl], or clusters [bl].

(112) a. *fiidik* ‘video cassette’
b. *pleyer* ‘video cassette player’
c. *yaablaq* ‘apple’

Although the reduplicant is generally faithful to the base, faithfulness may be overruled by the requirements of harmony. Since the replacement vowel is always [a] or [u], it follows that all vowels in the reduplicant must also be [+back]. If all vowels of the base are already [+back] they do not undergo harmony shift in the reduplicant (a-b). But if the base is [-back], the replacement vowel triggers harmony shift of every post-initial vowel in the reduplicant (c-f).

(9)

	base	base + reduplicant	gloss
a.	arat	arat-urat	'peasant'
b.	olar	olar-ular	'they'
c.	idik	idik-adɨk	'boot'
d.	inek	inek-anak	'cow'
e.	idiktig	idiktig-adɨktɨg	'boot'-ADJ
f.	xülümzüreer	xülümzüreer ~ xalɨmzɨraar	'smile'-P/F

All front vowels in the reduplicant will shift to become [+back], thus being unfaithful to the base. Clearly, it is more important for a reduplicant to be harmonic than to be faithful to the base.

An analogous harmony shift occurs on rounded vowels in post-initial syllables of the reduplicant. In (10 a-e) reduplication replaces the initial round vowel of the base with an unrounded [a] in the reduplicant. Any following rounded vowels must subsequently de-round, since they no longer possess a preceding trigger. In (f-h), reduplication introduces a rounded vowel [u] in the first syllable, and any adjacent high vowels subsequently become round, In (i), we see that if the potential target is non-high, the introduction of a RH trigger has no effect.

(10)

	base	base + reduplicant	gloss
a.	ulu	ulu-alɨ	'dragon'
b.	nomʒuur	nomʒuur-namʒɨɨr	'read'
c.	börü	börü-barɨ	'leaf'
d.	bürü	bürü-barɨ	'wolf'
e.	oqtaar	oqtaar-aqtaar	'throw'-FUT
f.	arɨzɨ	arɨzɨ-uruzu	'bee'-3
g.	eri	eri-uru	'man'-3
h.	ivi	ivi-uvu	'deer'
i.	arat	arat-urat	'herdsman'

Thus, with the application of backness and rounding harmony, a reduplicant can end up having all its vowels differ from their corresponding vowels in the base. The constraints that enforce harmony must therefore outrank base-reduplicant faithfulness.

ALIGN[BK], ALIGN[RD] >> IDENT.B-R

All the base-reduplicant pairs we have seen so far consist in native, harmonic forms. But many loanwords contain vowels that are disharmonic with respect to rounding, backness, or both. When speakers apply reduplication to disharmonic loanwords the resulting forms provide us with important insights into the nature of their underlying representations.

4.2 Disharmony and reduplication

Loanwords may contain vowels that are disharmonic with respect to rounding harmony. For example, high rounded vowels [ü] [u] are disharmonic if they appear in post-initial syllables but are not explicitly motivated by rounding harmony (e.g. not preceded by a RH trigger).

(113)	base	base + reduplicant	
a.	*dersu*	*dersu-darsɨ*	'hunter'
b.	*aquula*	*aquula-uquula*	'shark'

Unmotivated high rounded vowels are routinely derounded in reduplicants (21a). These vowels remain round, however, if they acquire motivation by having a rounded trigger introduced in the initial syllable of the reduplicant (21b).

Low rounded vowels [ö] [o] are disharmonic in any post-initial syllable, because they are banned from this position in Tuvan words.

(114)	base	base + reduplicant	
a.	*kurort*	*kurort-karart*	'spa'
b.	*qartoošqa*	*qartoošqa-qurtaašqa*	'potato'

Most speakers eliminate unmotivated low rounded vowels in reduplicants. For example, the post-initial [o] in *kurort* 'spa' violates the markedness constraint agains low rounded vowels. But the reduplicant, *karart* satisfies this constraint by eliminating the offending [o]. The reduplication process itself does not require [o] in *kurort* to change to [a]. Rather, the reduplicant is merely subjected to all active constraints, such that it emerges phonotactically more well formed than the base from which it derives. It has been well documented that reduplicants, for example in Tagalog (Ross 1996), are often more conservative in obeying native phonotactic constraints than are the bases from which they derive.

The behavior of round disharmonic forms (20, 21) does not force us to modify our proposed ranking. Harmony constraints still appear to outrank faithfulness constraints.

Summing up the data presented thus far, we observe that harmony constraints appear to be undominated in the lexicon, the regular morphology, and reduplicative morphology. Second, we observe that faithfulness relations between base and reduplicant always lose out to harmony. In other words, no matter what the base, harmony constraints will derive a harmonic reduplicant from it.

We now look at words that are disharmonic with respect to [+/- back]. These would seem to represent the simplest case, yet they force us to reevaluate our proposed constraint ranking and to reconsider the nature of underlying representations.

Back disharmonic segments, unlike round disharmonic segments, never undergo harmony shift in the process of reduplication. Disharmonic [-back] vowels unexpectedly remain disharmonic:

(115)	<u>base</u>	<u>base + reduplicant</u>	<u>gloss</u>
a.	*ali*	*ali-uli* **ali-ulɨ, *ali-ulu,* **ali-ulü*	'Ali'
b.	*mašina*	*mašina-mušina*	'car'
c.	*noyabrina*	*noyabrina-nuyabrina*	(name)
d.	*aybek*	*aybek-uybek*	'Aibek'
e.	*radiyo*	*radiyo-rudiya* *radiyo- *rud<u>ü</u>ya*	'radio'

4.3 Finnish Kontti Kieli

A similar result has been documented in a Finnish word game called *kontti kieli* 'knapsack language' (Campbell 1986, Vago 1988). This game takes any word, adds the word *kontti* 'knapsack' after the word, then exchanges the initial syllables and re-harmonizes both words:

(116) *mitä* → *mitä kontti* → *kota mintti* 'what'

Native Finnish words that are [-back] invariably undergo back harmony shift when reduplicated (vowels targeted by backness harmony are underlined):

(117)	base	base + reduplicant	
a.	*käy*	*kou käntti*	'he visits'
b.	*sikiö*	*kokio sintti*	'embryo'
c.	*mitä*	*kota mintti*	'what'

Loanwords that are back disharmonic pattern differently: instead of conforming to backness harmony, they retain their [-back] disharmonic segments:

(118)	base	base + reduplicant	
a.	*manööveri*	*konööveri mantti*	'manouver'
b.	*jonglööri*	*konglööri jontti*	'juggler'
c.	*hydrosfääri*	*kodrosfääri hyntti*	'hydrosphere'
d.	*klorofylli*	*korofylli klontti*	'chlorophyll'

Vago (1986:198) concludes the following regarding disharmonic segments:

> "...the exceptional nature of disharmony with respect to the basic restrictions imposed by vowel harmony is preserved. That is, the non-initial front harmonic vowel of a disharmonic word does not harmonize to the preposed back harmonic vowel of /kontti/."

In the Finnish *kontti kieli,* speakers thus replace the initial vowel then re-harmonize the remaining vowels, with the exception of the neutral [e] and [i]. As in Tuvan, speakers re-harmonize harmonic words, but consistently fail to do so with disharmonic words. Note that root vowels of Finnish never alternate except within the special context of *kontti kieli*. These results show that in the novel context of a language-external word game, speakers treat harmonic and disharmonic segments differently. Campbell (1986) used this to argue for the psychological reality of a rule of vowel harmony within Finnish, even in the face of numerous surface counter-examples to harmony. From our perspective, the Finnish data are comparable with Tuvan and may be analyzed in the same way, with the same implications.

4.3 A Ranking Paradox

Returning to Tuvan, we recall that back disharmonic segments also fail to undergo harmony shift under reduplication. Thus, for back disharmonic words, base-reduplicant faithfulness would appear to outrank the constraints responsible for back harmony.

IDENT-BR >> ALIGN[BK]

This ranking is exactly the reverse of that we proposed above. If we rely on output constraints alone, we are driven to posit two distinct constraint rankings, one for harmonic forms and another for disharmonic ones. This is tantamount to positing two distinct grammars. Native harmonic words, accidentally harmonic loanwords and round disharmonic loanwords must all belong to grammar A, where harmony outranks faithfulness. Back disharmonic loanwords must belong to grammar B, where the opposite holds true. The proliferation of grammars is clearly an undesirable outcome.

If we reject the notion of separate grammars (constraint rankings) for harmonic and disharmonic forms, then we need a formal theory that can account for the disparate patterning of disharmonic vowels. Such a grammar should generate all and only the attested forms with a single, uniform set of constraints.

We adopted in chapter 4 a formal model that calls for the features [+/-back] and [+round] to be aligned with the word edge. This model *can* account for harmony patterns in native lexical forms. It also readily accounts for the elimination of rounded vowels in reduplicants: e.g. *kino* > *kino-kan***a** (not **kino-kano*). But the alignment model cannot readily accommodate the presence of many disharmonic loanwords that are well integrated into the Tuvan lexicon. Nor does it account for the failure of disharmonic front vowels to undergo harmony shift. It does not therefore give us any principled way to capture the disparate patterning of back and round disharmonic vowels under reduplication.

We also proposed in chapter 4 that underlying representations are unspecified for harmonic features [+/- back] and [+round]. By adopting an underspecified UR for harmonic forms, but a fully specified UR for disharmonic forms, we can generate all attested cases with a single constraint ranking. We claim that the different patterning of harmonic and disharmonic words can be adequately captured by an underspecification analysis. We conclude that underspecification is not only a useful analytic tool for linguists, but that it is reflected in speakers' production of allomorphic forms such as Tuvan reduplicants. In the next section, we examine a proposed model of undespecification and argue that the Tuvan facts call for a different distribution of underspecification than that predicted by the model.

5.0 Underspecification and lexicon optimization

In this section, we show that current interpretations of Lexicon Optimization (Prince and Smolensky 1993), in particular that of Archiphonemic Underspecification (Inkelas 1995), incorrectly predict the distribution of underspecification in lexical entries. We present cases from three vowel harmony languages in which speakers treat harmonic and disharmonic roots differently under reduplication. The assumption of full specification entails a ranking paradox, which can be resolved if underspecification is admitted in certain contexts not predicted by the principles of Lexicon Optimization. We point the way towards an expanded model of Lexicon Optimization that would both allow for and predict such cases of underspecification.

Transparency to spreading (or assimilation), susceptibility to spreading, failure to initiate spreading and various other special behaviors prompted analysts working within a derivational model of phonology to hypothesize that certain featural specifications are absent for the relevant segments throughout at least a portion of the phonological derivation. Contrastive Underspecification (Steriade 1987) and Radical Underspecification (Archangeli 1984) were the dominant formal models designed to predict in a principled manner the incidence of underspecification. Contrastive Underspecification theory posited non-contrastiveness as the criterion for potential underspecification. A feature value might be missing from underlying representations if it failed to serve a contrastive function for the segment class in question. Under Radical Underspecification, the commitment to a redundancy-free lexicon had the consequence of eliminating all predictable feature values from underlying representations.

<u>Within Optimality Theory (Prince and Smolensky 1993), Richness of the Base specifically rules out the systematic exclusion of featural specifications from input representations. The input space is assumed to be infinite, thus unrestricted. The lexicon, by contrast, is assumed to be finite. A learner's construction of lexical representations is guided by Lexicon Optimization (Prince and Smolensky 1993), which heavily favors fully specified inputs. It is assumed that a speaker will choose the most harmonic input-to-output mapping. Outputs can be mapped to fully specified inputs without the accrual of gratuitous faithfulness violations. Given two competing input forms, one fully specified and one partially specified, the fully specified alternative will be preferred, all else being equal. The model nevertheless leaves room for the possibility that partially underspecified lexical entries will on occasion be posited.</u>

Archiphonemic Underspecification (Inkelas 1995) seeks to predict when underspecified inputs will in fact be deployed. It demonstrates that the principles of Lexicon Optimization dictate that predictable feature values will be underspecified only when they enter into surface alternations. In a backness harmony language like Turkish, for instance, only those vowels that are involved in alternations will be underspecified for backness in lexical entries. In the words given in (1), all post-initial vowels agree in backness with the initial vowel. However, it is only the suffix vowels that both alternate and have a predictable value for backness. Archiphonemic Underspecification predicts that only they will be underspecified for backness, as indicated in (2):

(119) <u>Turkish</u>

a. kilim-im 'rug-my' b. kɨlɨc-ɨm 'sword-my'

Assumed Lexical Representations

a. kilim-im (i, i: -B -B) b. kɨlɨc-ɨm (ɨ, ɨ: +B+B)

This model essentially claims that harmony targets only suffix vowels in Turkish, because it is only for suffix vowels that a backness value lacking in the input representation is introduced in the corresponding output. Root vowels, whether harmonic or not, will be fully specified (and presumably identical) in input and output representations. Root vowels thus cannot be thought of as undergoing harmony at all. Clements and Sezer (1982, p. 226), motivated by entirely different theoretical considerations, take a similar position, stating that "…the burden of proof is on the linguist who wishes to demonstrate that roots [in Turkish] are governed by vowel harmony at all."

In this chaopter we present data that challenge the position advanced by Clements and Sezer and the predictions of Archiphonemic Underspecification. The patterns that we discuss indicate that harmony is indeed active within roots. Thus, we claim that the incidence of underspecification is not that which is predicted by Archiphonemic Underspecification. In particular, we introduce patterns of "re-harmonization" which demonstrate that speakers underspecify vowels not only in affixes but also in roots. The data involve the distinct treatment of harmonic and disharmonic roots in three harmony languages: Tuvan, Finnish and Turkish.

5.2 Harmony, Reduplication and Re-harmonization

Empirical evidence from three vowel harmony languages—Finnish, Tuvan and Turkish—poses a problem for the predictions of Archiphonemic Underspecification. In these languages, we argue, predictable segments must be underspecified even though they are non-alternating. We further argue that if non-alternating segments are underspecified then alternation is not an adequate diagnostic of underspecification. Our claim that speakers underspecify non-alternating segments rests on the different patterning of harmonic vs. disharmonic vowels in the novel context of reduplication and re-harmonization.

As we have shown, harmonic and disharmonic vowels pattern differently under reduplication. Our claims about underspecification and harmony rest on this difference in patterning. As shown above, there is clearly a faithfulness relation between the base and reduplicant: except for the replacement vowel, the two are always identical. However, polysyllabic roots (along with any suffixal morphology) may be considerably less faithful to their base. In polysyllabic forms, post-initial vowels generally agree in backness with the replacement vowel. Note that since the replacement vowel is either [a] or [u], it is always [+back]. To achieve this agreement, speakers subject post-initial vowels to re-harmonization. To call attention to potential re-harmonization effects, we underline all post-initial vowels in reduplicants.

(120) Full reduplication of polysyllabic bases with re-harmonization

	a.	idik	idik-adɨ̱k (*adi̱k)	‘boot’
	b.	fiidik	fiidik-faadɨ̱k (*faadi̱k)	‘video cassette’
	c.	teve	teve-tava̱ (*tave̱)	‘camel’
	d.	tevelerim	tevelerim-tava̱la̱rɨ̱m (*tave̱le̱ri̱m)	‘camel’-PL-1SG

Disharmonic segments, whether native or borrowed, fail to undergo re-harmonization, and remain disharmonic.

(121) Full reduplication of polysyllabic bases with no re-harmonization

	a.	maʃina	maʃina-muʃi̱na (*muʃɨ̱na, *muʃu̱na)	‘car’
	b.	ajbek	ajbek-ujbe̱k (*ujba̱k)	‘Aibek’
	c.	ʒiguli	ʒiguli-ʒagu̱li̱ (*ʒagu̱lɨ̱,*ʒagu̱lu̱)	‘Zhiguli’ (car name)
	d.	aal=ʒe	aal=ʒe-uul=ʒe̱ (*uul=ʒa̱)	‘yurt’=ALL

Adopting an Optimality Theoretic framework, we model Tuvan harmony and reduplication with the following constraints:

(122) Constraints

Replacement vowel	*IDENT-BR(V1)
Input-Reduplicant Identity	**IDENT-I/R**
Harmony	**ALIGN[BK]**
Base-Reduplicant Identity	**IDENT-B/R**

To begin, let's consider an input form such as *idik* 'boot'. In accord with the predictions of Lexicon Optimization, we assume a fully specified representation for this harmonic form, even though the backness value of the second vowel is predictable on the basis of the backness of the initial vowel. For *idik*, as for all harmonic words, the following constraint ranking allows us to capture the pattern of re-harmonization:

(123) Ranking

*IDENT-BR(V1) >> ALIGN[BK] >> IDENT-I/R, IDENT-B/R

To show how these constraints interact, consider tableau 3, where we have included all but the Tuvan-specific morphological constraint.

(124) Tableau A: Harmonic form with fully specified input.

/idik, RED/ | / -B	ALIGN[BK]	IDENT-IR	IDENT-BR
a. idik-adik	*!	*	*
☞ b. idik-adɨk		**	**

The attested output form will only emerge if the harmony constraint ALIGN[BK] outranks both faithfulness constraints. This is due to the fact that the unattested candidate (a) is more faithful to the input, both in terms of IDENT-I/R and IDENT-B/R. Re-harmonization under reduplication provides ample evidence for the undominated status of the harmony constraint.

<u>But for a disharmonic form with a fully specified input, the ranking shown in (15) selects the wrong candidate, as shown in tableau 4.</u>

(125) Tableau B: Disharmonic form with fully specified input.

	/ajbek, RED/ | | -B +B	ALIGN[BK]	IDENT-IR	IDENT-BR
☹	a. ajbek-ujbek	*!	*	*
💣	b. ajbek-ujbak		**	**

The attested disharmonic candidate (a) should win, but does not. For candidate (a) to win, the harmony constraint must rank below at least one of the faithfulness constraints. We thus propose an alternative ranking in which ALIGN[BK] ranks below input-reduplicant faithfulness. Again, we assume a fully specified input.

(126) Tableau C: Disharmonic form with fully specified input (new ranking)

	/ajbek, RED/ | | +B -B	IDENT-IR	ALIGN[BK]	IDENT-BR
☞	a. ajbek-ujbek	*	*	*
	b. ajbek-ujbak	**!		**

This ranking correctly selects the attested disharmonic form. We are thus faced with a ranking paradox by which harmonic and disharmonic sequences seem to require separate constraint rankings (i.e., separate grammars). This apparent paradox may be resolved if we allow harmonic words to be represented by underspecified inputs (contra the predictions of Lexicon Optimization and Archiphonemic Underspecification).

(127) Tableau D: Harmonic form with underspecified input

	/idik, RED/ | -B	IDENT-IR	ALIGN[BK]	IDENT-BR
	a. idik-adik	*	*!	*
☞	b. idik-adɨk	*		**

Since disharmonic forms do not undergo re-harmonization (tableau 5), while harmonic forms do (tableau 6), we have assumed the former are fully specified for the harmonic feature and the latter are partially underspecified. Underspecification thus has a desirable result in that it obviates the need to posit separate constraint rankings for various subsets of the lexicon. Partial underspecification has the following consequence, illustrated in tableau 6. A violation is incurred only for the vowel of the initial syllable in both candidates. The underspecified non-initial vowels undergo "cost-free" re-harmonization (19b). In tableau 5, on the other hand, both vowels are fully specified such that an output which obeys harmony (18b), does so at a cost, namely the violation of input-reduplicant faithfulness. The same disparity in the treatment of harmonic and disharmonic vowels is found in Finnish (section §2.2) and Turkish (section §2.3).

We have claimed that speakers of vowel harmony languages such as Tuvan and Finnish underspecify predictable but non-alternating segments within roots. This claim entails that vowel harmony must be an active process even in cases where it appears to be little more than a static, phonotactic pattern. In Tuvan and Finnish, we have presented evidence that the quality of root vowels is established by means of alignment, except in the cases of disharmonic roots. We turn now to Turkish, where we extend our argument that harmony (rendered formally as alignment) actively determines feature values of all harmonic segments, even those that never enter into surface alternations. We show that Turkish vowel co-occurrence patterns within roots are not merely static, phonotactic patterns but active harmonic processes that have psychological reality and accessibility for speakers (contra Clements and Sezer 1982).

5.3 Turkish Root Harmony

Turkish lacks a Tuvan or Finnish-style process that would subject roots to novel alternations. Turkish root vowels thus never alternate in any context. As a pilot study, we taught a Tuvan-style reduplication rule to a speaker of Turkish. The speaker was instructed to replace the initial vowel of a set of Turkish words with [a] or [u]. He could then make any other changes—or no changes at all—to the reduplicant. The resulting form was to be a nonsense word that "sounds like a good Turkish word." The speaker produced multiple reduplicants for most bases; he then selected the best-sounding ones. Preliminary results of our pilot study show effects quite similar to those of Tuvan. The Turkish speaker showed a clear preference for re-harmonizing harmonic words (a, b), but not disharmonic words (c, d).

(128) Turkish novel reduplication

a.	kibrit	kibrit-kabrɨ̲t (*kabri̲t)	'match'
b.	bütün	bütün-batɨ̲n (*batü̲n, *bati̲n)	'whole'
c.	mali	mali-muli̲ (*mulɨ̲ *mulü̲)	'Mali'
d.	butik	butik-bati̲k (*batɨ̲k)	'boutique'

Our Turkish speaker, like Tuvan or Finnish speakers, apparently manipulated underlyingly unspecified segments by re-harmonizing them in this novel, reduplicative context. The fuller study of Turkish speakers' harmony preferences under reduplication will, we predict, provide additional empirical evidence that speakers of harmony languages underspecify predictable but non-alternating segments. They do so, we suspect, in response to an observed harmonic pattern attested in most (but by no means all) lexemes. The presence of a surface pattern of vowel co-occurrence in roots, in combination with regular alternations in suffixes is sufficient to drive speakers to posit a general system of vowel harmony that obtains across both roots and affixes.

5.4 Towards a model of Pattern Responsive Lexicon Optimization

Speakers can thus arrive at underspecification by various means: the first would be as a result of predictable alternations, as in Turkish suffixes. Archiphonemic Underspecification correctly predicts that such cases will give rise to underspecification. A second way speakers can arrive at underspecification is essentially by analogy. Specifically, the presence of alternations in one part of the grammar (e.g., suffixes) in combination with segmental predictability in roots leads speakers to underspecify non-alternating root segments. We thus argue that it is an overall *pattern* of alternations, not morpheme-specific alternations, that triggers underspecification in lexical entries. A third type of situation that could trigger underspecification might be one that lacks any alternations but that shows a predictable pattern of surface distribution. This would be the case in a language having strict vowel co-occurrence patterns like those of Turkish and Tuvan, but lacking any affixal morphology.

We do not yet know whether a surface pattern alone, in the absence of any alternations, would provide “clear guidance” to speakers (see Yip 1998), inducing them to posit an active harmony process and to underspecify segments for the harmonic feature. We suggest that such a *pattern-responsive* system should be considered as a logical possibility and that further research should look for such patterns in phenomena such as tone, stress, reduplication and various types of harmony. An important goal is to construct a model of Lexicon Optimization that would allow us to characterize precisely the circumstances under which speakers will posit abstract lexical entries. An adequate theory of lexicon optimization should anticipate speakers’ propensity to underspecify in response to surface-true patterns.

6.0 The emergence of reserve constraints in reduplication

We stated in chapter 4 that RH obtains robustly without regard to height, backness or length of trigger vowels. In the data below, we illustrate that the length of the trigger vowel plays no role in the application of RH.

(129)

trigger	target	word	gloss
long	short	ooʒum	'slow(ly)'
long	short	boo-zu	'gun'-3
short	short	udup	'sleep'-CV
short	short	ulu-zu	'dragon'-3
short	long	uruu	'daughter'-3
short	long	idii-uduu	'boot'-3-REDUP-3
long	long	tooruu	'sunflower.seed'-3

Length plays no apparent conditioning role in the application of Tuvan RH. This is not the case in the Tungus languages Bayinna Oroch (Li 1996) and Evenki (Nedjalkov 1997), or in Daur Mongolian (Wu 1996) where long vowels fail to trigger RH, although they consistently undergo it. Shortness of trigger vowels thus seems to be a factor favoring harmony. We know of no languages where long vowels trigger RH but short vowels do not. The relationship between trigger goodness and length is implicational: if long vowels are good RH triggers in a given language then short vowels are too. The reverse is not true. A constraint that spreads rounding from short vowels would thus seem to be universally ranked above one that spreads from long vowels. We propose the following constraints to account for quantity-sensitive RH as attested in Evenki, Bayinna Orochen and Daur:

(130) Quantity-sensitive RH constraints

ALIGN[RD]/V *Align the feature [+round] from a short vowel*
ALIGN[RD] *Align the feature [+round]*

In the normal application of Tuvan RH, there is no evidence for the distinct status or separability of these two constraints. They pattern together (essentially, as a single constraint) in all areas of the lexicon and affixal morphology. RH is triggered by all rounded vowels, regradless of their length. Since the set of constraints is thought to be universal, however, we should conclude that these constraints do indeed remain distinct in Tuvan, but simply never pattern distinctly. It is uniquely in the context of reduplication, as we will show, that these constraints may fail to apply in an identical manner.

We conclude that the constraints responsible for rounding spread are virtually undominated in Tuvan. Nowhere in the lexicon, borrowed lexicon, or morphology does RH fail to apply when the triggering conditions are met. Given the robust and pervasive nature of Tuvan RH, it is somewhat surprising that the output of reduplication provides unique examples of forms where rounding harmony fails to apply even though conditions for it appear to be present.

(131)

base	base+reduplicant I	base+reduplicant II	
azɨ	*azɨ-uzu*	* *azɨ-uzɨ*	'if'
arɨ	*arɨ-uru*	* *arɨ-urɨ*	'bee'-3
aazɨ	? *aazɨ-uuzu*	*aazɨ-uuzɨ*	'mouth'-3
aarɨ	? *aarɨ-uuru*	*aarɨ-uurɨ*	'burden'-3
	RH APPLIES	**RH FAILS**	

For all speakers, *azɨ* 'if' is reduplicated as *azɨ-uzu*, with obligatory rounding harmony. But the nearly homophonous form *aazɨ* 'mouth'-3 is, for many speakers, reduplicated as *aazɨ-uuzɨ* without rounding harmony. Underapplication of RH can occur only if the triggering vowel is long. This effect is quite robust in reduplicants, but does not surface anywhere else in Tuvan.

As noted in chapter 4, rounding harmony when viewed cross-linguistically is frequently seen to be conditioned by various phonological features of the target and trigger (e.g. height, backness). We model the conditioned nature of harmony systems using constraints that directly encode various features of target and trigger.

We proposed in chapter 4 that rounding harmony may be perceptually motivated as a strategy for dealing with a difficult contrast such as unenhanced rounding. We will now capitalize on that notion by proposing that it may be more crucial to spread the feature [round] from a short vowel than from a long one. A long vowel has greater perceptual salience, and may be better able to support a difficult contrast. This hypothesis provides the motivation for the two RH constraints proposed in (10).

Our formal model of Tuvan reduplication and RH relies on five constraints:

*IDENT-BR(V1)	*The replacemen vowel must differ from the base vowel*
IDENT-BR	*The reduplicant should be identical to the base*
ALIGN[RD]/V	*Align [+round] from a short vowel*
ALIGN[RD]	*Align [+round]*
*MARKED	*Avoid marked vowels*

The full range of attested Tuvan reduplication patterns can be readily generated by the appropriate rankings of these constraints. For example, speakers showed three different patterns of RH from a long trigger:

(132)

base		base+reduplicant		
speaker 1	(RH)	*taaqpɨ*	*taaqpɨ- tuuqpu̲*	'tobacco'
speaker 2	(no RH)	*taaqpɨ*	*taaqpɨ-tuuqpɨ*	
speaker 3	(optional RH)	*taaqpɨ*	*taaqpɨ- tuuqpu̲*	
			~ *taaqpɨ-tuuqpɨ*	

Speaker 1 represented a minority of those we surveyed. He consistently applied RH regardless of the length of the trigger. We model this behavior with the following sub-ranking:

ALIGN[RD]/V, ALIGN[RD] >> IDENT-BR

(133) Tableau E

	aazɨ 'mouth'-3	*MARKED	*IDENT-BR(V1)	ALIGN[RD]/V	ALIGN[RD]	IDENT-BR
	aazɨ-aazɨ		* !			
	aazɨ-uuzɨ				* !	*
☞	aazɨ-**uuzu**					* *
	aazɨ-aaza		* !			*
	aazɨ-iizɨ	*!				* *

Most speakers we surveyed were like speaker two, in that they always applied RH when the trigger was short, but consistently failed to apply RH when the trigger was long. The effect we describe is thus quite robust, though absent from the regular lexicon and morphology of the language. Speaker two *never aligned* rounding from a long trigger vowel in reduplicants. In his grammar, IDENT-BR is interspersed between the two alignment constraints.

(134) Tableau F

	aazɨ 'mouth'-3	*MARKED	*IDENT-BR(V1)	ALIGN[RD]/V	IDENT-BR	ALIGN[RD]
	aazɨ-aazɨ		* !			
☞	aazɨ-**uuzɨ**				*	*
	aazɨ-uuzu				* *!	
	aazɨ-aaza		* !		* *	
	aazɨ-ɨɨzɨ	* !			* *	

Finally, speaker three *optionally aligned* rounding from a long trigger vowel in reduplicants. We found such speakers consistently able to alternate between forms with RH and forms without RH, thus employing both of the above rankings.

In the case of reduplicants with short harmony triggers, all speakers consistently applied RH, and judged proposed forms without RH to be ill-formed. All speakers *always aligned* rounding from short trigger vowels in reduplicants. In all grammars, ALIGN[RD]/V outranks IDENT-BR.

(135) Tableau G

azɨ 'if'	*MARKED	*IDENT-BR(V1)	ALIGN[RD]/V	ALIGN[RD]	FAITH-BR
azɨ-azɨ		* !			
azɨ-uzɨ			* !		*
☞ azɨ-**uzu**					* *
azɨ-aza		* !			*
azɨ-ozɨ	* !		*		*
azɨ-ozu	* !				* *
azɨ-ɨzɨ	* !				*

6.1 Discussion

Tuvan speakers have no evidence from the regular lexicon and morphology of their language that RH may fail to obtain when the potential trigger is long. Such cases are unattested in the standard language. Yet in novel contexts, when called upon to apply the reduplication rule, speakers produce forms in which RH fails to apply precisely when the trigger is long. The underapplication of RH when all conditions for it are present is clearly an anomaly in Tuvan. This output reflects "on-line" decisions made by speakers about the ranking of RH constraints relative to faithfulness constraints which are otherwise inactive.

We propose that phonotactic constraints that may be entirely dormant or have no active role in the native lexicon or morphology of a language can nonetheless surface and play an active role in more marginal or novel areas of a language such as word games, reduplicants, and loanwords. When these hidden phonological constraints do come into play, speakers may exhibit some uncertainty about their ranking, or may assign them a novel ranking. The assigned ranking can have the effect of demoting other constraint(s) that would otherwise be dominant.

In the case discussed herein, a sizeable number of speakers appear to be able to demote one of the RH constraints, causing RH to fail in some cases. In demoting this constraint, they effectively break up the family of RH constraints, which everywhere else in the language pattern as a uniform block. This raises theoretical issues about the behavior and (non)uniform patterning of constraint families.

The generally dormant or inactive status of IDENT-BR and the fact that it becomes active only under reduplication, may account for speakers' uncertainty about its ranking. Clearly, no consensus has yet been reached in the speech community about where this constraint should rank relative to RH constraints. Dialectal and individual differences in ranking persist. Furthermore, many speakers seem content to employ at least two alternative rankings on different occasions, and do not seem to be in rush to assign a ranking. This raises important issues about the stability of variable constraint rankings and speakers' tolerance for uncertainty.

Summary

In this thesis we have aimed for *descriptive adequacy* by documenting the phonology, morphology and (to a lesser extent) syntax of Tuvan. In chapters 2 through 5, we have also aimed for *explanatory adequacy* by adapting formal models of grammar to account for our data. In this section we review the major findings of our research on Tuvan and point to some outstanding problems which merit further investigation.

In our study of the pitch accent system in chapter two, we proposed that the *length* of the pitch accented vowels was event-driven. Namely, the accented vowels have to be longer in order to accommodate the pitch contour. This hypothesis has found support in new data we obtained after the completion of the chapter. Where Tuvan has (long) pitch accented vowels, we found that the closely related language Tuha employs short nasalized vowels. Similarly, we found that the Tsengel Tuvan dialect employs short pharyngealized vowels. A fuller acoustic study of peripheral Tuvan dialects is needed to ascertain just how the pitch accent system breaks down and develops into these other laryngeal/nasal qualities. Also, as noted in chapter 2, we hope to undertake a comparative historical study to shed light on the genesis of Tuvan pitch accent.

The study of hiatus resolution and velar deletion in chapter 3 may shed light on the phonology-semantics interface. We have shown, for example, that a phonological process is blocked when it would interfere with recovery of ‘small’ morphological elements, thus leading to a loss of semantic content. We adopted a somewhat vague criterion of ‘recoverability’, arguing that this can trigger blocking of otherwise robust phonological processes. Clearly, we will need a better formal model of ‘recoverability’, formulated in terms of specific constraints and their interaction with phonological constraints.

Our study of vowel harmony in chapter 4 attempts to include both dialect and individual variation. We note that such ‘messy’ data has typically been swept under the rug in the interest of reducing harmony to an elegant, uniform phenomenon. Yet, we claim, it is precisely this kind of data that will prove crucial in understanding the nature of harmony.

More work needs to be done to organize such data and investigate patterns that emerge from inter- and intra-speaker variation. Although we adopted the notion of ‘optionality’ to explain some kinds of variation in harmony, we have not yet developed the necessary formal model to explain it. For example, does optionality require crucially unranked constraints or simply flexible rankings? If so, how do speakers arrive at such non-canonical ranking patterns?

A second important result of our study of harmony is the evidence for directionality. Even though a given harmony system may show no directional assymetries at all, speakers keep in *reserve* the distinct constraints that enforce harmony directionally. Under the right circumstances (e.g. epenthesis-driven harmony) these may emerge and show effects. We argue that while these effects are analogous to TETU effects, they are fundamentally different in that the emergent constraint is an alignment constraint, not a markedness constraint. This raises the possibility that harmony constaints remain in reserve for all languages, and could emerge and show effects even in a non-harmony language.

A third important result of our harmony study emerges in the special context of reduplication and re-harmonization patterns presented in chapter 5. We argue that these provide strong evidence for underspecification in input representations. Moreover, we argue that this underspecification is driven not by alternations (since it seems to arise in non-alternating roots as well) but by a pervasive pattern of vowel co-occurence. In future research we will continue to develop a model of ‘pattern-responsive’ grammar, to determine what kinds of patterns can guide speakers to posit underspecification.

Finally, we turn to a potentially fruitful area of research that we were not able to include in this thesis. Tuvan employs a number of enclitics. These bound or semi-bound elements lack the morpho-phonological properties of suffixes and are never stressed. Some exhibit partially assimilatory behavior, for example they may undergo vowel harmony or voicing assimilation of the onset consonant. Others may lack any phonological interaction with the immediately preceding host. As such, they may constitute cases of apparent underapplication of expected phonological processes. For example, if an enclitic has a lateral onset [l] that consonant may fail to undergo an expected geminate dissimilation process (whereby [ll] becomes [ld]) when it follows a word ending in [l].

We limit ourselves here to a few examples that illustrate the range of behavior of two enclitic elements. The emphatic clitic /-LA/ and an equative clitic /-LA/ are homonymous. They differ, however, in their placement within reduplicated forms. The emphatic clitic always attaches at the right edge of its host word.

(1) *kɨp-kɨzɨl=***la**
[REDUP-'red'-EMPH]
'very red' (emphatic)

The equative clitic also attaches at the right, but if there is already a prefixed reduplicant, the equative intervenes between the reduplicant and its base, in what is likely a prosodically determined position (e.g. to fill out a foot).

(2) *kɨp=***la***-kɨzɨl*
[REDUP- EQU -'red']
'very red-like'

Ambiguity may arise between the two homophonous forms:

(3) *xavan=***na**
['pig'-EQU] or ['pig'-EMPH]
'pig-like' or 'pig' (emphatic)

When both appear in a single domain, the Equative obeys prosodic conditioning and appears in the first foot.

(4) *kɨp=***la***-kɨzɨl=***la**
[REDUP- EQU -'red'-EMPH]
'**really** very red-**like**'

(5) *ol idik.tiɣ men up=***la***-uzun=***na** *kiži tömeyleštim*
['that'] ['boot'-ADJ] [REDUP=EQU-'long'] ['person'] ['resemble'-RECIP.PAST.I.1]
'With those boots, I resembled a **very** tall-**like** person'

In addition to the two examples given here, a number of Tuvan lexemes and affixes appear to be developing into clitics. Tuvan, along with its cognate language Xakas (Anderson 1998), thus provides a range of clitic-like behaviors which may prove challenging to divide up into formal categories. These and at least a dozen other putative Tuvan enclitics will require further study to assess their morpho-phonological properties.

References

Adam, Lucien. 1874. De l'harmonie des voyelles dans les langues ouralo-altaïques. Paris: Maisonneuve & C[ie].

Anderson, Gregory D.S. 1998. *Xakas*. Languages of the World/Materials, vol. 251. München: LINCOM Europa.

_____ 2000. Turkic languages of Altai-Sayan: A report from the field. Talk presented at LSA annual meeting, January 2000.

Anderson, Stephen R. 1980. Problems and perspectives in the description of vowel harmony. In R. Vago (ed.) *Issues in vowel harmony*. (pp. 1-48). Amsterdam: John Benjamins.

_____ 1992. *A-morphous Morphology*. Cambridge: University Press.

Archangeli, Diana. 1984. Underspecification in Yawelmani phonology and morphology. Doctoral dissertation, MIT, Cambridge, Mass.

Babushkin, G. F. 1959a. O znachenii i sintakticheskix funkcijax deeprichastij v tuvinskom yazyke. In *Uchenye Zapiski tuvinskogo nauchno-issledovatel'skogo instituta yazyka, literatury i istorii [TNIIYaLI] VII*: 93-104. Kyzyl: TNIIYaLA.

_____ 1959b. Predvaritel'nye dannye po zapadnomu dialektu tuvinskogo yazyka. In *Uchenye Zapiski TNIIYaLI VII*: 211-16. Kyzyl: TNIIYaLA.

_____ 1960. O strukture pridatochnyx predlozhenij v tuvinskom yazyke. In *Uchenye Zapiski TNIIYaLI VIII*: 127-38. Kyzyl: TNIIYaLA.

_____ n.d. (1960?). Morfologicheskaja struktura xakasskix dialektov v sravnenii s dialektami Altaja-Sayanskoj grupy tyurkskix yazykov. Unpublished-ms. Abakan: XakNIIYaLA.

Baskakov, N. A. 1958. *Altajskij yazyk*. Moscow: AN SSSR.

_____ 1965. Forma glagola na <u>-chyx/chik</u>, <u>-chu/chü</u> v xakasskom, tuvinskom i kirgizskom yazykax. In *Voprosy tyurkologia.* Tashkent: "N" UzSSR, pp. 5-10.

_____ 1978. Forma proshedego vremeni glagola na -chyx/chu v tyurkskix yazykax i ee proisxozhdenie. In <u>Sovetskaya Tiurkologiya</u> Sovetskaya tiurkologiya #4, pp. 3-8.

Beckman, Jill. 1998. Positional Faithfulness. Ph.D. Dissertation. Amherst: Univ. of Mass.

Bergel'son, M. A. and A. A. Kibrik. 1987a-b. Sistema perekluchenіya referentsii v tuvinskom yazyke. In Sovetskaya tiurkologiya #2, 16-32 (=1987a); #4, 30-45 (=1987b).

Bicheldei, K. A. 1985. Osobennosti zvukovogo stroja yugo-vostochnogo dialekta tuvinskogo jazyka. In *Fonetika sibirskix jazykov*, pp. 25-36. Novosibirsk: AN SSSR SO.

_____ 1980b. Faringalizatsija glasnykh v tuvinskom jazyke. In *Zvukovoj stroj sibriskikh jazykov*. Novosibirsk: AN SSSR, SO IIFF, 39-65.

_____ 1986. Chetyrye aspekta glasnykh tuvinskogo jazyka v potoke rechi. In *Fonetika jazykov Sibiri i sopredel'nyx regionov*, pp. 33-7. Novosibirsk: "Nauka".

_____ 1980a. Faringalizatsija glasnykh v tuvinskikh zaimstvovanijakh iz mongol'skogo i russkogo jazykov. In *Jazyki Narodov Sibiri*, pp. 248-53. Kemerovo: KGU.

_____ 1993. Tìva bižikti bolgaš šɨn bižilgeni sayzɨradɨr talazɨ=bile sanaldar. In *Voprosy tuvinskogo yazykoznaniya*, pp. 40-49. Kyzyl: TNIIYaLA.

_____ 1985. Osobennosti zvukovogo stroya yugo-vostochnogo dialekta tuvinskogo yazyka. In *Fonetika sibirskix yazykov*, pp. 25-36. Novosibirsk: AN SSSR SO.

_____ 1980b. Faringalizatsiya glasnyx v tuvinskom iazyke. In *Zvukovoj stroj sibriskix yazykov*. Novosibirsk: AN SSSR SO IIFF, 39-65.

_____ 1980c. Nazalizovannye glasnye sovremennogo tuvinskogo yazyka. In *Issledovaniya zvukovyx sistem sibirskix yazykov*, pp. 29-34. Novosibirsk: AN SSSR SO.

_____ 1984. Nazalizovannye glasnye v tuvinskom yazyke i ix sootvetstviya v tyurkskix yazykax. In *Issledovaniya zvukovyx sistem yazykov Sibiri*, pp. 35-9. Novosibirsk:AN SSSR SO.

Boguroditskij, A. 1953. *Vvedenie v tatarskoe yazykoznanie.* Kazan’: Tatgosizdat.

Budaeva, A. N. 1972. Russkie slova so stecheniem soglasnykh w proiznoshenii evenkov. In *Voprosy yazyka i fol’klora narodov severa.* Yakutsk: Sib. Otdel. AN.

Bulatova, N. Ya. & Lenore Grenoble. 1999. Evenki. München: LINCOM Europa, 1999.

Browman, Catherine & Louis Goldstein 1989. Articulatory gestures as phonological units. *Phonology* 6:201-251.

_____ 1992a. Articulatory phonology: An overview. *Phonetica* 49: 155-180.

Campbell, Lyle. 1986. Testing phonology in the field. In *Experimental phonology,* ed. John J. Ohala and Jeri J. Jaeger, 163-186. Orlando, Fl: Academic Press.

Casali, Roderic. 1996. Resolving Hiatus. Ph.D. Dissertation. Los Angeles: UCLA.

Castrén, M. A. 1857. *Versuch einer koibalischen und karagassischen Sprachlehre*. Edited by A. Schiefner. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Chadamba, Z. B. 1974. *Todzhinskij dialekt tuvinskogo jazyka*. [The Tozhan dialect of Tuvan] Kyzyl: Tuvknigizdat.

Cheremisina, M. A. [ed.] 1995. *Yazyki korennyx narodov Sibira.* Novosibirsk: RAN.

Cheremisina, M. A. 1995. Osnovnye tipy analiticheskix konstrukcij skazuemogo v tyurkskix yazykax yuzhnoj Sibira. In Cheremisina [ed.], pp. 3-21.

Cheremisina, M. A. et al. [eds.] 1984. *Predikativnoe sklonenie prichastij v altajskix yazykax*. Novosibirsk: "Nauka".

Chomsky, Noam & Morris Halle. (1961) The Sound pattern of English. New York: Harper & Row.

Clements, George N. and Engin Sezer. (1982). Vowel and consonant disharmony in Turkish. In Harry van der Hulst and Norval Smith (eds.) *Structure of phonological representations* (Vol. II), pp. 213-255. Dordrecht, Holland: Foris Publications.

Cole, Jennifer & Charles Kisseberth. 1994a. An optimal domains theory of harmony. Cog.Sci. Technical Report UIUC-BI-CS-94-02, University of Illinois, Urbana-Champaign. ROA-22.

Cole, Jennifer & Charles Kisseberth. 1994b. Paradoxical strength conditions in harmony systems. ROA-48.

Comrie, Bernard. 1981. *The languages of the Soviet Union.* London: Routledge

Darɨmaa, O. K-C. 1976. *Tɨva ulustuŋ tɨvɨzɨktarɨ.* [Tuvan people's riddles.] *Kɨzɨl. TNUCh.*

Delger-Ool, A. K. 1960. Osnovnye tipy glagolnyx oborotov v tuvinskom predlozhenia. In *Uchenye Zapiski TNIIYaLI VIII*: 118-26. Kyzyl: TNIIYaLA.

Doerfer, G. 1973. Bemerkungen zu den sojonischen Anlautsklusilen. In *Ural-Altaische Jahrbücher* XLV: 254-60.

Donegan, Patricia. 19. On the Natural Phonology of Vowels. New York: Garland.

Efremov, N. N. 1984. *Polipredikativnye predlozheniya vremeni v yakutskom yazyke*. Moscow: "Nauka".

Elert, Claes-Christian. 1964. Phonologic studies of quantity in Swedish.

Forsyth, J. 1992. *A History of the Peoples of Siberia*. Cambridge: Cambridge University Press.

von Gabain, Annemarie. 1941. *Alttürkische Grammatik.* Leipzig.

Gvozdanovic, Jadranka. 1980. Tone and accent in standard Serbo-Croatian. Vienna: Austrian Academy of Sciences.

Hahn, Reinhard F. 1991. Spoken Uyghur. Seattle: University of Washington Press.

Harrison, K. David. 1999. Tuvan reduplication and harmony. In *Proceedings of the Berkeley Linguistics Society 30: Parasession on Turkic, Dravidian and Caucasian Linguistics.*

_____ 2000a. Harmony and underspecification in Tuvan. In MIT Working Papers in Linguistics. Proceedings of the Eleventh Student Conference in Linguistics SCIL 11.

_____ 2000b. When the optimal output isn't: Velar deletion in Siberian Turkic. Talk presented at LSA annual meeting, Chicago, January 2000.

Harrison, K. David & Abigial Kaun. 2000. Pattern Responsive Lexicon Optomization. In *Proceeings of NELS 30*. Also in Rutgers Optimality Archive: ROA-392-06100

Howard, Irwin. 1970. Kasem nominals and the ordering of phonological rules. In *University of Hawaii Working Papers in Linguistics.* Vol. 2, no. 3.

Inkelas, Sharon. 1995. The consequences of optimization for underspecification. In *Proceedings of the Twenty-fifth Northeastern Linguistics Society*. ed. Eugene Buckley and Sabine Iatridou. Amherst, Mass.: GLSA.

_____ 1998. Phonotactic blocking through structural immunity. Handout, Lexicon in Focus.

Iskhakov, F. G. & A. A. Pal'mbakh 1961 *Grammatika tuvinskogo jazyka*. [A grammar of Tuvan] Moscow: Nauka.

Itô, Junko, Armin Mester & Jaye Padgett. (1995) NC: licensing and underspecification in optimality theory. ROA-38-1194.

Katanov, N. F. 1903. *Opyt issledovaniya uriangxajskogo yazyka s ukazaniem glavnejshix rodstvennyx otnoshenij ego k drugim yazykam tyurkskogo kornya*. Kazan'.

Kaun, Abigail R. (1995). *The typology of rounding harmony: An optimality theoretic approach.* Ph.D. dissertation. UCLA.

_____ (1999a) Turkish Epenthesis Harmony. In *Proceedings of the Berkeley Linguistics Society 30: Parasession on Turkic, Dravidian and Caucasian Linguistics.*

_____ (1999b). Inactive constraints: "discharged" or "in reserve"?. Talk presentet at AT&T Laboratories colloquium.

_____ Forthcoming. Rounding harmony. In *The phonetic basis of phonology,* ed. Bruce Hayes, Robert Kirchner and Donca Steriade. Cambridge: Cambridge University Press.

Kelepir, M. (1999) Emphatic non-identical reduplication in Turkish. handout, LSA Annual Meeting.

Kenin-Lopsan, Mongush Boraevich. 1993. *Shamanic songs and myths of Tuva*. Budapeszt.

Kiparsky, Paul (1993) Blocking in non-derived environments. In S. Hargus and E. Kaisse (eds.), Studies in Lexical Phonology. San Diego: Academic Press.

Kirchner, Robert. 1993. Turkish vowel harmony and disharmony: An optimality theoretic account. ROA-4.

Kirk, Paul J., Jenny Ladefoged & Peter Ladefoged. 1993. Quantifying Acoustic Properties of Modal, Breathy, and Creaky Vowels in Jalapa Mazatec. In *American Indian Linguistics & Ethnography in Honor of Laurence C. Thompson*. University of Montana Occasional Papers in Linguistics No. 10

Krueger, John R. 1977 *Tuvan Manual: Area handbook, grammar, reader, glossary, bibliography*. Uralic and Altaic Series 126. Bloomington: Indiana University Press.

Kunaa, A. Ch. 1970. *Prostoe predlozhenie sovremennogo tuvinskogo yazyka*. Kyzyl: Tuvknizdat.

Letyagina, N. A. 1989. Vyrazhenie kolichestvennyx otnoshennij v tuvinskom yazyke (vzaimodejstvie yazykovyx sredstv raznyx urovnej). In *Grammaticheskie issledovaniya po otdel'nym altajskim yazykam*, pp. 59-92. Leningrad: "Nauka".

Letyagina, N. A. & D. M. Nasilov. 1974. Passiv v tuvinskom yazyke. In Sovetskaya tiurkologiya. #1, 13-24.

Levin, Theodore C. & Michael E. Edgerton. 1999. The Throat-Singers of Tuva. In *Scientific American*, Vol. 281 no. 3, September 1999.

Li, Bing. (1996). *Tungusic Vowel Harmony*. The Hague: Holland Acad. Graphics.

Lopsaŋ, S. 1941. Tiva dildiŋ ažik Ynneriniŋ tugajInda šinčilel. [Research on Tuvan vowels] In *Ertem oruu*, 22-23. Kyzyl.

Lucien, Adam. 1898. Sur l'Harmonie des Voyelles dans les langues Toungouses. Paris.

Maddieson, I. (1984).*Patterns of sounds.* Cambridge UK: Cambridge Univ. Press.

Martan-Ool, M. B. 1986. Kharakteristika izoglossnogo javlenija "faringalizatsija" glasnykh v tuvinskom jazyke. [Characteristics of the isogloss phenomenon of vowel "pharyngealization" in Tuvan.] In *Fonetika jazykov Sibiri i sopredel'nyx regionov*, pp. 37-40. Novosibirsk: "Nauka".

McCarthy, J. & A. Prince (1994) The emergence of the unmarked: Optimality in prosodic morphology, In M. Ginzalez, ed., Procedings of NELS 24, Amherst: GLSA, pp. 333-379.

McCarthy, John & Alan Prince 1995. Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh, 249-384. Amherst, Mass.: GLSA.

Menges, K. H. 1955 The South Siberian Turkic Languages, I: General Characteristics of their Phonology. In *Central Asiatic Journal* 1, 107-36.

_____ 1956. SSTL II: Notes on the Samoyed Substratum. In *Central Asiatic Journal* 2, 161-75.

_____ 1958. Die süd-sibirischen Türk-Sprachen, III: Die Tuba-Sprachen Sojon und Karagas. Zur Charakteristik einer einzelnen süd-sibirischen Gruppe. In *Central Asiatic Journal* 4, pp. 90-129.

_____ 1959. Das Sojonische und Karagassische. In J. Deny et al. [eds.] *Philologicae Turcicae Fundamenta* I pp. 640-70. Wiesbaden: Franz Steiner Verlag.

_____ 1956. SSTL II: Notes on the Samoyed Substratum. In *Central Asiatic Journal* 2, 161-75.

Mongush. D. A. 1959. O vremennyx formax v tuvinskom yazyke. In *Uchenye Zapiski TNIIYaLI VII:* 85-92.

_____ 1968. Ob analiticheskoj forme predikacii v tuvinskom yazyke. In *Uchenye Zapiski TNIIYaLI XIII*: 242-48. Kyzyl, Tuva: TNIIYaLI

_____ 1983. O sluzhebnyx funktsiya slov kizhi, ulus, i chüve v tuvinskom yazyke. In *Tyurkskie yazyki Sibiri*, pp. 12-35. Novosibirsk: AN SSSR SO .

_____ 1987. O nekotoryx modelyax predlozhenij v tuvinskom yazyke, obrazovannyx pri uchastii proizvodnyx ot glagola de=/di= 'say'. In *Pokazateli svyazej v slozhnom predlozhenii (na materiale yazykov raznyx sistem),* pp. 83-88. Novosibirsk: AN SSSR.

_____ 1989. O skazuemyx tuvinskogo yazyka, vyrazhayushchix modal'noe znachenie dostovernosta. In *Predlozhenie v yazykax Sibiri*, pp. 19-35. Novosibirsk: AN SSSR SO.

_____ 1993. Ob odnom tipe predlozheniya frazeologizirovannoj struktury v tuvinskom yazyke. In *Voprosy tuvinskogo yazykoznaniya*, pp. 55-63. Kyzyl: TNIIYaLA.

_____ 1995. Skazuemye tuvinskogo yazyka so znacheniyami irreal'nosti, modal'nosti i voprositel'nosta. In Cheremisina [ed.], pp. 22-51.

Mongush, D. A. 1958. Nastoyashchee zaglaznoe vremya v tuvinskom yazyke. In *Uchenye Zapiski T*NIIYaLI VI: 111-14. Kyzyl: TNIIYa*LA*.

Mongush, D. A. and K. B. Kuular 1992. *Tɨva dìl 6-7* Kyzyl: Tuvanyng nom ündürer cheri.

Mongush, D. A. and Sh. Ch. Sat. 1967. *Sovet üede tɨva dɨldɨŋ xögžüldeza*. Kyzyl: Tuvanyng nom ündürer cheri.

Nadelaev, V. M. 1986. *U istokov tuvinskogo yazyka.* [At the sources of Tuvan]. Kyzyl.

Nedjalkov, Igor. 1997. *Evenki.* New York: Routledge.

Novikova, K. A. 1960. Ocherki dialektov evenskogo yazyka. Leningrad: Nauka.

Orgun, Orhan and Ronald Sprouse. 1997 [2000] From MParse to Control: Deriving ungrammaticality. MS. Univ. of California, San Diego.

Orgun, Orhan. 1996. Sign-based morphology and phonology. Ph.D. Dissertation. Univ. of California, Berkeley

Oyun, M .V. 1983. Pridatochnye opredelitel'nye predlozheniya, vvodimye sluzhebnymi slovami <u>dep, deer, deen</u> v tuvinskom yazyke. In *Tyurkskie Yazyki Sibiri* , pp. 46-55. Novosibirsk: AN SSSR SO.

_____ 1985. Zavisimoe opredelitel'nye predlozheniya v russkom i tuvinskom yazykax. In *Russkij yazyk v Tuve*, pp. 123-41. Kyzyl: Tuvknizdat.

Padgett, Jaye. 1995. Feature Classes. UOMP 18, 385-420. ROA-112.

Piggott, G. 1999. Harmony and Markedness: A Conspiracy in Optimality Theory. Phonoloy 17.

Pokrovskaia, L.A. 1964. *Grammatika gagauzskogo jazyka.* Moscow: Nauka.

Poppe, Nicholas. 1966. *Buriat Grammar.* Indiana University Press.

Potapov, L. V. 1964. The Tuvans. In Potapov and Levin [eds.] *Peoples of Siberia.*, pp. 380-422. Chicago: University of Chicago Press.

Prince, Alan S. & Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. ms. Rutgers University and University of Colorado, Boulder.

Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical Report #2 of the Rutgers Center for Cognitive Science, Rutgers University.

Pulleyblank, Douglas. 1998. Yoruba vowel patterns: Deriving asymmetries by the tension between opposing constraints. ROA.

Radloff, V. V. 1882. *Phonetik der nördlichen Türksprachen.* Leipzig.

Rassadin, V. I.1995. Problemy vozrozhdeniya tofalarskogo yazyka. In *Aborigeny Sibiri*, pp. 62-3. Novosibirsk: NGU.

_____ 1971 *Fonetika i leksika tofalarskogo jazyka.* Ulan Ude: BKI.

_____ 1983. *Morfologiya tofalarskogo yazyka v sravnitel'nom osveshchenie.* Moscow: Nauka.

Ross, Kie. (1996) Floating Phonotactics: Variability in infixation and reduplication of Tagalog loanwords. MA thesis, UCLA.

Samdan, Z. B., Ed. (1994) *Tuvinskoe narodnoe skazki.* Novosibirsk: RAN.

Sarɨkay, M. X. 1973. Svedeniya o rechi tandynskix tuvincev. In *Uchenye Zapiski TNIIYaLI XVI*: 242-50. Kyzyl: TNIIYaLI

Sat, Sh. Ch. 1960. *Sintakticheskie funkcii prichastij v tuvinskom yazyke*. Kyzyl: Tuvknizdat.

_____ 1973. *Formirovanie i razvitie tuvinskogo nacional'nogo literaturnogo yazyka*. Kyzyl: Tuvknizdat.

_____ 1981. Rol' sluzhebnyx slov <u>dep/deesh</u> v slozhnopodchinennom predlozhenii tuvinskogo yazyka. In *Padezhi i ix ekvivalenty v stroe slozhnyx predlozheniyax v yazykax narodov Sibiri*, pp. 112-15. Novosibirsk.

_____ 1983. *Amgì tìva literaturlug dìl: Sintaksis. Pedinstituttuŋ toogu-dìl fakultedinge xergleer ööredilge nomu*. Kyzyl: Tìvanìŋ nom ündürer cheri. [1966: 1st ed.]

_____ 1984. Foneticheski osobennosti terexol'skogo govora tuvinskogo yazyka. In *Issledovaniya zvukovyx sistem yazykov Sibiri*, pp. 187-91. Novosibirsk.

_____ 1959. Prichastie v tuvinskom yazyke. *Uchenye Zapiski TNIIYaLI VII*: 76-84. Kyzyl: TNIIYaLA.

_____ 1987. *Tɨva Dialektologiya.* [Tuvan dialectology] Kyzyl, Tuva: Tyvanyŋ nom ündürer cheri.

Sereedar, N. Ch. 1995. Sposoby vyrazheniya znachenij vremeni v predlozheniyax s imennym skazuemym v tuvinskom yazyke. In Cheremisina [ed.] , 68-77.

Seren, P. 1996. Tsaatan nuur tɨvalar. [The Tsaatan Tuvans of Mongolia] In *Ulug xem.* Kyzyl, Tuva.

_____ 1993. "Tsagaan-Nuur tyvalarynda" [On the Tsagaan-Nuur Tuvans] in Uchyonniye Zapiski TNIIYaLI. Kyzyl, Tuva.

Sezer, Engin. 1981. The k/0 alternation in Turkish. Harvard Studies in Phonology. G.N. Clements, ed.354-382. Bloomington: Indiana University Linguistics Club.

Shamina, L. A. 1983. O modelyax tuvinskix slozhnyx predlozhenii s formy na =gyzhe v zavisimoj chasta. In *Tyurkskie yazyki Sibiri*, pp. 36-45. Novosibirsk: AN SSSR SO.

_____ 1987. *Vremennye polipredikativnye konstrukcii tuvinskogo yazyka*. Novosibirsk: "Nauka".

_____ 1993. Slozhnye predlozheniya v tuvinskom yazyke, vyrazhayushchie otnosheniya oboslovennosti mezhdu sobytiama. In *Voprosy tuvinskogo yazykoznaniya*, pp. 63-75. Kyzyl: TNIIYaLA.

_____ 1995. Analiticheskie konstrukcii modal'no-infinitivnogo tipa v tuvinskom yazyke. In Cheremisina [ed.], pp. 52-67.

Shamina, L. A. & M. A. Cheremisina 1982. Prichastie na -galak (-kalak) v tuvinskom yazyke. In Ubryatova (ed.) , pp.57-70.

Shcherbak, A. M. 1994. *Vvedenie v sravnitel'noe izuchenie tyurskikh jazykov.* [Introduction to the comparative study of Turkic languages] Sankt-Peterburg: Nauka.

Smolensky, Paul. 1993. Harmony, markedness and phonological activity. Paper presented at the Rutgers Optimality Workshop-1, Rutgers University, October 1993.

_____ 1995. The initial state and "richness of the base" in Optimality Theory. Technical report, Cognitive Science Department, Johns Hopkins University, Baltimore, MD. also on ROA-154-1196

Steriade, Donca. 1993. Positional Neutralization. Paper delivered at *NELS* 24.

_____ 1995. Underspecification and markedness. In *The handbook of phonological theory*, ed. John A. Goldsmith, 114-174. Cambridge: Blackwell Publishers.

Stevens, Kenneth, S. J. Keyser & H. Kawasaki. (1986). Toward a phonetic and phonological theory of redundant features. In J. Perkell and D. Klatt (eds.), *Invariance and variability in speech processes.* (pp. 426-449). Hillsdale, NJ: Lawrence Erlbaum.

Tatarintsev, B. A. 1974. *Russkie leksicheskie zaimstvovaniya v sovremennom tuvinskom yazyke*. Kyzyl: Tuvknigoizdat.

_____ 1986. K voprosu o sootvetstvii tipa "faringalizatsija-pervichnaja dolgota". [On the question of the correspondence of pharyngealization with original length] In *Fonetika jazykov Sibiri i sopredel'nyx regionov*, pp. 153-56. Novosibirsk: "Nauka".

_____ 1976. *Mongol'skoe yazykovoe vliyanie na tuvinskuyu leksiku*. Kyzyl:TKI

Tekin, Talat. 1968. *A Grammar of Orkhon Turkic*. Bloomington: Indiana University

Ubryatova, E. A. & F. A. Litvin [eds.] 1986. *Strukturnye tipy sinteticheskix polipredikativnyx konstrukcij v yazykax raznyx sistem*. Novosibirsk: "Nauka".

Vago, Robert. 1988. Vowel harmony in Finnish word games. In *Features, segmental structure and harmony processes part II*, ed. Harry van der Hulst and Norval Smith, 185-205. Dordrecht: Foris.

Verner, G. K. 1972. Problema proiskhozhdenija faringalizatsii v tuvinskom i tofalarskom jazykakh. [The problem of the origin of pharyngealization in Tuvan and Tofalar] *Sovetskaya Tiurkologia* 5: 17-24.

Walker, Rachel. 1999. On the Connection between Round Licensing, Harmony and Bisyllabic Triggers in Altaic. ROA-336-02100.

Wu, Chaolu. (1996). *Daur*. Languages of the World/Materials vol. 93. München: LINCOM EUROPA.

Yip, Moira. 1998. Lexicon optimization in languages without alternations. In *Current trends in phonology*, ed. J. Durand. Paris: Royaumont.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. ROA-282.

Abbreviated author citations used in chapter 1:

Ch	Cheremisina
I&P	Iskhakov & Pal'mbakh
Sh	Shamina