

Coronal consonant harmony in Indus Kohistani*

Paul Arsenault
University of Toronto
& SIL International
paul.arsenault@utoronto.ca

1. Introduction

- This paper investigates co-occurrence restrictions on coronal obstruents in Indus Kohistani, an Indo-Aryan language of the Dardic sub-group in northern Pakistan.
- The study reveals that Indus Kohistani exhibits a pattern of coronal place harmony within roots that is sensitive to the relative similarity of coronal obstruents in terms of their manner of articulation.
- The paper focuses on the methodological and empirical aspects of the study, but also provides a brief discussion of typological and theoretical implications.

2. Background: Coronal harmony, retroflexion, and similarity effects

- Consonant harmony is an assimilatory effect or co-occurrence restriction holding between consonants that are separated by a vowel, and possibly other segments.
- Recent studies have highlighted the following (Hansson 2001, Rose & Walker 2004):
 - Coronal place harmony is by far the most common form of consonant harmony cross-linguistically, especially among sibilants.
 - Patterns of consonant harmony (of all types) are sensitive to the relative similarity of participating consonants. That is, consonants are more likely to harmonize with each other if they are already highly similar, having many features in common.
- Within the published literature, all previously identified cases of coronal harmony involving retroflex obstruents apply to either coronal stops or coronal sibilants (affricates and/or fricatives), but not both (see Table 1).

* I would like to thank Alexei Kochetov, Keren Rice and Ron Smyth for helpful discussion of some of the methodological and theoretical issues addressed in this paper. Any errors are my own.

Table 1. Cases of retroflex consonant harmony reported in Hansson (2001) and Rose & Walker (2004). The term “sibilants” refers to affricates and/or fricatives.

	<i>Language (genus, family)</i>	<i>Retroflex vs. non-retroflex harmony</i>	
		<i>Stops</i>	<i>Sibilants</i>
a.	Malto (Northern Dravidian, Dravidian)	√	--
	Javanese (Sundic, Austronesian)	√	--
	Pohnpeian (Oceanic, Austronesian)	√	--
	Gaagudju (Gaagudju, Australian)	√	--
	Gooniyandi (Bunuban, Australian)	√	--
	Mayali (Gunwinygic, Australian)	√	--
b.	Benchnon Gimira (Omotic, Afro-Asiatic)	--	√
	Rumsen (Costanoan, Penutian)	--	√
	Nebaj Ixil (Mayan)	--	√
	South Peruvian Quechua (Quechuan)	--	√
	Wanka Quechua (Quechuan)	--	√
	Capanahua (Panoan)	--	√

- Malto and Gimira are representative of the typology of retroflex harmony:
 - Malto: coronal stops in a root must be either retroflex or dental (1).
 - Gimira: posterior sibilant affricates and/or fricatives within a root must be either retroflex or palato-alveolar (2).

(1) Retroflex harmony in Malto (Mahapatra 1979; Hansson 2001)

- a. to:totri ‘quickly’ tu:d ‘tiger’
 dudu ‘mother’ danda ‘staff’
- b. *t...t, *d...d, etc.

(2) Retroflex harmony in Gimira (tones omitted) (Breeze 1990; Rose & Walker 2004)

- a. ʃatʃ ‘stretcher’ ʃaʃ ‘vein’
 tʃʰaʃt ‘be pierced’ tʃʰutʃʰ ‘louse’
- b. *ʃ...ʃ, *tʃʰ...tʃʰ, *ʃ...tʃʰ, etc.

- However, in each of the cases listed in Table 1, the contrast between retroflex and non-retroflex consonants is limited to one manner class: either to stops, as in Malto (3a), or to (a subset of) sibilant affricates and fricatives, as in Gimira (3b).

(3) Coronal obstruents in Malto (a) and Gimira (b) (Mahapatra 1979; Breeze 1990)

a.	t	t			
	d	d			
				tɕ	
				ɕ	
	ð	s			

b.	t	tʲ			
	tʰ	tʰʲ			
	d	dʲ			
	ts			tʃ	tʃʰ
	tsʰ			tʃʰ	tʃʰʰ
	s	sʲ	sʷ	ʃ	ʃʰ
	z	zʲ		ʒ	ʒʰ

- Thus, in each case harmony holds between all coronal segment types that exhibit contrastive retroflexion.
- As a result, it is not clear whether similarity plays any role at all, or whether harmony simply applies everywhere that retroflexion is contrastive. Stated differently, it could be that all segments that are contrastive for retroflexion are targets of retroflex harmony, regardless of their similarity to the triggering segment.
- The question is: What might coronal harmony look like in a language with contrastive retroflexion in more than one manner of articulation? Will harmony target all segments that are contrastive for retroflexion, regardless of manner, or will it apply only between consonants that share the same (or similar) manners of articulation?
- In what follows, I present a study of coronal harmony in a language with contrastive retroflexion in both stops and sibilants. The results reveal that under these circumstances a clear similarity effect does emerge.

3. Coronal obstruents in Indus Kohistani

- Indus Kohistani is an Indo-Aryan language of the Dardic sub-group in northern Pakistan.
- Contrasts within the coronal obstruent system include (see (4)):
 - 3 manner classes: stops, affricates, and fricatives
 - 3 place classes: dentals, retroflexes (apical post-alveolars), and ‘palatals’ (laminal post-alveolars)
- Based on Zoller 2005, cf., Hallberg & Hallberg (H&H) 1999, Bashir 2003.

(4) Consonant phonemes of Indus Kohistani

Labial	Dental	Retroflex	Palatal	Dorsal	Uvular	Laryngeal
p p ^h	t t ^h	ɽ ɽ ^h		k k ^h	(q)	
b b ^h	d d ^h	ɖ ɖ ^h		g g ^h		
(f)	ts ts ^h	ʈ ʈ ^h	tʃ tʃ ^h		(x)	h
	s	ʂ	ç		(y)	
	z z ^h	ʐ ʐ ^h	ʒ ʒ ^h			
m	n					
	r	ɽ				
	l					
w			j			

(5) Vowel phonemes of Indus Kohistani

Front				Back			
i	ĩ	i:	ĩ:	u	ũ	u:	ũ:
e	ẽ	e:	ẽ:	o	õ	o:	õ:
				a	ã	a:	ã:

- Some dialects of Indus Kohistani (not represented in the present study) tend to pronounce dental affricates as palatal (H&H 1999: 3–4; Zoller 2005: 3–4).

- There is no contrast for manner among voiced sibilants (i.e., affricate vs. fricative). The voiced palatal fricatives /z, zʰ/ tend to be pronounced as affricates [dʒ, dʒʰ] word-initially.

“[T]here is a strong tendency in word-initial position to pronounce z and z as voiced fricatives and to pronounce ʃ as an affricate. The phoneme ʃ is, however, usually pronounced as a fricative word-finally. There thus exists a common asymmetry between voiced and voiceless affricates/fricatives and within the class of voiced affricates/fricatives in many languages of northern Pakistan. The asymmetry is due to an lenition process which is more advanced in the case of the voiced phonemes than in the case of the voiceless phonemes.” (Zoller 2005: 34; cf., H&H 1999: 11, footnote 14).

- Aspiration is contrastive word-initially, but some instances of final aspiration may not be phonemic. There is a tendency to pronounce voiceless stops and affricates with a slight aspiration word finally (Zoller 2005: 37).
- Voiced consonants tend to be devoiced in final position (Zoller 2005: 37).
- Vowels are contrastive for length, nasalization, and pitch accent.
- To the best of my knowledge, coronal consonant harmony has not been reported for Indus Kohistani in the published literature. The present study is possibly the first attempt to investigate and document consonant harmony in the language.

4. A study of coronal co-occurrence patterns in Indus Kohistani

4.1 Method

4.1.1 Data

- In order to determine whether Indus Kohistani shows synchronic co-occurrence restrictions on coronal obstruents, a corpus of relevant roots was compiled.
- Data for the study was drawn from Zoller’s (2005) dictionary of Indus Kohistani, which contains approximately 8,000 lexical entries.
- Zoller’s (2005) transcription includes some sub-phonemic details such as: (i) word-final aspiration, (ii) word-final devoicing, (iii) final ultra-short vowels (transcribed as superscript) and (iv) the vowel [ʌ], which is an allophone of /a/. The transcription was modified to fit IPA conventions but no attempt was made to phonemicize the data.
- A search of the dictionary revealed 665 lexical entries with word-initial C₁V(N)C₂ sequences, where both C₁ and C₂ are coronal obstruents. This list was reduced to a more restrictive set of 252 roots. Lexical entries were included or excluded from this set based on the following criteria:
 - Neighboring languages: Zoller’s dictionary of Indus Kohistani includes lexical entries for words drawn from the neighboring and closely related languages Gabār (G) (a.k.a. Gowro) and Bhaṭṣīṣē (B) (a.k.a. Baṭṣīṣā). These were excluded on the grounds that G and B are generally regarded as distinct languages (not dialects of Indus Kohistani). Their lexical similarity with Indus Kohistani is only 61% and 58%, respectively (*Ethnologue*, Gordon 2005, Cf., H&H 1999:7).
 - Dialectal variants: The vast majority of data in Zoller’s dictionary are based on the Jijālī (J) dialect of Indus Kohistani. However, the dictionary also includes some data from the Šāṭōṭī (Š) dialect. Wherever variations of the same word were listed for both dialects, only the J word was counted in order to avoid duplicates of a single root. A few words from Š were retained, but only when the J variant was not

available. Five roots were listed as showing variation in the place of articulation of either C_1 or C_2 . In these cases the harmonic variant was counted (i.e., a root was counted as harmonic if it was at least *potentially* so). See examples in §4.3.2.

- Derived and inflected variants were excluded in order to avoid duplicates of a single root. For instance, if masculine and feminine forms of a word were both listed, or nominal and adjectival forms, then only one representative of each was counted. E.g., the masculine adjective $t\lambda^h$ ‘hot’ was counted, but its feminine counterpart $t\lambda^h i$, and its derived nominal counterparts $t\lambda^h i$: ‘heat’ and $t\lambda^h r\acute{u}$: ‘hot ashes’ were not.
- Morphologically complex words were excluded from both lists if an identifiable morpheme boundary occurred between C_1 and C_2 . Among other things, this criterion excluded a large number of reduplicated forms, most of them onomatopoeic (e.g., $t\epsilon\ddot{u}$:- $t\epsilon\ddot{a}$: ‘inarticulate sounds produced by a dumb person’, $t\mathring{s}\acute{a}$:- $t\mathring{s}\acute{a}$: ‘sound of panicking goats’, $z\ddot{o}$:- $z\ddot{o}$: ‘sound of a buzzing or humming insect’, etc.).
- Loanwords: The vocabulary of Indus Kohistani derives primarily from Old-Indo-Aryan, while most Persian and Arabic loans were introduced more recently and account for only about 15% of the lexicon (Zoller 2005: 16). As a result, Persian and Arabic loanwords were excluded, along with English loanwords.¹ All other items were retained, including a few potential loanwords from other Indo-Aryan or non-Indo-Aryan languages of the area (e.g., Shina, Urdu, Burushaski, etc.).

4.1.2 Counts and statistical analysis

- Coronal obstruents were collapsed into 5 natural classes based on place and manner, and ignoring laryngeal distinctions (e.g., dental stops, retroflex stops, dental sibilants, palatal sibilants, and retroflex sibilants).
- Each root was then classified as belonging to one of 15 logically possible combinations of C_1 and C_2 , ignoring the relative order of consonants. For instance, the words $t\lambda s$ ‘a poisonous plant’, $q^h\acute{u}s$ ‘bruised’, $si\ddot{t}i$: ‘a whistle’, $z\grave{a}t^h$ ‘body hair, wool’, $t\grave{e}t^h s^h$ ‘flint’, and $tsa\ddot{f}\acute{a}\ddot{v}$ ‘to lick’ were all classified as representing the combination of retroflex stops with dental sibilants.
- ‘Observed values’ (O): counts of roots for each consonant pair.
- ‘Expected values’ (E): values that would be statistically expected under random co-occurrence of the phonemes.
- Observed-to-expected ratios (O/E ratios) were computed for each consonant combination to determine whether some combinations are over represented or underrepresented in the corpus (cf. Frisch et al. 2004, Kawahara et al. 2006, etc.).²
 - Ratio of 1.00: no difference between observed and expected frequency.
 - Ratio of more than 1.00: the combination is over-represented.
 - Ratio of less than 1.00: the combination is under-represented.
 - Ratio of 0.00: categorical under-representation (complete absence).

¹ In actual fact, the study was conducted both with and without the Persian and Arabic loanwords. Apart from a minor difference in the results for fricative/fricative combinations, the inclusion or exclusion of these items did not alter the overall pattern that emerged.

² Some studies have used the Chi-square test to determine the statistical significance of over and under-representations (e.g., Kawahara et al. 2006). However, there is some doubt as to whether this test has been applied properly. Thus, for the purpose of the present study I will not attempt to calculate the statistical significance of over and under-representations, and will let the observed counts and O/E ratios speak for themselves.

4.2 Results

4.2.1 #CV(N)C sequences with sibilants (affricates and fricatives) collapsed

- The results for word initial CV(N)C sequences are shown in Tables 2–3, where affricates and fricatives have been collapsed into a single manner class.

Table 2. Observed counts of coronal obstruents in #CV(N)C sequences, with sibilants collapsed. Values expected under random co-occurrence are shown in brackets (n = 252).

C1/C2	t t ^h d d ^h	t t ^h ɖ ɖ ^h	ts ts ^h s z z ^h	tʃ tʃ ^h ʃ ʒ ʒ ^h	tʃ tʃ ^h ʃ z z ^h
t t ^h d d ^h	28 (16.3)	2 (29.5)	37 (24.4)	25 (27.9)	10 (15.6)
t t ^h ɖ ɖ ^h		27 (12.4)	28 (24.8)	30 (29.0)	9 (14.8)
ts ts ^h s z z ^h			10 (7.3)	4 (15.6)	1 (10.6)
tʃ tʃ ^h ʃ ʒ ʒ ^h				21 (8.3)	0 (11.9)
tʃ tʃ ^h ʃ z z ^h					20 (3.6)

Table 3. Ratio of observed to expected values, with sibilants collapsed.

C1/C2	t t ^h d d ^h	t t ^h ɖ ɖ ^h	ts ts ^h s z z ^h	tʃ tʃ ^h ʃ ʒ ʒ ^h	tʃ tʃ ^h ʃ z z ^h
t t ^h d d ^h	1.72	0.07	1.51	0.90	0.64
t t ^h ɖ ɖ ^h		2.17	1.13	1.03	0.61
ts ts ^h s z z ^h			1.38	0.26	0.09
tʃ tʃ ^h ʃ ʒ ʒ ^h				2.53	0.00
tʃ tʃ ^h ʃ z z ^h					5.61

Key observations

- First, combinations of two coronal obstruents that share the same manner and *agree* in place of articulation are over-represented, in most cases substantially. This includes:
 - 2 stops (both retroflex O/E = 2.17, both dental O/E = 1.72);
 - 2 sibilants (both retroflex O/E = 5.61, palatal O/E = 2.53, dental O/E = 1.38).
- Second, combinations of two coronal obstruents that share the same manner and *disagree* in place of articulation are either categorically absent or substantially under-represented.
 - 2 stops (dental/retroflex O/E = 0.07);
 - 2 sibilants (palatal/retroflex O/E = 0.00, dental/retroflex O/E = 0.09, dental/palatal O/E = 0.26).
 - The four lowest O/E ratios in the table correspond to the four possible combinations of same-manner coronals that disagree in place.
- Third, combinations of two coronal obstruents with different manners of articulation (i.e., stops with sibilants of any kind) do not show categorical co-occurrence restrictions and, overall, are not substantially under-represented or over-represented.
 - Although some cells representing stop/sibilant combinations show a degree of over or under-representation, they lack the extreme polarization evident in the same-manner combinations where over-representation of pairs agreeing for place of articulation is accompanied by categorical absence or substantial under-representation of pairs disagreeing for place.

- In sum, Indus Kohistani exhibits a pattern of coronal place harmony that is sensitive to the similarity of participating obstruents in terms of their manner of articulation. Coronal obstruents are harmonic for place of articulation, but only if they share the same manner (i.e., both stops or both sibilants).

4.2.2 #CV(N)C sequences with sibilants (affricates and fricatives) expanded

- Tables 4–5 show the results for word-initial CV(N)C sequences when sibilants are expanded into two manner classes: affricates and fricatives.

Table 4. Observed counts of coronal obstruents in #CV(N)C sequences, with sibilants expanded. Values expected under random co-occurrence are shown in brackets (n = 252).

C1/C2	t t ^h d d ^h	t̪ t̪ ^h d̪ d̪ ^h	ts ts ^h	tʃ tʃ ^h	tʂ tʂ ^h	s z z ^h	ʃ ʒ ʒ ^h	ʂ ʐ ʐ ^h
t t ^h d d ^h	28 (16.3)	2 (29.5)	5 (6.1)	7 (11.5)	5 (6.0)	32 (18.4)	18 (16.5)	5 (9.5)
t̪ t̪ ^h d̪ d̪ ^h		27 (12.4)	9 (6.6)	18 (12.4)	2 (6.2)	19 (18.2)	12 (16.6)	7 (8.7)
ts ts ^h			2 (0.2)	0 (0.9)	0 (0.7)	3 (2.5)	0 (2.0)	0 (1.8)
tʃ tʃ ^h				3 (0.9)	0 (1.3)	4 (4.8)	5 (4.0)	0 (3.3)
tʂ tʂ ^h					2 (0.4)	1 (2.8)	0 (2.4)	10 (1.8)
s z z ^h						5 (4.6)	0 (7.9)	0 (5.4)
ʃ ʒ ʒ ^h							13 (3.4)	0 (4.8)
ʂ ʐ ʐ ^h								8 (1.4)

Table 5. Ratio of observed to expected values, with sibilants expanded.³

C1/C2	t t ^h d d ^h	t̪ t̪ ^h d̪ d̪ ^h	ts ts ^h	tʃ tʃ ^h	tʂ tʂ ^h	s z z ^h	ʃ ʒ ʒ ^h	ʂ ʐ ʐ ^h
t t ^h d d ^h	1.72	0.07	0.82	0.61	0.83	1.74	1.09	0.52
t̪ t̪ ^h d̪ d̪ ^h		2.17	1.35	1.45	0.32	1.05	0.72	0.81
ts ts ^h			9.33	0.00	0.00	1.21	0.00	0.00
tʃ tʃ ^h				3.27	0.00	0.83	1.26	0.00
tʂ tʂ ^h					4.80	0.36	0.00	5.70
s z z ^h						1.10	0.00	0.00
ʃ ʒ ʒ ^h							3.82	0.00
ʂ ʐ ʐ ^h								5.73

Observations concerning sibilant harmony

- First, all exceptions to sibilant harmony involve affricate/fricative pairs, while sibilants with identical manner (affricate/affricate and fricative/fricative) exhibit a more

³ When sibilants are expanded into affricates and fricatives, the overall low count for certain sibilants leads to very low expected frequencies for some combinations. This in turn leads to exaggerated O/E ratios that are not statistically reliable (a problem that was avoided in Tables 2–3 by collapsing the sibilant class). For instance, the expected frequency for dental affricate pairs is only 0.2. As a result, an observed count of 2 leads to an O/E ratio of 8.38. This figure is probably over-inflated and unreliable from a statistical point of view. Nevertheless, the observed values in affricate/affricate and fricative/fricative combinations exhibit a clear and categorical pattern of place harmony.

categorical pattern of place agreement.⁴ This suggests that, even among sibilants, harmony is enforced more strictly between segments with identical manner of articulation, than between those with slightly different manners of articulation.

- Second, all exceptions to sibilant harmony involve *dental fricatives* co-occurring with palatal or retroflex affricates. There are no exceptions to sibilant harmony in the data set that involve combinations of palatal and retroflex sibilants (whether affricates or fricatives). This suggests that harmony among posterior sibilants might be enforced more strictly than harmony between anterior (dental) and posterior (palatal/retroflex) sibilants.
- Third, dental fricatives appear to display a gradient form of harmony in combination with affricates: they occur most freely with dental affricates (O/E = 1.21), less often with palatal affricates (O/E = 0.83), and least often with retroflex affricates (O/E = 0.36).⁵

4.2.3 *Co-occurrence of laryngeal features*

- Voicing and aspiration are contrastive among the obstruents of Indus Kohistani. However, Zoller's transcription includes some predictable, sub-phonemic instances of final devoicing and aspiration.
- Aspiration clearly does not participate in harmony. There are few cases of #C₁V(N)C₂ sequences where both C₁ and C₂ are aspirated (regardless of whether cases of final aspiration are taken as phonemic or not).
- This is not surprising given that a form of Grassman's law – a dissimilatory rule prohibiting sequences of two aspirates in successive syllables – operated in Sanskrit (MacEachern 1997) and is still evident in New Indo-Aryan Dardic languages such as Palula (Liljegren 2008).
- For the purpose of determining co-occurrence effects for voicing between coronal obstruents, only #CVC sequences were considered (n=201). #CVNC sequences were excluded (n=51) because obstruents are predominantly voiced after nasals.
- Tables 6–7 show the results for #CVC sequences with obstruents classified according to manner and voicing. Word-final phonetically devoiced obstruents (e.g., [d̥] in Zoller's transcription) were classified as voiced for the purpose of this analysis on the grounds that they are voiced phonemically.

⁴ If Persian and Arabic loanwords are admitted in the data, then this generalization is weakened because they introduce a few disharmonic combinations of dental and palatal fricatives. These excluded exceptions are discussed in § 4.3.2.

⁵ If Persian and Arabic loanwords are admitted in the data, they introduce some exceptional forms involving co-occurrence of dental and palatal fricatives. In this case, the surface pattern that emerges for fricative/fricative combinations shows the same gradient effect: dental/dental pairs occur most often, dental/palatal pairs less often, and dental/retroflex pairs least of all.

Table 6. Observed counts of coronal obstruents in #CVC sequences, with obstruents classified according to manner and voicing. Values expected under random co-occurrence are shown in brackets (n = 201).

C1/C2	t t ^h t t ^h	d d ^h d̥ d̥ ^h	ts ts ^h tʃ tʃ ^h tʂ tʂ ^h s ʃ ʂ	z z ^h ʒ ʒ ^h z̥ z̥ ^h
t t ^h t t ^h	22 (17.9)	7 (22.3)	76 (66.5)	9 (11.3)
d d ^h d̥ d̥ ^h		14 (4.0)	20 (23.7)	4 (4.8)
ts ts ^h tʃ tʃ ^h tʂ tʂ ^h s ʃ ʂ			33 (34.8)	12 (14.2)
z z ^h ʒ ʒ ^h z̥ z̥ ^h				4 (1.4)

Table 7. Ratio of observed-to-expected values with obstruents classified according to manner and voicing.

C1/C2	t t ^h t t ^h	d d ^h d̥ d̥ ^h	ts ts ^h tʃ tʃ ^h tʂ tʂ ^h s ʃ ʂ	z z ^h ʒ ʒ ^h z̥ z̥ ^h
t t ^h t t ^h	1.23	0.31	1.14	0.79
d d ^h d̥ d̥ ^h		3.46	0.84	0.83
ts ts ^h tʃ tʃ ^h tʂ tʂ ^h s ʃ ʂ			0.95	0.85
z z ^h ʒ ʒ ^h z̥ z̥ ^h				2.96

Observations concerning co-occurrence of laryngeal features

- Combinations of two coronal stops that agree in voicing are somewhat over-represented, while those that disagree are somewhat under-represented:
 - Both voiced O/E = 3.46, both voiceless O/E = 1.23
 - Voiced/voiceless O/E = 0.31
- Combinations of two coronal sibilants do not show a clear pattern of voicing agreement or disagreement. While pairs of voiced sibilants are over-represented, pairs of voiceless sibilants and pairs of sibilants that disagree in voicing both have observed counts that are not far below their expected frequencies.⁶
 - Both voiced O/E = 2.96 (but note the low expected frequency)
 - Both voiceless O/E = 0.95, voiced/voiceless O/E = 0.85
- In sum, aspiration plays no role in coronal harmony and voicing plays only a marginal role at best. Stop/stop combinations show a tendency toward voicing agreement, but sibilant/sibilant combinations do not. Moreover, even in the case of stop/stop combinations the pattern of voicing agreement is far from categorical.

⁶ The lack of voicing agreement among sibilants in Tables 6–7 is not the result of grouping affricates and fricatives into a single sibilant class. Of the 12 observed roots with sibilant/sibilant combinations that disagree in voicing, 7 are affricate/fricative combinations and 5 are fricative/fricative combinations. Thus, the pattern is not substantially different for sibilants with identical manners as compared to sibilants with slightly different manners.

4.3 Representative Examples

4.3.1 Harmonic and disharmonic roots

(6) Stop/stop combinations (harmonic)

a.	to:tá:	‘a blue-colored butterfly’	tʌtú:	‘a type of small horse’
	tʌtʰ	‘hot; heat’	tʌ:tʰ	‘a small rug’
	tʰatʌr	‘smallpox’	tʰatʌr	‘shallow’
	đít ^{hi}	‘given’	đí:t ^{hi}	‘span of hand’
	dadál	‘measles’	da:dí:	‘trustworthy’
	dʰʌtr̩:	‘burnt food’	dʰà:đ	‘a woodpecker’
b.	tùnd	‘a kind of basket’	taŋdáv	‘to beat’
	dùnd	‘a flock, herd’	dá:ŋđ	‘a stick’
	dʰʌndaʰ	‘dealings, business’	dʰã:ŋđ	‘a pond’

(7) Affricate/affricate combinations (harmonic)

	tsíts ^{hi}	‘nipple, breast’	tʂítʂ ^h	‘grey, spotted’
	tʂi:tʂʌkʰ	‘smallpox’	tʂ ^{hi} i:tʂáv	‘to learn’
	tʂã:tʂú:	‘a dwarf’		

(8) Fricative/fricative combinations (harmonic)

a.	sʌzú:	‘sister’s son’	ʂiʂ	‘a head’
	zʰã:z	‘a branch of a holm oak’	ʂiʂã:	‘left over’
	ʂiçáv	‘to dry (up)’	ʂiʂò:	‘a kind of rhubarb’
	çò:çʌ:	‘decoration’	ʂù:ʂ	‘decent, fine, proper’
	z̩iʂi:l	‘type of pheasant’	zʰʌʂ	‘a shift, move, pull’
b.	zʰʌnzé:r	‘a kind of brown bird’	zʰʌnzà:l	‘confusion, disorder’
	çʌmçʌtá:	‘a turtle’		

(9) Affricate/fricative combinations (harmonic)

a.	tsàs	‘a pinch; a sudden pull’	tʂí:ʂ	‘thirst’
	tsãz	‘soft (kernel, bread)’	tʂ ^{hi} i:z	‘a hill’
	tʂuçt̩i:	‘absorption’	tʂòeʂ	‘mother-in-law’
	tʂ ^h ʌç-tʂʌç	‘sit down!’	tʂo:ʂáv	‘to wring out’
	tʂ ^h ʌʂ̩	‘a winnowing tray’	tʂ ^h ãzáv	‘to make a mistake’
	çã:tʂ	‘a kind of mungo’ (Š)	zà:tʂ	‘a grape’
b.	tʂ ^h ʌnzò:	‘a curry comb’	zʌmtʂú:	‘a son-in-law’

(10) Stop/affricate combinations (disharmonic)

a.	tʂatáv	‘to lick’	tʂè:tʂ ^h	‘a flint’
	tʂatʌn	‘a lintel’	tʂʌtú:	‘a grater for spices’
	tʂ ^h atáv	‘to plaster, throw mud’	ta:tʂ ^h áv	‘to carve’
b.	tsʌŋđ	‘forelock of a man’	tʂé:ŋđ	‘a crack, fissure’

(11) Stop/fricative combinations (disharmonic)

a.	dʌjɕ ⁱ	‘ten’	dù:ʂ	‘a sin’
	tʰosà:	‘a fist, punch’	siʃi:	‘a whistle’
	ɕòtʰ	‘a bump, swelling’	ɕʌtʰ	‘strength, power’
b.	dʌnɕilè:	‘the spleen’	ɕʌndà:l	‘a roof’
	sá:ŋǽ	‘a bull’	ɕʌŋǽ	‘barren, castrated’

4.3.2 Exceptions

- Exceptions involve sequences of two coronal obstruents with the same manner of articulation that disagree in place.
- A list of all exceptional forms that were included in the data set is provided in (12).

(12) Exceptional disharmonic forms included in the data set

a.	dʌe:tʰ	‘a prostitute’	tʌŋǽ	‘sound of a fired bullet’
b.	sʌtɕ	‘true, right’	tɕʌskà ^h	‘pleasure’
	sʌtɕʌṽ	‘a marble game’	sù:i:tʂ	‘female (animal)’
	sù:tɕ	‘a thought’		

- Other exceptions were found in Zoller (2005) but did not meet the requirements for inclusion in the data set. These exceptions fall into three categories:
 - (i) Disharmonic forms with harmonic variants (13). For these roots, only the harmonic variant was included.
 - (ii) Morphologically complex forms in which an identifiable morpheme boundary intervenes between C₁ and C₂ (14).
 - (iii) Loanwords from English, Persian, or Arabic (15).

(13) Roots showing variation between harmonic and disharmonic forms

a.	tʰó:tʰ	~ tʰó:tʰ	~ tʰò:to (Š)	‘fur shoes’
	ǫj:tʰi	~ ǫj:tʰ	(Š)	‘span of hand’
b.	ʂũ:ʂ	~ sũ:ʂ		‘straight’
	zʰʌʂ	~ zʰʌs		‘a shift, move; pull’
	tʂʌʂ	~ tʂʌs		‘sudden fierce or stabbing pain’

(14) Disharmony in morphologically complex forms

a.	ǫǫj-tapxʌj	‘spine’	
b.	ʂʌ-s	‘this (OBL)’	← ʂũ ^h ‘this’ (NOM)
	ʂʌ-ze:	‘here, at this place’	← ʂʌ-s ‘this (OBL)’ (← ʂũ ^h) + zé: ‘place’
	ʂʌ-zior:	‘here along’	← ʂʌ-ze: ‘here’ (← ʂũ ^h)
	ʂú-ɕà:nā:	‘similar’	← ʂũ ^h ‘this’ + ɕa:n ‘resembling’
	ʂa-z-ʌṽ	‘to be attached’	← ʂáṽ ‘to attach’ + passive morpheme -z-
	ʂʌ-z-è:l	‘closed, stuck’	← ʂáṽ ‘to attach’ + passive morpheme -z-
	zʰa:-zʌj	‘brother's wife’	← zʰà: ‘brother’

- (15) Disharmony in loanwords
- | | | |
|----------------------|------------------|--|
| sΛçmà ^h | ‘Tuesday’ | < <i>se</i> ‘three’ + <i>šambah</i> ‘Saturday’ (Persian) |
| suždà ^h | ‘divine worship’ | < <i>sijda</i> (Arabic) |
| çaza:dà ^h | ‘a prince’ | < <i>šāhzāda</i> (Persian) |
| žisam | ‘body’ | < <i>jism</i> (Arabic) |
| žù:s | ‘(fruit) juice’ | < ‘ <i>juice</i> ’ (English) |
| ž ^h á:z | ‘an airplane’ | < <i>jahāz</i> ‘ship, vessel’ (Arabic) |

5. Historical sources of harmony

- Consonant harmony in Indus Kohistani is manifested as a static morpheme structure constraint. The study found no evidence of active alternations in the synchronic grammar.
- Limited diachronic evidence concerning the source of harmony is listed in (16). Most of the examples involve retroflex harmony.

(16) Evidence of diachronic harmony between obstruents with the same manner⁷

- a. t/d → t̪/d̪
- | | | | |
|-----------------------------------|------------------------------|----------------------------------|-------------------------------------|
| t̪Λt̪ú: | ‘type of small sturdy horse’ | < * <i>taṭṭu</i> ⁸ | ‘pony’ (cf., Marathi <i>taṭṭū</i>) |
| t̪ ^h à:t̪ ^h | ‘clever, cheerful’ | < <i>taṭṭha-</i> < <i>taṣṭá-</i> | ‘fashioned, formed in mind’ |
| ḍj̪:t̪ ^{hi} | ‘span of hand’ | < <i>diṣṭi</i> | ‘a measure of length’ |
| ḍa:ḍj̪: | ‘trustworthy’ | < <i>dārdhya-</i> | ‘hardness, firmness, strength’ |
| ḍá:ṇḍ̪ | ‘a stick’ | < <i>daṇḍá</i> | ‘stick, club’ |
- b. c[t̪ç]/s[ç] → t̪ç⁹
- | | | | |
|----------------------------|-----------------------------|---------------------------------------|---------------------------|
| t̪ç̪i:t̪ç̪ ^h | ‘grey, spotted, variegated’ | < * <i>ci[t̪ç̪]-</i> < <i>citrá-</i> | ‘conspicuous, variegated’ |
| t̪ç̪ ^h i:t̪ç̪á̃ | ‘to learn’ | < * <i>śi[t̪ç̪]-</i> < <i>śikṣate</i> | ‘learns’ |
- c. s → ç/ṣ, ś[ç] → ṣ
- | | | | |
|---------------|--------------------------|---------------------|----------------|
| çò:ça: | ‘decoration’ | < * <i>suśobha-</i> | ‘splendid’ |
| ṣù:ṣ | ‘decent, fine, proper’ | < <i>sūkṣma-</i> | ‘minute, fine’ |
| ṣũ:ṣ | ‘straight’ | < * <i>sūṅkṣa-</i> | ‘straight’ |
| ṣiṣ | ‘a head’ | < <i>śirśá-</i> | ‘head, skull’ |
| ṣiṣā: | ‘left over’ | < <i>śeṣa-</i> | ‘remaining’ |
| ṣi:ṣΛã̃ ḥΛVá: | ‘a light wind or breeze’ | < <i>śuṣi</i> | ‘breath’ |

⁷ The etymologies in (16) are based on Zoller (2005) and Turner (1966). Forms with an asterisk (*) are hypothetical reconstructed forms, while all others are attested in ancient Vedic and/or Classical Sanskrit texts.

⁸ Zoller (2005) identifies this word as **taṭṭu* (not **taṭṭu*). However, Zoller’s source is Turner (1966), who reconstructs it as either **taṭṭu* or **taṭṭu*. I assume that the disharmonic form **taṭṭu* is original on the grounds that word-initial retroflex consonants were rare or non-existent at one time in OIA (Masica 1991: 157). The disharmonic form is attested in at least one other other Indo-Aryan language, i.e., Marathi *taṭṭū* ‘pony’.

⁹ Old Indo-Aryan did not have retroflex affricates. They developed in Indus Kohistani and other Dardic languages primarily from original *-kṣ-* and *-tr-* sequences. E.g., t̪ç̪^hà:r ‘waterfall’ < *kṣará* ‘melting away’, t̪ç̪à: ‘three’ < *tráyaḥ* ‘three’.

- The evidence in (16) is compatible with either (i) “anticipatory” right-to-left harmony and/or (ii) a trigger/target asymmetry among coronal places in which retroflexes are dominant, dentals are recessive, and palatals are intermediate.
- I tentatively assume the dominant-recessive interpretation for the following reasons:
 - Similar dominant-recessive asymmetries occur in patterns of local coronal assimilation in Old Indo-Aryan (e.g., Sanskrit). In those cases, direction is not a factor (Whitney 1967 [1889]: 68).
 - The apparent directionality in (16) might be a by-product of the fact that word-initial retroflexes were rare at one time in Old-Indo-Aryan (Masica 1991: 157).
- Harmony did not apply diachronically between obstruents with different manners (i.e., stops and sibilants) (17).

(17) No diachronic harmony between obstruents with different manners

taʈʂʰáṽ	‘to carve’	< *ta[ʈʂ]-	< <i>tákṣati</i>	‘forms by cutting’
daʈʂʰó: ~ daʂõ: (Š)	‘(at the) right’	< *da[ʈʂ]-	< <i>dákṣiṇa-</i>	‘right (hand)’
dʌj̥e ⁱ ~ dʌç (Š)	‘ten’		< <i>dása</i>	‘ten’
dù:ʂ	‘a sin’		< <i>dōṣa</i>	‘fault’
siʈi:	‘a whistle’		< * <i>sīṭṭa-</i>	‘whistle’
çòʈʰ	‘a bump, swelling’		< * <i>śoṭṭha-</i>	‘defective’

- Intervocalic retroflex stops of OIA developed into retroflex flaps in Indus Kohistani via a process of lenition, while geminate stops were reduced to singletons (e.g., berʂà^h ‘a boat’ < *bedā-*, buḍà: ‘an old man’ < **budḍha-*).
- Given the historical relation of retroflex flaps to stops, it is interesting to note that stops fail to harmonize with retroflex flaps (18a). Only one example shows variable harmony between a stop and a retroflex flap (18b).

(18) No harmony between stops and retroflex flaps

a. tʌʂʌq ^h	‘a blow, knock, stroke’	< * <i>taḍati</i>	‘strikes’
d ^h ʌʂ	‘a body’	< * <i>dhaḍa-</i>	‘trunk of body’
d ^h a:ʂà ^h	‘a cattle raid’	< <i>dhāṭi-</i>	‘assault’
d ^h úʂ	‘dust’	< * <i>dhūḍi-</i>	‘dust, powder’
b. tʌʂa: ~ ʈʌʂa:	‘a cliff’	< <i>taṭa-</i>	‘slope, shore’

6. Discussion: Theoretical and typological issues

6.1 The similarity effect

- The study of Indus Kohistani reveals a pattern of coronal place harmony that is sensitive to the similarity of participating obstruents in terms of their manner of articulation.
 - Retroflex harmony does not simply target any segment type for which retroflexion is contrastive. Rather, it holds only between segment types that share identical, or near-identical, manners of articulation.
- Theoretical models of consonant harmony must account for similarity effects.

- Earlier ‘feature spreading’ accounts of consonant harmony (e.g., Shaw 1991, Gafos 1999) did not predict similarity effects (cf., discussion in Rose & Walker 2004, Arsenault & Kochetov 2009, to appear).
- This is not necessarily a problem for ‘feature spreading’, but only for the feature spreading models as they have been applied to consonant harmony.
- Recent ‘agreement by correspondence’ (ABC) models do encode similarity effects (e.g., Hansson 2001, Rose & Walker 2004), but some issues remain unresolved:
- How is similarity evaluated?
 - Similarity metric: number of shared feature classes divided by the sum of shared and unshared feature classes (Frisch et al. 2004).
 - Similarity metric with differential feature weighting (Gallagher & Coon 2008).
 - Natural classes based on contrastive feature specifications (Mackenzie 2009).
- Why does manner count in the evaluation of similarity, but not laryngeal features?
 - Could we expect to find a language where coronal place assimilation is conditioned by similarity of laryngeal features (e.g., coronal obstruents harmonize for place only if they share the same voicing and/or aspiration)?
 - The ABC model does not rule this out. But I am not aware of any cases of coronal place harmony conditioned by voicing/aspiration.¹⁰
- The lack of harmony between stops and retroflex flaps, despite their historical relation, suggests either:
 - (i) Phonetic properties (i.e., not just phonological) can determine similarity, or...
 - (ii) Flaps had developed independent phonemic status as sonorants at the time when harmony took effect, and were thus no longer phonologically related to stops.

6.2 Coronal harmony and reduplication

- Indus Kohistani displays many examples of reduplication, most of them onomatopoeic (e.g., t^hù:-t^hu: ‘sound of spitting’, t̪s̪ã:-t̪s̪ã: ‘sound of panicking goats’), but some of them possibly serving an emphatic function (e.g., dudù^h ‘very quickly’).
- Reduplication results in the same general surface pattern as coronal harmony, i.e., a sequence of two coronal obstruents with the same manner and place (and laryngeal features), separated minimally by a vowel.
- The surface pattern cannot be attributed solely to reduplication.
 - All identified or suspected cases of reduplication were excluded from the current study on the grounds of morphological complexity (§ 4.1.1).
 - The diachronic evidence clearly shows a process of harmony at work (§ 5).
- However, the interaction between harmony and reduplication bears a resemblance to so-called phonological ‘conspiracies’ in which diverse phonological processes yield the same surface pattern.
- Within the framework of Optimality Theory such conspiracies are often subsumed under a single output constraint.
- Question: Is there any formal relation between harmony and reduplication (e.g., an output correspondence constraint)?

¹⁰ Chol (Mayan) comes close. Pairs of plain or ejective sibilants must be identical for place and manner. However, mixed plain/ejective combinations are still subject to coronal place harmony, but can differ in manner (Gallagher & Coon 2008).

6.3 Coronal harmony as a genetic or areal trait

- A recent study by Arsenault & Kochetov (2009, to appear) found very similar results for Kalasha, another Dardic language closely related to Indus Kohistani.
- Kalasha has a very similar phonemic inventory, and displays a nearly identical pattern of coronal place harmony conditioned by manner (with minor differences in sibilants).
- This suggests that coronal harmony might be a genetic trait of the Dardic (and most likely other Indo-Aryan) languages, or even an areal trait affecting non-Indo-Aryan languages of the area (which tend to have similar coronal obstruent inventories).
- Both of these possibilities are currently being explored:
 - As part of my doctoral research on coronal assimilation in South Asia, I am conducting similar studies on two other Dardic languages: Kalami (data courtesy of Joan Baart) and Palula (data courtesy of Henrik Liljegren).
 - In collaboration with Alexei Kochetov (SSHRC Research Grant 416-2006-1006, “Mechanisms of consonant assimilation”) I am exploring the possibility of harmony in non-Indo-Aryan languages of northern Pakistan including Wakhi (Irianian) and Burushaski (Islolate).

7. Conclusion

- The study of Indus Kohistani reveals a pattern of coronal place harmony that is sensitive to the similarity of coronal obstruents in terms of their manner of articulation.
- To the best my knowledge, this pattern has not been previously reported or documented for Indus Kohistani, and no pattern like it has been reported for any language in the literature on consonant harmony (apart from Arsenault & Kochetov’s recent study of Kalasha (2009, to appear)). Thus, the study makes an important empirical contribution.
- The evidence from Indus Kohistani highlights the role of similarity in consonant harmony systems, but many questions remain unanswered.
 - How is similarity evaluated?
 - Does harmony hold over longer domains (i.e., other than #CV(N)C sequences)? If so, do any intervening segments display blocking effects?
 - Is coronal harmony a genetic or areal trait? If so, how widespread is it among languages of the Indo-Aryan family or other languages of the area?
- These and other questions are the subjects on ongoing research.

References

- Arsenault, Paul and Alexei Kochetov (to appear). Retroflex harmony in Kalasha: Agreement or spreading? Proceedings of NELS 39.
- Arsenault, Paul and Alexei Kochetov. 2009. Retroflex harmony in Kalasha. Paper presented at the 83rd Annual Meeting of the Linguistic Society of America, San Francisco, January 8-11, 2009.
- Bashir, Elena. 2003. Dardic. In The Indo-Aryan languages, ed. George Cardona and Dhanesh Jain, 818–894. London and New York: Routledge.
- Breeze, M. J. 1990. A sketch of the phonology and grammar of Gimira (Benchnon). In Omotoc language studies, ed. R. J. Hayward, 1–67. School of Oriental and African Studies, University of London.
- Frisch, Stefan A., Janet B. Pierrehumbert, and Michael B. Broe. 2004. Similarity avoidance and the OCP. Natural Language and Linguistic Theory 22: 179–228.
- Gafos, Adamantios I. 1999. The articulatory basis of locality in phonology. New York & London: Garland Publishing.
- Gallagher, Gillian and Jessica Coon. 2008. Distinguishing total and partial identity: evidence from Chol. Unpublished manuscript.
- Gordon, Raymond G., Jr. (ed.). 2005. Ethnologue: languages of the world, fifteenth edition. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>.
- Hallberg, Daniel G. and Calinda E. Hallberg. 1999. Indus Kohistani: a preliminary phonological and morphological analysis. Islamabad: National Institute of Pakistan Studies and Summer Institute of Linguistics.
- Hansson, Gunnar Ólafur. 2001. Theoretical and typological issues in consonant harmony. Doctoral dissertation, University of California, Berkeley.
- Kawahara, Shigeto, Hajime Ono, and Kiyoshi Sudo. 2006. Consonant co-occurrence restrictions in Yamato Japanese. Japanese/Korean Linguistics 14: 27-38. Stanford: CSLI Publications.
- Liljegren, Henrik. 2008. Towards a grammatical description of Palula: an Indo-Aryan language of the Hindu Kush. Doctoral dissertation, Stockholm University.
- MacEachern, Margaret R. 1997. Laryngeal cooccurrence restrictions. Doctoral dissertation, University of California, Los Angeles.
- Mackenzie, Sara. 2009. Contrast and similarity in consonant harmony processes. Doctoral dissertation, University of Toronto.
- Mahapatra, B.P. 1979. Malto: an ethnosemantic study. Mysore: Central Institute of Indian Languages.
- Masica, Colin P. 1991. The Indo-Aryan languages. Cambridge: Cambridge University Press.
- Rose, Sharon and Rachel Walker 2004. A typology of consonant agreement as correspondence. Language 80: 475–531.
- Shaw, Patricia A. 1991. Consonant harmony systems: the special status of coronal harmony. In Phonetics and phonology, Vol. 2, The special status of coronals: internal and external evidence, ed. Carole Paradis and Jean-François Prunet, 125–157. San Diego: Academic Press.
- Turner, R. L. 1966. A comparative dictionary of Indo-Aryan languages. London: Oxford University Press.
- Whitney, W. D. 1967 [1889]. Sanskrit grammar. Cambridge, Mass.: Harvard University Press.
- Zoller, Claus Peter. 2005. A Grammar and Dictionary of Indus Kohistani. Volume 1: Dictionary. Berlin and New York: Mouton de Gruyter.