MOT Phonology Workshop
University of Toronto
March 9–11, 2012

An Evolutionary Perspective on Retroflex Phonotactics

Paul Arsenault
University of Toronto
& Tyndale University College
parsenault@tyndale.ca

1 Introduction

- Retroflex consonants are typologically marked; they occur in only about 11% of the world's languages, typically in larger inventories that include at least one other coronal series (e.g., dental or alveolar) (Maddieson, 1984, p. 32; Hamann, 2003, p. 3).
- They tend to have a limited distribution relative to their non-retroflex counterparts.
- This paper examines two contradictory phonotactic constraints on retroflex and other apical consonants.

(1) Phonotactic constraints on retroflexion/apicality
   a. No apical/retroflexes in strictly prevocalic (CV) positions
   b. No apical/retroflexes in strictly postvocalic (VC) positions

- Main Point: These contradictory patterns can be explained if phonotactic constraints on retroflexion are a direct result of the evolution of retroflexion in a language.
- This stands in contrast to previous studies, which posit (universal?) synchronic constraints on retroflexion, grounded in speech perception (Hamilton, 1996; Steriade, 2001; Hamann, 2003).

2 Coronal articulations

- Cross-linguistically, languages can distinguish up to four coronal 'place' series.

(2) Coronal place articulations

	ANTERIOR	POSTERIOR
LAMINAL	dental t̪	palatal c / tʃ / tɕ
APICAL	alveolar t	retroflex ʈ

3 Phonotactic constraints on retroflexion

- Distribution of consonants can be defined in relation to a $C_1VC_2C_3VC_4$ template:
  - $\{C_1, C_3\}$ = the set of strictly pre-vocalic (CV) positions
  - $\{C_2, C_4\}$ = the set of strictly post-vocalic (VC) positions
- Intervocalic consonants are both pre- and postvocalic (VC_5V); they are typically unconstrained and will not be discussed.
- Abstracting away from special exceptions (e.g., homorganic -CC- clusters, loanwords, etc.), we find the following typology of constraints on apicality/retroflexion:

(3) Two distribution patterns affecting apical/retroflex consonants

	C_1	V	C_2	C_3	V	C_4
Type A:	*		✓	*		✓
Type B:	✓		*	✓		*

3.1 Type A: No initial or other strictly prevocalic (CV) retroflexes

- In some languages, retroflex consonants are categorically banned or infrequent in word-initial and other strictly prevocalic positions $\{C_1, C_3\}$; they are overwhelmingly preferred in postvocalic positions $\{C_2, C_4\}$ (and intervocalically).
- Caveat: languages of this type may allow retroflex consonants in C_3 position if C_2 is also retroflex; i.e., in homorganic consonant clusters.
  - These are typically the result of progressive assimilation, diachronically if not synchronically (e.g., ʈt > ʈʈ, ɳt > ɳʈ).
- This is the most common phonotactic constraint on retroflexion cross-linguistically; some form of it occurs in Dravidian, Indo-Aryan and Australian languages.

3.1.1 Australian

- Australian languages contrast 2–4 coronal places of articulation.
  - Minimally: 1 laminal (dental or postalveolar), 1 apical (alveolar or retroflex).
  - Maximally: t̪, t, ʈ, c
- Hamilton (1996): Apicals (alveolar /t/, retroflex /ʈ/) are the least preferred segments in prevocalic $\{C_1, C_3\}$ positions, and the most preferred in postvocalic $\{C_2, C_4\}$ positions.
  - Some languages ban apicals altogether in prevocalic $\{C_1, C_3\}$ positions.
  - If they are allowed in prevocalic $\{C_1, C_3\}$ positions, then: (i) apicals are the least frequent segments in those positions; and (ii) most languages neutralize the alveolar/retroflex contrast in those positions (in favour of one or the other).

(4) Martuthunira: Distribution of coronals (Hamilton, 1996, p. 278)

	C_1	C_2	C_2C_3	C_4
/t̪, n̪, l̪/	t̪-, n̪-	—	-n̪t̪-	—
/t, n, l/	—	-nk-, -lk-	-nt-	-n, -l
/ʈ, ɳ, ɭ/	—	-ɳk-, -ɭk-	-ɳʈ-	-ɳ, -ɭ
/c, ɲ, ʎ/	c-, ɲ-	-ɲk-, -ʎk-	-ɲc-	-ɲ, -ʎ

3.1.2 Dravidian

- Proto-Dravidian:
  - Four coronal places of articulation: t̪, t, ʈ, c
  - No apicals (alveolar/retroflex) word-initially.
- Contemporary Dravidian languages:
  - Some South Dravidian languages preserve the 4-way coronal system with its phonotactic constraint on initial apicals.
  - Others have merged the alveolar series variously with the dental or retroflex series, but preserve the ban on word-initial retroflexes.
- In Dravidian, the constraint shows up mainly as a ban on word-initial apicals (C_1).
  - Medial consonant clusters are typically limited to homorganic geminate or NC sequences, so the distinction between C_2 and C_3 is often obscured.
  - Most words end in vowels. Stems ending in underlying consonants are often pronounced with an epenthetic "enunciative" vowel (typically [ɯ], [ɨ] or [u]), thereby eliminating (or at least reducing) occurrences of C_4.

(5) Tamil, Kanniyakumari dialect: Distribution of coronals (Christdas, 1988)

	C_1	C_2C_3	C_2C_3	C_4
/t̪/ [n̪-]	t̪-, [n̪-]	-tt-, [-n̪n̪-]	-[n̪]t̪-	—
/t, n, l/	—	-tt-, -nn-, -ll-	-nt-	-n, -l
/ʈ, ɳ, ɭ/	—	-ʈʈ-, -ɳɳ-, -ɭɭ-	-ɳʈ-	-ɳ, -ɭ
/c, ɲ/	c-, ɲ-	-cc-, -ɲɲ-	-ɲc-	—

- Tamil: dental [n̪] occurs as an allophone of alveolar /n/ word-initially, before /t̪/, and in derived geminates (e.g., [pun̪n̪el] 'rice that has just been harvested' < /puN + [n̪]ell/).

3.1.3 Indo-Aryan

- Old Indo-Aryan (OIA, Sanskrit) had a 3-way coronal system: dental, retroflex, palatal.
- Retroflex consonants did not occur word-initially (Schwarzschild, 1973; Masica, 1991).

- Most New Indo-Aryan languages now allow retroflex plosives initially, but still avoid initial retroflex sonorants (e.g., *ɳ-, *ɭ-, *ɽ-).

3.2 Type B: No final or other strictly postvocalic (VC) retroflexes

- In some languages, retroflex consonants are prohibited in final and other strictly postvocalic positions {C_2, C_4}; they are preferred in initial and other prevocalic positions {C_1, C_3} (and intervocalically).
- This pattern is typologically rare and has not received much attention. However, it is attested in some Tibeto-Burman languages, particularly those of the Tibetic group in the western Himalayas.

3.2.1 Tibeto-Burman

- Most Tibeto-Burman languages of the western Himalayas contrast dental and retroflex plosives. Many of these ban retroflex segments from syllable codas (i.e., {C_2, C_4}). [1]

(6) Lhomi: Distribution of coronals (Vesalainen & Vesalainen, 1976)

	C_1	C_2	C_2C_3	C_3	C_4
/t̪, t̪ʰ/	t̪-, t̪ʰ-	-t̪C- [-ʔC-]	-t̪t̪-	-Ct̪-	-t̪ [-ʔ]
/ʈ, ʈʰ/	ʈ-, ʈʰ-	—	—	-Cʈ-	—

- In Lhomi, dental /t̪/ is typically realized as glottal [ʔ] is syllable codas (see footnote 1).

(7) Lhomi: Dental and retroflex plosives in syllable onsets (i.e., prevocalic CV positions)

Dental /t̪/		Retroflex /ʈ/	
t̪á	'horse'	ʈák	'button'
sóp.t̪ok	'ring'	sìp.ʈok	'comb of chicken'
sà.t̪u	'to eat'	pʰí.ʈa	'wild cat'

(8) Lhomi: No retroflex consonants in syllable codas (i.e., strictly postvocalic VC positions)

Dental /t̪/		Retroflex /ʈ/
pèt̪	'is'	*pèʈ
sórit̪	'stomach (hon.)'	*sóriʈ
lít̪.maŋ	'he didn't come'	*líʈ.maŋ

[1] Coda consonants have been subject to glottalization or complete elision in some Tibeto-Burman languages. This obscures the distribution pattern to some degree because both dental and retroflex consonants may be absent in syllable codas in such cases. Nevertheless, historical-comparative evidence indicates that the dental series did occur in codas prior to glottalization/elision, whereas the retroflex series did not.

- The same pattern is attested in other Tibeto-Burman languages of the western Himalayas including: Tshangla (Andvik, 2003), Nar Phu (Noonan, 2003), Tamang (Mazaudon, 2003), Dolpo (Kevin Kopp, p.c.), Humla (Wilde, 2001), and Dolakha Newar (Genetti, 2007), among others.

4 Diachronic sources of retroflexion

- The two phonotactic patterns in (3) are contradictory: where retroflexes are avoided in one, they are preferred in the other, and *vice versa*.
- These contradictory patterns can be explained if phonotactic constraints on retroflexion are a direct result of the evolution of retroflexion in a language.
- The most common diachronic sources of retroflexion are listed in (9).

(9) Common diachronic sources of retroflexion (Bhat, 1973; Hamann, 2003; 2005)

a. retroflexion in liquid context e.g., rt / tr > ʈ
b. retroflexion in back vowel context e.g., ut / tu > uʈ / ʈu

- Liquids are the most common triggers, especially rhotics (e.g., English *cart* [kʰɑɹʈ]), but also laterals.
- The triggering liquid or back vowel typically precedes the targeted consonant (e.g., rt > ʈ), but in some cases it follows (e.g., tr > ʈ).

(10) Central hypothesis

Pattern	Constraint on Retroflexion	Source of Retroflexion
Type A	preferred in VC contexts avoided in CV contexts	progressive assimilation from a preceding liquid/back vowel (e.g., rt > ʈ; ut > uʈ)
Type B	preferred in CV contexts avoided in VC contexts	regressive assimilation from a following liquid (or back vowel?) (e.g., tr > ʈ)

4.1 Type A

4.1.1 Dravidian

- Retroflexion first emerged phonetically in the class of liquids and spread from liquids to following nasals and plosives through progressive assimilation across morpheme boundaries (Zvelebil, 1970, pp. 101–104, 174–175; Tikkanen, 1999; Levitt, 2010).

(11) Sources of retroflexion in (pre-) Proto-Dravidian

a. retroflexion of liquids: **l > *l, ɭ, ɻ
b. coronal assimilation: *l̪t̪ > *nt, *(t)t
*ɭt, ɻt > *ɳʈ, *(ʈ)ʈ

- Step 1: Original **/l/ developed a retracted/retroflex allophone [l̻], which in turn developed an approximant variant [ɻ̻]. Conditions are unclear, but similar developments are well-attested elsewhere, especially in *non-initial, postvocalic* positions:
  - English: /l/ → [ɬ] in codas (i.e., non-initial, mostly postvocalic positions)
  - Norwegian & Swedish dialects: /l/ → [ɬ] in postvocalic position (Hock, 1991, p. 79)
  - Indo-Aryan: MIA /-l-/ > NIA /-l̻-/; MIA /-ll-/ > /-l-/ (Masica, 1991, p. 193).
- Step 2: Progressive assimilation in heteromorphemic clusters */lt̻, l̻t̻, ɻt̻/ produced new apico-alveolar and retroflex segments from dental stops and nasals. Supported by:
  - Morphophonological alternations in old literary Tamil (Levitt, 2010, pp. 64–69).
  - Stem alternations in South Dr. languages (Zvelebil, 1970, pp. 101–104, 174–175).
- Step 3: Simplification of clusters with loss of the conditioning liquid.
- Result: 3-way dental-alveolar-retroflex contrast in stops, with the original dental series in both initial and non-initial positions and the new alveolar and retroflex series limited to non-initial, post-vocalic environments.

(12) Development of coronal contrasts in Proto-Dravidian

a.	Starting point:	t̻-	-(t̻)t̻-	-lt̻-	-l̻t̻-, -ɻt̻-
b.	Progressive assimilation:	---	---	-lt-	-l̻t̻-, -ɻt̻-
c.	Loss of conditioning liquid:	t̻-	-(t̻)t̻-	-(t)t-	-(t̻)t̻-

- Near-identical developments are independently attested in Central Norwegian and Swedish dialects, where /l/ → retracted/velarized [ɬ] in postvocalic position.

(13) Development of coronal contrast in Norwegian & Swedish dialects (Hock, 1991, p. 79)

a.	Starting point:	t̻, d̻	rt̻, rd̻	ɬt̻, ɬd̻
b.	Progressive assimilation:	---	rt, rd	ɬt̩, ɬd̩
c.	Loss of conditioning liquid:	t̻, d̻	t, d	t̩, d̩

4.1.2 Indo-Aryan

- The main internal developments are summarized in (14) (Misra, 1967, pp. 28–29, 63ff; Bhat, 1973; Hamp, 1996; Tikkanen, 1999).

(14) Sources of retroflexion in Proto-Indo-Aryan

a.	ruki:	IE s, z > IA ṣ, z̩ / r, u, k, i ___
b.	n-retroflexion:	IA n > ṇ / r, ṣ (…) ___
c.	coronal assimilation:	IA ṣt, z̩d > ṣṭ, z̩ḍ
d.	sibilant laxing:	IA z̩ḍ > iʳḍ, uʳḍ > iḍ, uḍ
e.	Fortunatov's law:	IE lt, ls, ln > l̻t̻, l̻s̻, l̻n > l̻t̻, l̻s̻, l̻ṇ > IA ṭ, ṣ, ṇ

- Retroflexion first emerged in sibilants following /r, u, k, i/ and spread to following nasals and plosives through progressive assimilation (14)(a)-(d).
- Fortunatov's law: retroflexion of liquid /l/ with progressive assimilation of following dentals (nearly identical to the Dravidian pattern).
- Later: OIA -rt̪-, -rd̪- > MIA -ʈʈ-, -ɖɖ- > NIA -ʈ-, -ɖ- (Masica, 1991, pp. 176, 187–188)
- Result: dental vs. retroflex contrast with dentals occurring initially and non-initially and retroflexes occurring only in non-initial environments.

4.1.3 Australian

- Proto-Australian is generally reconstructed with two coronal series, one laminal and one apical. The origin of the apical vs. laminal contrast is not generally addressed.
- However, it is widely recognized that the retroflex series developed from the apico-alveolar series after rhotics, and possibly after back vowels (Dixon, 1980; 2002).

4.2 Type B

4.2.1 Tibeto-Burman

- Retroflexion developed from coronal and non-coronal consonants before liquids, i.e., mostly from Cr- sequences (but also Cl- in some cases) (Bhat, 1973; Matisoff, 2003).
- The examples below demonstrate the correspondence between Cr sequences in Classical Tibetan and retroflex stops in contemporary Central Tibetan. Classical Tibetan reflects the phonological structure of the language in the seventh century CE (DeLancey, 2003).

(15) Retroflex plosives from Cr- sequences in Tibetan (data from Bhat, 1973, p. 34)

Gloss	Classical 'Written' Tibetan		Central Tibetan
'child'	phru-gu	>	ʈʰu-gu
'before'	drung-du	>	ʈung-du
'belly'	grod-pa	>	ɖʰö-pa

- Cr- and Cl- sequences occurred only in syllable onsets. Thus, the retroflex consonants that they produced are limited to onsets.

(16) Evolution of coronal contrast in Tibetan languages

a. Starting point:	t̪-	Cr-	-t̪
b. Regressive assimilation:	---	ʈr-	---
c. Loss of conditioning liquid:	t̪-	ʈ-	-t̪

4.3 Summary

- Type A languages, which prohibit retroflex/apical segments in initial and other strictly prevocalic positions {C_1, C_3}, are those that developed retroflexion from consonants *after* liquids/back vowels (e.g., Dravidian, Indo-Aryan, Australian).
- rC and lC sequences are typically non-initial and postvocalic because they constitute well formed syllable codas, but not well formed syllable onsets.
- Retroflexion of consonants after liquids/back vowels produces retroflexion only in non-initial, postvocalic {C_2, C_4} positions:
  - Vrt, Vlt > Vʈ
  - ut > uʈ
- Type B languages, which prohibit retroflex segments in final and other strictly postvocalic positions {C_2, C_4}, are those that developed retroflexion from consonants *before* liquids (e.g., Tibeto-Burman).
- Cr and Cl sequences are typically syllable initial and prevocalic because they constitute well formed syllable onsets, but not well formed syllable codas.
- Retroflexion of consonants before liquids produces retroflexion only in initial and other prevocalic onset {C_1, C_3} positions:
  - CrV, ClV > ʈV

5 Synchronic perceptual accounts of retroflex phonotactics

- Previous studies have posited synchronic constraints that are grounded in perception (Hamilton, 1996; Steriade, 2001; Hamann, 2003).
- The general hypothesis runs as follows:
  - Neutralization of a given contrast targets those positions in which the contrast, if realized, would be less salient.
  - The acoustic cues critical to apical/retroflex contrasts are most salient in VC transitions, and least salient in CV transitions (Ahmed & Agrawal, 1969).
  - Thus, apical/retroflex contrasts are avoided in strictly CV positions (where they lack the critical VC transitions); and they are preferred in VC positions (where they benefit from the cues in VC transitions).
- Problem: provides a plausible explanation of the dominant Type A pattern (Australian, Dravidian, Indo-Aryan); but it cannot explain the Type B pattern (Tibeto-Burman).
- Type B pattern: retroflex segments are restricted to those environments where their cues are least salient; and prohibited in those environments where their cues are most salient.
- The existence of the Type B pattern suggests that the perceptual account cannot b generalized into any kind of universal constraint ranking or implicational universal.

- E.g., Steriade (2001) posits the (universal?) "Law" of apical contrast and neutralization.

(17) Patterns of apical contrast and neutralization (Steriade, 2001, p. 226)
   a. **The Law**: if the t/ʈ contrast occurs in a language, it occurs after V
   b. **The General Case**: t/ʈ contrast only after V
   c. **The Initial Deviation**: t/ʈ contrast only after V and in #___
   d. **The I-Deviation**: t/ʈ contrast after central and back V; reduced to [t] after [i]

- Tibeto-Burman languages with the Type B pattern violate this law.
- The distribution of perceptual cues may still play an important role in the evolution of retroflex phonotactics; it may explain the cross-linguistic frequency of the Type A pattern over the Type B pattern.
- Apical/retroflex assimilation is expected to be predominantly progressive because progressive assimilation preserves the salient VC transitions at the expense of the less salient CV transition ($VC_2C_3V > VC_2C_2V$), whereas regressive assimilation does just the opposite ($VC_2C_3V > VC_3C_3V$).
- Thus, liquids/back vowels are more likely to induce retroflexion in a following segment (e.g., Vrt > Vʈ) than in a preceding segment (e.g., CrV > ʈV).

6 Conclusion

- Contradictory phonotactic patterns involving apical/retroflex consonants can be explained if the patterns are a direct result of the evolution of apical/retroflex contrasts in each language.
- Perceptual accounts of apical/retroflex phonotactics provide a plausible explanation of only the dominant Type A pattern; they leave the Type B pattern unexplained and unexpected.
- Perceptual factors may still play an important role: To the extent that they predict a natural bias toward progressive apical/retroflex assimilation over regressive assimilation, they may account for the cross-linguistic frequency of the Type A pattern over the Type B pattern.

References

Andvik, E. (2003). Tshangla. In G. Thurgood, & R. J. LaPolla (Eds.), *The Sino-Tibetan langauges* (pp. 439–455). London & New York: Routledge.

Bhat, D. N. (1973). Retroflexion: an areal feature. *Working Papers on Language Universals , 13*, 27-67.

Christdas, P. (1988). The phonology and morphology of Tamil. Doctoral dissertation. Ithaca, NY: Cornell University.

DeLancey, S. (2003). Classical Tibetan. In G. Thurgood, & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 255–269). London & New York: Routledge.

Dixon, R. M. (2002). *Australian languages: Their nature and development.* Cambridge: Cambridge University Press.

Dixon, R. M. (1980). *The languages of Australia.* Cambridge: Cambridge University Press.

Genetti, C. (2007). *A grammar of Dolakha Newar.* Berlin & New York: Mouton de Gruyter.

Hamann, S. (2005). The diachronic emergence of retroflex segments in three languages. *Link , 15* (1), 29-48.

Hamann, S. (2003). *The phonetics and phonology of retroflexes.* Utrecht: LOT (Landelijke Onderzoekschool Taalwetenschap).

Hamilton, P. (1996). *Phonetic constraints and markedness in the phonotactics of Australian aboriginal languages.* Doctoral dissertation, University of Toronto.

Hamp, E. P. (1996). On the Indo-European origins of the retroflexes in Sanskrit. *Journal of the American Oriental Society , 116* (4), 719–723.

Hock, H. H. (1991). *Principles of historical linguistics.* (2nd, Ed.) Berlin & New York: Mouton de Gruyter.

Jain, B. D. (1934). *A phonology of Panjābī as spoken about Ludhiana and a Ludhiānī phonetic reader.* Lahore: The University of the Panjab.

Levitt, S. H. (2010). Fortunatov's law and Dravidian. *International Journal of Dravidian Linguistics , 39* (1), 19–90.

Maddieson, I. (1984). *Patterns of sounds.* Cambridge: Cambridge University Press.

Masica, C. P. (1991). *The Indo-Aryan languages.* Cambridge: Cambridge University Press.

Matisoff, J. A. (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction.* Berkeley: University of California Press.

Mazaudon, M. (2003). Tamang. In G. Thurgood, & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 291–314). London & New York: Routledge.

Misra, B. G. (1967). *Historical phonology of Modern Standard Hindi: Proto-Indo-European to the present.* Doctoral dissertation, Cornell University.

Noonan, M. (2003). Nar Phu. In G. Thurgood, & R. J. LaPolla (Eds.), *The Sino-Tibetan Languages* (pp. 336–352). London & New York: Routledge.

Schwarzschild, L. A. (1973). Initial retroflex consonants in Middle Indo-Aryan. *Journal of the American Oriental Society , 93* (4), 482-487.

Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume, & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219-250). New York: Academic Press.

Tikkanen, B. (1999). Archaeological-linguistic correlations in the formation of retroflex typologies and correlating areal features in South Asia. In R. Blench, & M. Spriggs (Eds.), *Archaeology and language IV: Language change and cultural transformation* (pp. 138-148). London & New York: Routledge.

Vesalainen, O., & Vesalainen, M. (1976). *Lhomi phonemic summary.* Kathmandu: Summer Institute of Linguistics and Institute of Nepal and Asian Studies.

Wilde, C. P. (2001). Preliminary phonological analysis of the Limi dialect of Humla Bhotia. MA thesis. University of Helsinki.

Zvelebil, K. (1970). *Comparative Dravidian phonology.* The Hague: Mouton.