The diachronic emergence of retroflex segments in three languages*

Silke Hamann
ZAS (Centre for General Linguistics), Berlin
silke@zas.gwz-berlin.de

The present study shows that though retroflex segments can be considered articulatorily marked, there are perceptual reasons why languages introduce this class into their phoneme inventory. This observation is illustrated with the diachronic developments of retroflexes in Norwegian (North-Germanic), Nyawaygi (Australian) and Minto-Nenana (Athapaskan). The developments in these three languages are modelled in a perceptually oriented phonological theory, since traditional articulatorily-based features cannot deal with such processes.

1 Introduction

Cross-linguistically, retroflexes occur relatively infrequently. From the 317 languages of the UPSI database (Maddieson 1984, 1986), only 8.5 % have a voiceless retroflex stop [ʈ], 7.3 % a voiced stop [ɖ], 5.3 % a voiceless retroflex fricative [ʂ], 0.9 % a voiced fricative [ʐ], and 5.9 % a retroflex nasal [ɳ]. Compared to other segmental classes that are considered infrequent, such as a palatal nasal [ɲ], which occurs in 33.7 % of the UPSID languages, and the uvular stop [q], which occurs in 14.8 % of the languages, the percentages for retroflexes are strikingly low. Furthermore, typically only large segment inventories have a retroflex class, i.e. at least another coronal segment (apical or laminal) is present, as in Sanskrit, Hindi, Norwegian, Swedish, and numerous Australian Aboriginal languages. Maddieson's (1984) database includes only one exception to this general tendency, namely the Dravidian language Kota, which has a retroflex as its only coronal fricative.

Low frequency and the restriction to large inventories, among other factors such as late appearance in language acquisition and diachronic instability, are traditionally considered to be indices for the markedness of a segmental class, see Greenberg (1966, 1978) and Lass (1984). Segmental markedness is often explained by articulatory complexity (see Lindblom & Maddieson 1988, among others): a segmental class is marked because it departs more from a default mode of production than an unmarked counterpart. Applying this articulatorily based idea of markedness to the class of retroflexes, it could be argued that retroflexes are more marked than apical alveolars (or apical dentals), since retroflexes involve a raising and displacement of the tongue tip towards the post-alveolar region, whereas an apical alveolar involves only a tongue tip raising.

If retroflexes are indeed articulatorily marked, then the question arises how languages should employ such a marked category at all. Bhat (1973: 55) writes that in the majority of cases retroflexes are "introduced into a language mainly through the assimilatory influences of neighboring sounds such as back vowels, velar consonants, r, or at a latter stage by other retroflexed consonants." According to Bhat, retroflexes thus emerge via articulatory

* I would like to thank T. A. Hall for comments on an earlier version. Furthermore, I am thankfull to the audience of the the *3rd Uiltjes-Dagen* in Utrecht, the *11th Manchester Phonology Meeting*, the *16th International Congress of Historical Linguistics* in Copenhagen, especially Patrick Honeybone and Aditi Lahiri, and the ZAS phonology reading group, i.e. Laura Downing, Tonio Green, and Marzena Zygis. All errors remain my own.

assimilation to adjacent segments. However, Bhat's hypothesis fails to explain why languages are attested in which a front apical turns into a retroflex if it assimilates to a back vowel, a velar, or an r. This is problematic because the latter segments are articulatorily not closer to a retroflex than to a front apical. For instance the retroflexion of an apical in back vowel context would imply that retroflexes have a tongue body configuration that is closer to back vowels than front apicals. Since Bhat does not assume that retroflexes are inherently retracted (see Bhat 1974 who explicitly rejects this assumption), and front apicals can be retracted, i.e. velarized, and thereby be similarly close to back vowels, there is no obvious reason for such an assimilation.

The present study argues that diachronic developments introducing retroflexes are primarily perceptually motivated and illustrates this claim with three processes whereby languages introduced retroflex segments into their phoneme inventory. These processes and a sample language for each are given in (1).

(1)	a) retroflexion in rhotic context	rt > ʈ	Norwegian (North-Germanic)
	b) retroflexion in back vowel context	ut > uʈ	Nyawaygi (Australian)
	c) retroflexion via secondary labialization	k^w > ʈɽ	Minto-Nenana (Athapaskan)

It will be shown below that the three processes in (1) show a large perceptual similarity between their respective earlier and later forms. This and a gestural simplification that occurred in the later forms are argued to have caused the diachronic change. A phonological modelling of the processes in (1) and their conflicting perceptual and articulatory drives are represented in an Optimality-theoretic framework (Prince & Smolensky 1993, henceforth: OT). The diachronic developments in (1a) and (1b) are analyzed as a re-ranking of constraints, as done by Bermúdez-Otero (1999), Boersma (1998, 2003), Green (1997, 2001), Ham (1998), and Jacobs (1995). The development in (1c), however, is shown to be due to the child's association of acoustic cues to a relevant phonological feature that differs from the cues that the adult has associated with this feature.

Departing from traditional OT work, underlying perceptual representations are assumed, following Boersma (1998), since they provide a tool to model the need for perceptual similarity between earlier and later forms for processes that involve a change in the articulatory realization of the segments. Furthermore, a distinction between articulation and perception grammar is made, also following Boersma.

The study is structured as follows. In section 2, the alleged articulatory and perceptual markedness of retroflexes will be discussed. Section 3 gives the data for the three processes under consideration. In the following section, perceptual motivations in diachrony in general and in the three processes in particular are elaborated. The last section concludes.

2 Alleged markedness of retroflexes

This section looks at evidence and counterevidence for the alleged markedness of retroflexes, dealing separately with articulatory complexity (section 2.1) and perceptual salience (section 2.2).

2.1 Articulatory complexity

As elaborated above in section 1, retroflexes can be considered articulatorily more marked than apical alveolars since they involve a displacement of the tongue tip towards the post-alveolar region in addition to the tongue tip raising that is also present in apical alveolars. However, x-ray tracings of the voiceless retroflex fricative in Standard Chinese by Ladefoged & Wu (1984), of the same sound class (voiced and voiceless) in Polish by Puppel et al. (1977), and of the same voiceless sound in Tamil by Svarný & Zvelebil (1955) illustrate that retroflex fricatives do not necessarily involve a displacement of the tongue tip. What is more, the tongue tip is often not even discernible as active articulator in these sounds. This observation lead Hamann (2003b) to the conclusion that displacement of the tongue tip cannot be a necessary criterion in the definition of retroflex sounds. Based on a careful comparison of x-ray data of retroflex sounds in numerous languages, Hamann introduces tongue back retraction besides a raising of the tongue tip and a posterior place of articulation (i.e. behind the alveolar ridge) as defining criteria for retroflex sounds.

Since front apicals are not articulated behind the postalveolar region and only optionally retracted, they can still be assumed to be articulatorily less complex and therefore less marked than retroflexes.

2.2 Perceptual weakness

In the phonetic literature, the term 'weak feature' is often used to denote a feature that is perceptually less salient than others and is detected by the auditory system only after the so-called robust or primary features (Stevens & Keyser 1989). In this sense, weak features are perceptually marked, a generalization that is attested by Ohala's (1993: 89) remark that the distinction between robust vs. weak features correlates nicely with the way segment inventories in languages are constructed: "those with a small number of phonemes use the robust features almost exclusively; those with many phonemes use the same robust features but also exploit weaker, slower features". According to this definition, retroflexion should be considered perceptually weak or marked since it occurs rarely and even then only in large inventories, as elaborated in section 1. If we look at the perceptual cues for retroflexes, the question arises how exactly these weak cues should look like that differentiate retroflexes from front apicals or laminal coronals.

Compared to other coronal sounds, retroflex sonorants and stops are distinguished by very low third formant values (F3) and accordingly low F3 vowel transitions into and out of the segment, as discussed in Hamann (2003 chapter 3). Besides this prominent cue, retroflex closures are usually shorter than those of other coronal articulations (see Anderson & Maddieson 1994 for Tiwi coronal stops and McDonough & Johnson 1997 for Tamil laterals). Retroflex fricatives show different cues than stops and sonorants, they are characterized by a far lower starting point of spectral friction noise and a lower first peak than other coronals (Hamann 2003; see also Zygis & Hamann 2003 for centre of gravity measurements of Polish that support this point).

Following the argumentation by Stevens & Keyser and Ohala layed out above, low F3 is the distinguishing and therefore weak feature of retroflex stops and sonorants. Furthermore, this feature should be perceived by the listener only after the stronger, primary feature of coronality (Stevens & Keyser 1989: 81), which is realized by mid to high second formant values (F2). However, acoustic studies showed that listeners process general formant patterns and do not perceive single formants separately, see the experiments with synthesized speech by Cooper et al. (1952) and Liberman (1957).

Since all sibilants are differentiated by noise location (and formant transitions), there is no evidence that the listener processes the low location of noise for retroflex fricatives later than any other place of noise for other fricatives, though it might be argued that listeners parse the friction noise first, and only then determine the place of articulation.

Perception tests prove that retroflexes are detected as easily and quickly as any other place of articulation. This is at least the case for speakers of a native language that either has a retroflex phoneme (such as Hindi, see Ohala & Ohala 2001), or no retroflex phoneme or allophone at all (such as German, see Hamann 2003a). Native speakers of American English with its retroflex allophones of 'r' do show problems with discriminating retroflex tokens from dental ones (Werker et al. 1981), a fact probably due to their unawareness of having a retroflex allophone, but their performance could be largely improved by specific perceptual training (Pruitt et al. 1990).

A further point that has to be discussed with respect to the alleged perceptual markednesss of retroflexes is the asymmetric spread of retroflex cues. Most retroflex manners (such as stops, taps, laterals) show already a retraction of the tongue tip during a preceding vowel. At the time of the closure phase, the tongue tip is fronted, i.e. the retroflex gesture is released. For this reason, the preceding vowel is often slightly retroflex and provides additional, early cues to the following retroflex segment. The transitions from the retroflex into the following segment (CV cues), however, are often identical to front apical transitions, due to the fronting of the tip. This observation was used by Steriade (1995, 2001) to explain the restriction of retroflexes to prevocalic position (where their strong VC cues are perceivable) in a number of languages. The availability of retroflex cues as early as in the preceding vowel is a further argument against their alleged late perception and weakness.

Lastly, some languages introduced retroflex segments in their inventory via areal spread (Bhat 1973: 55). Areal spreading of this class supports the observation that retroflexes do not have weak perceptual features, since languages are expected to borrow segments with strong perceptual features (compare for instance the areal spread of the click sounds in Bantu languages) rather than segments with weak features.

In sum, no perceptual evidence could be found that the cues of retroflexion are harder to detect than those of any other place of articulation, thus no evidence exist for retroflex cues being weak.

3 Three diachronic changes

In the following section, three diachronic changes are investigated that introduced retroflexes. This is done by looking at one sample language for each process. For further examples of retroflexion in a rhotic context (discussed for Norwegian in section 3.1) and retroflexion in a back vowel context (discussed for Nyawaygi in section 3.2), see Bhat (1973) and Hamann (2003). Languages other than Minto-Nenana (in section 3.3) which introduce retroflexion via secondary rounding could not be found.

The three processes are by no means meant to cover all diachronic developments of retroflexes. Further developments involve for instance retroflexion of velar segments and voiced (implosive) stops, and retroflexion via areal spread.

3.1 Norwegian: retroflexion of rhotic and front apical sequences

In Urban East Norwegian, orthographic forms of <r> plus <t, d, s, n, l> in monomorphemic words are realized by the respective retroflex segments [ʈ, ɖ, ʂ, ɳ, ɭ], see the examples in (2)a) (from Kristoffersen 2000). For a discussion why the retroflexes in these words are assumed to be underlying, see Hamann (2003: 84f.). Historically, these forms used to be pronounced as sequences of rhotic plus alveolar, see Haugen (1982).

(2) *Norwegian*

a)	kart	[kɑʈ]	‘map’		
	kors	[kɔʂ]	‘cross’		
	barn	[bɑːɳ]	‘child’		
	sardin	[sɑ.ˈɖiːn]	‘sardine’		
b)	bror+s	[bɾuːʂ]	‘brother’ + possessive	cf. [bɾuːɾ]	bror
	vår+dag	[ˈʋoː.ˌɖɑːg]	‘spring day’	cf. [ʋoːɾ]	vår
	Per ser	[ˌpeː.ˈʂeːɾ]	‘Per sees’	cf. [peːɾ]	Per

Taking the retroflex voiceless stop as example, the following development can be assumed:[1]

(3)	*Old Scandinavian*		*Norwegian*
	*/ɾt/	→	/ʈ/

The examples in (2)b) illustrate that besides the diachronic retroflexion rule, there is also an active process of retroflexion across morpheme and word boundaries.

3.2 Nyawaygi: retroflexion of front apicals in back vowel context

Proto-Australian is assumed to have had one apical and one laminal series of stops, nasals, and laterals, the apical being articulated at the alveolar ridge (Dixon 1980). In a number of daughter languages, the apical series developed retroflex allophones after back vowels, as is still the case in the languages from Eastern Queensland (Dyirbal, Yidiny, etc.). Languages from Western Australia (such as Western Destert and Walmatjari) futher developed two separate apical series from these alveolar and retroflex allophones (Dixon 1980: 155).

The Eastern Queensland language Nyawaygi is one that retained the single apical series but which developed a retroflex flap allophone [ɽ] from the stop /d/ before an /u/ in intervocalic position (Dixon 1983: 449f.). This is observable when compared to its neighbour language Wargamay, which has preserved the original /d/, see (4)a).

(4)	*Nyawaygi*	*Wargamay*	*gloss*
a)	wuɽu	wudu	‘nose’
	giɽul	gidul	‘cold’

[1] For possible intermediate stages such as /ɽt/ and /ɽʈ/ and their analysis, see Hamann (2003: 87f. and 174f.).

b)	garala	gadala	‘dry’
	baːri	baːdi	‘to cry’
c)	ruŋan	duŋan	‘stinging tree’

In Nyawaygi, the apical stop changed to an apical trill before /a/ and /i/ (the only other vowels besides /u/), see the examples in (4)b), and also word-initially, see the example in (4)c). The two outcomes of Pre-Nyawaygi */d/ are stated in (5).

(5)	*Pre-Nyawaygi*		*Nyawaygi*	
a)	*/d/	→	/ɽ/	/ V_u
b)	*/d/	→	/ɾ/	

Since the present article is concerned with the development of retroflexes, the process in (5)b) will not be discussed in the following sections.

3.3 Minto-Nenana: retroflexion via secondary rounding

In the Athapaskan language Minto-Nenana spoken in Alaska (Krauss 1962, Tuttle 1998), retroflex fricatives (voiced and voiceless) and affricates (voiced, voiceless and voiceless ejective) were introduced via secondarily rounded segments. The resulting segment classes have rhotic releases, see the examples in (6)a) for the fricatives and in (6)b) for the affricates, where the first column gives the orthographic form and the second a phonetic transcription of the underlined graphemes.

(6)			
a)	srisr	[ʂ^{ɽ}]	‘sheefish’
	zren	[ʐ^{ɽ}]	‘black’
b)	tr’axa	[ʈ^{ɽ}’]	‘woman’
	tretr	[ʈ^{ɽ}]	‘nail; dry wood’
	dran	[ɖ^{ɽ}]	‘day’

The retroflex affricate series (6)b) in Minto-Nenana stems from Proto-Athapaskan segments which are assumed to be either rounded velar stops (Krauss 1973, 1979), rounded alveolar stops (Tharp 1972) (derived from the Proto-Athapaskan-Eyak */k^{w}/), or rounded postalveolar affricates (Cook & Rice 1989a, Cook & Rice 1989b, Krauss & Leer 1981), see (7)a) – c) respectively.

(7)	*Proto-Athapaskan*		*Minto-Nenana*
a)	*/k^{w}/, */k^{w}’/	→	/ʈ^{ɽ}/, /ʈ^{ɽ}’/
b)	*/t^{w}/, */t^{w}’/	→	/ʈ^{ɽ}/, /ʈ^{ɽ}’/
c)	*/t͡ʃ^{w}/, */t͡ʃ^{w}’/	→	/ʈ^{ɽ}/, /ʈ^{ɽ}’/

The present study follows the view in (7)c) that the pre-Minto-Nenana stage had rounded postalveolar affricates, since this is most widely spread among Athapaskan linguists (see the discussion in Cook & Rice 1989).

The Proto-Athapaskan fricatives */ʃ^{w}/ and */ʒ^{w}/ changed to the retroflex fricatives /ʂ^{ɽ}/ and /ʐ^{ɽ}/ with rhotic releases, respectively, see the development in (8).

(8) *Proto-Athapaskan* *Minto-Nenana*

*/$ʃ^{w}$/	□	/$ʂ^{ɽ}$/
*/$ʒ^{w}$/	□	/$ʐ^{ɽ}$/

In contrast to Minto-Nenana, the neighbouring language Chena (also called Lower Tenana) did not develop retroflex segments, but seems to have collapsed the rounded postalveolars with the Proto-Athapaskan plain alveolars ([s, z, t͡s]), compare the Minto-Nenana words with their Chena cognates in the second column in (9) (Kari & Tuttle 1993).

(9)

Minto-Nenana	*Chena*	*gloss*
$ʂ^{ɽ}$a	sa	‘sun/moon’
dəlk’ə$ʂ^{ɽ}$	dəlk’əs	‘it is red’
nə$ʐ^{ɽ}$unh	nəzu	‘it is good’
$ʈ^{ɽ}$’exe	t͡s’exe	‘woman’
$ʈ^{ɽ}$əx	ət͡səx	‘he cries’ vs. ‘I’m crying’

4 Perception in diachrony

Perceptually-based diachronic changes have not received much attention with respect to formal analysis in the phonological literature (but see e.g. Flemming 2002). In the phonetic literature, they are treated as some kind of misperception or misparsing of cues on the listener’s side. Ohala (1993: 89f.), for example, divides such misperceptions into dissociation parsing errors, where cues are not associated with a segment, and false association parsing errors, where cues of one segment are associated with another segment. A schematic representation for the false association of rhotic cues to the apical coronal yielding a retroflex is given in figure 1, based on a general schema by Ohala (1993: figure 7.1).

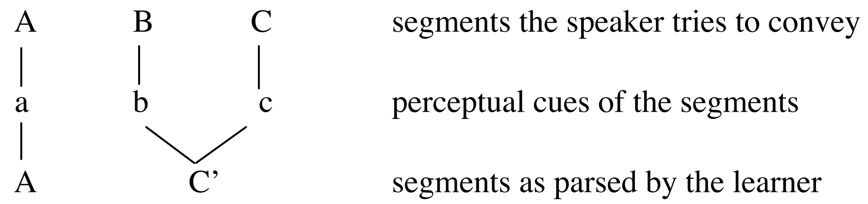

Figure 1 An example for a misparsing error yielding a retroflex segment (C’), with A being a vowel (with its vowel cues a), B a rhotic (with its lowered formant cues b), and C a front apical (with its cues c).

The reason for the neglect of perceptually-motivated developments in phonological theories is that traditional phonological descriptions take articulatory features as their primitives and describe articulation, only; an underlying articulatory form /x/ is realized as a surface articulatory form [x] which has a perceptual output. The acquisition of an underlying form /y/ via a perceptual output of the underlying form /x/ cannot be accounted for, even if the two forms /x/ and /y/ are perceptually similar. Such a change is thus ascribed to non-phonological processes, see figure 2.

UF /articulatory/		/x/	
↓	*Phonology & Phonetic implementation*	↓	
SF [articulatory]		[x]	
↓	automatic process	↓	
acoustic form		x	
↓	perceptual decoding	↓	misperception
UF /articulatory/		/y/	

Figure 2 Standard assumption of the phonological model on the left, and the misperception of an acoustic form which can lead to a different underlying representation (UF) on the right. Italics indicate modules of the grammar. SF = surface form.

The proposal made here departs from this traditional phonological view of perceptually-based sound changes in several ways. First of all, it assumes that speech perception is part of the phonological module and must therefore be modelled in a separate perception grammar as proposed by Boersma (1998 et seq.) in his Functional Phonology model. An overview of the model as relevant for the present article is given in figure 3, based on Boersma (1998: 144).

UF |perceptual|	
↓	*production grammar*
OF [articulatory]	
↓	automatic process
[acoustic]	
↓	*perception grammar*
SF /perceptual/	

Figure 3 Boersma's Functional Phonology model. UF = underlying form, OF = overt form, and SF = surface form. Italics indicate modules of the grammar.

As can be seen from figure 3, the underlying representation of phonological forms is perceptual, i.e. the primitives assumed in Boermsa's model are perceptual features.

Besides a production and perception grammar, Boersma also assumes a recognition grammar, which applies in the model in figure 3 to the perceptual surface form (last row), and guides the listener in constructing a perceptual underlying representation with the help of her lexical knowledge. All three phonological modules, i.e. articulatory, perception, and recognition grammar in Boersma's approach are modelled in Optimality Theory (McCarthy & Prince 1993, Prince & Smolensky 1993; henceforth: OT).

Assuming Boersma's model of phonology, sound changes can now be understood to happen in two ways. The first possibility starts off with a change on the part of the speaker: due to factors such as speech style, speech rate, etc., the speaker shows variation in the output forms via a reranking of constraints in her production grammar. Variation of forms can also be due to allophonic distribution, which does not require a constraint reranking. Taken on its own, output variation or alternation does not constitute a diachronic change in a language, it has to be accompanied by a reanalysis on the side of the learning child. From the alternating surface forms, the child constructs a different underlying representation than the one the former generation had established. The child which acquires this different grammar then contributes to more occurrences of the new form, which again leads further children to post the new form as underlyingly. If this process is continued, the old form will occur less and less often and an increasing number of learners will acquire the new form, see Kiparsky's (1968) assumption that diachronic change is the result of a child grammar that survives and spreads throughout the speech community. The sound change is complete if there are no speakers with an underlying old form anymore, and the actual speakers will not produce any instances of the old form as variant of their underlying new form, anymore.

The second possibility for sound change in Boersma's model sets in at an earlier age than the deduction of underlying forms, namely at the point when the child constructs phonological categories. For a change to happen, the child creates a segmental category with cues that are present in the acoustic signal but are not used as distinguishing cues for this category by the adults, or they use slightly different but perceptually very similar cues to the adult ones'. This 'wrong' category is not corrected at a later stage if its produced form is recognized and accepted by the speech community and does not endager any lexical distinctions. An example of the creation of a different category is the class of alveolar fricatives /f, v/ for the British English dentals /θ, ð/ (Sweet 1874), which are perceptually nearly identical in their frequency noise (Shadle et al. 1992).[2] This new creation started a sound change which is spreading from London to the adjacent counties (Trudgill 1999: 79).

Single speakers that acquire different categories than the majority of the speech community do not trigger a sound change, but add to variant input forms for the next generation of learners. This variation then might lead again to the construction of different underlying forms, which in the long run can create a sound change as elaborated above.

In the following subsections, I describe and model the variation of surface forms in the production grammars of Norwegian and Nyawaygian (sections 4.1 and 4.2, respectively) and the malacquistion of a category at an early stage of the perception grammar in Minto-Nenana, see section 4.3.

4.1 Norwegian retroflexion

According to Haugen (1982), retroflexion in a rhotic context as described in section 3.1 occurred in the transition from Old to Modern Scandinavian. Before looking at the actual diachronic process, we first have to clarify the articulation of the relevant segmental classes at that time. The rhotic /r/ in Old Scandinavian is described as a trill (Haugen 1982: 64), whereas the respective sound in Modern Norwegian is realized as a tap (Kristoffersen 2000: 24). The present study assumes that the Old Scandinavian rhotic in coda position that merged with the following coronal was realized as tap or flap, since a trill is very unlikely to have merged with another segment due to its strong perceptual cues and its more specific articulatory requirements (for the former see Lindau 1985, for the latter Recasens 1991). The front coronals /t, d, s, n, l/ in Old Scandinavian are described as apical in Haugen (ibid.), the actual place of articulation is not explicitly mentioned, though presumably alveolar. In present-day Norwegian, this class is laminal denti-alveolar (Kristoffersen 2000: 22), the change of articulator is probably a reaction to the introduction of the apical postalveolar series. The

[2] Note that this change lead to a merger of two different phoneme classes, namely /θ d/ and /f v/.

resulting retroflex category in Modern Norwegian is not a subapical postalveolar or palatal as the retroflex segments e.g. in Dravidian languages, but a less retracted apical postalveolar variant (see Simonsen et al. 2000). A detailed description of these segments in Old Scandinavian could not be found.

The present article proposes that in the synchronic grammar of Old Scandinavian, speakers displayed some variation in the output of rhotic – alveolar sequences, i.e. they produced both faithful forms with rhotic plus alveolar sequences and unfaithful forms with retroflexes only. The retroflex forms were possible due to some articulatory reduction: instead of two gestures of rhotic and alveolar, only one gesture, a retroflex, had to be articulated. This articulatory reduction is accepted, as there is a perceptual similarity between underlying and surface form: both rhotic flaps and retroflexes share a low third formant (Lindau 1985, Stevens 1998). The low F3 of the rhotic was reassigned to the apical coronal, which was consequently interpreted as retroflex, recall section 2.2 above, where low F3 was shown to be the distinctive feature for retroflexion. All cues relevant for the process are given in table 1 with the example of apical stops:

ɾ	t	ʈ
lowF3	midF3	lowF3
	burst	burst
short closure	long closure	long closure

Table 1 The perceptual cues for the segments involved in Norwegian retroflexion, with the example of an apical stop and a resulting retroflex stop.

While the rhotic and the retroflex stop share the low F3, the apical and retroflex stops share the same manner cues of burst and long closure (for other manners of articulation this also holds, though the cues look differently).

In order to model the production grammar with its varying stop output forms in an OT framework, we use Boersma's (1998: 187f.) *DELETE as faithfulness constraints to the underlying perceptual features. A definition of this constraint is given in (10)a).

(10) a) *DELETE (*feature*: *value*):
"An underlyingly specified *value* of a perceptual *feature* appears (is heard) in the surface form."
b) *DELETE (F3: low) or *DELETE (lowF3)
c) *DELETE (long closure & burst) = *DELETE (stop)
d) *DELETE (short closure & low F3 & low F2 & voicing) = *DELETE (ɾ)

The present study assumes that the perceptual cue [low F3] is interpreted as a phonological feature; it is shared both by apical rhotics and retroflexes, and its reassociation occurred across a whole segmental class and resulted in a new class. The *DELETE constraint in (10)b) has this feature as argument. Furthermore, it is assumed that combinations of features make up new features, thus [long closure] and [burst] combine to [stop], which is the argument of the constraint in (10)c). Lastly, features can define a phoneme, which in turn can be the argument of a *DELETE constraint, see (10)d).

As an articulatory markedness constraint, Boersma's (1998: 150) *DISTANCE is employed, see the definition in (11)a). For Norwegian retroflexion, a *DISTANCE (manner) constraint is needed, as defined in (11)b).

(11) a) *DISTANCE (*articulator*: *a* | *b*): "An *articulator* does not move from location *a* to *b*, away from the neutral position."

b) *DISTANCE (*tongue tip*: $manner_1$ | $manner_2$) = *DISTANCE (manner): "The *tongue tip* does not move from location $manner_1$ to $manner_2$ (e.g. from flap to stop)."

A production grammar for Old Scandinavian that produces faithful outputs with a rhotic and alveolar stop is given in the tableau in (12), with the example *kart* 'map'. Before looking at this tableau in detail, some shorthand conventions have to be elaborated. In a Functional Phonological production grammar, the input is the lexically stored perceptual representation of the word in question, represented in pipes |*spec*|. The output candidates consist of two forms each, namely an articulatory and a corresponding perceptual form. The articulatory output is given in brackets [*art*] and the corresponding perceptual output in slashes /*perc*/, all following Boersma (1998: 143ff.) and used in figure 3 already. The articulatory markedness constraints (here *DISTANCE (manner)) apply to the articulatory form [*art*], whereas the perceptual faithfulness constraints (*DELETE) refer to the corresponding perceptual form /*perc*/.

In the tableau in (12), all faithfulness constraints are ranked above the markedness constraint *DISTANCE (manner), thus the most faithful candidate wins.

(12)

|kart|	*DELETE (low F3)	*DELETE (stop)	*DELETE (ɾ)	*DISTANCE (manner)
☞ [kaɾt] /kaɾt/				*
[kaʈ] /kaʈ/			*!	
[kaɾ] /kaɾ/		*!		
[kat] /kat/	*!		*	

The assumed variation in the output forms of Old Scandinavian can only be modelled, if the candidate with the retroflex form also wins in some instances, which is the case if a reranking of *DISTANCE (manner) above *REPLACE (ɾ) takes place. This reranking embodies Green's (1997, 2001) idea on diachronic change as Promotion of the unmarked, an extension of the theory of the Emergence of the Unmarked (McCarthy & Prince 1994), where a higher ranking of the relevant markedness constraint results in a less marked output form.[3]

According to Boersma (1998) and Boersma & Hayes (2001), reranking of constraints within one grammar is possible if the two constraints under question are ranked close together in the constraint hierarchy, see figure 4, where the constraints are assigned a (hypothetical) ranking value along a continuous scale. A random stochastic element (noise) in constraint

3 An alternative to reranking is the assumption of partial constraint-ranking, as in the free-variationist model (Anttila 1995, Anttila & Cho 1998). See Green (2001: 9ff.) for a critical discussion of this view.

evaluation leads to the lower constraint being ranked above the higher constraints in a certain number of cases, depending on the closeness of the two constraints, see Boersma (1998: 330ff.) for the calculation of the distribution.

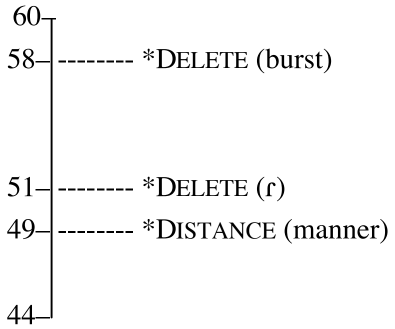

Figure 4 Partial constraint hierarchy: the constraints with hypothetical ranking values that result in surface alternation.

With a reversed ranking of *DELETE (ɾ) and *DISTANCE (manner), the retroflex stop would win, see the tableau in (13).

(13)

|kart|	*DELETE (low F3)	*DELETE (stop)	*DISTANCE (manner)	*DELETE (ɾ)
[kaɾt] /kaɾt/			*!	
☞ [kaʈ] /kaʈ/				*
[kaɾ] /kaɾ/		*!		
[kat] /kat/	*!			*

Thus, the two tableaus in (12) and (13) model the variation in the output forms, which might lead a learning child to postulate an underlying retroflex form instead of the original rhotic plus alveolar.

4.2 Nyawaygi retroflexion

Nyawaygi developed a rhotic trill /r/ from the Proto-Australian voiced stop /d/, as illustrated in section 3.2 above. This rhotic has a retroflex allophone before a back high vowel [u]. The present analysis will not deal with the change in manner from stop to trill/flap, but focus on the alternation between front apical and retroflex rhotic. An articulatory motivation for the development of a retroflex allophone in [u] context is that the tongue back position of a retroflex is very close to that of a back vowel [u], see Hamann (2003: 92f.). This can be formalized by a markedness constraint *DISTANCE (as defined in (11)a)) referring to the tongue back position, see (14).

(14) *DISTANCE (*tongue back*: $\textit{position}_1$ | $\textit{position}_2$) = *DISTANCE (back):
“The *tongue back* does not move from location $\textit{position}_1$ to $\textit{position}_2$ (e.g. from *front* for an apical rhotic to *retracted* for a back vowel).”

The articulatory assimilation of rhotic to back vowel has to be perceptually acceptable. We saw for the apical rhotic in Norwegian already, that front apical rhotics are characterized by the cues low F3 and low F2. The retroflex flap has an even lower F3 than the apical rhotic, see Ladefoged & Maddieson's (1996: 240f.) x-ray data of both segments in the Australian language Warlpiri. Butcher (1992) observes that Australian retroflexes are rather back in their place of articulation, i.e. postalveolar to palatal, which results in a lower F3 than retroflexes that are less back, e.g. the Norwegian ones. Thus both back vowel and retroflex share the cue of a very low F3, see the summary of the relevant cues in table 2, where manner cues are summarized as the features [trill], [flap] and [vowel].

r	ɽ	u
lowF3	lower F3	lowerF3
trill	flap	vowel

Table 2 The relevant perceptual cues and features for the segments involved in Nyawaygi retroflexion.

A faithfulness constraint necessary for the following modelling of retroflexion in Nyawaygi is *DELETE (trill), according to the general *DELETE constraint defined in (10)a).

The production of the standard form before allophonic variation, i.e. of the apical trill in back vowel context, is given in a Functional Phonological production grammar in tableau (15), with the assumed pre-Nyawaygi example */girul/ 'cold'.

(15)

|girul|	*DELETE (trill)	*DISTANCE (back)
☞ [girul] /girul/		*
[giɽul] /giɽul/	*!	

Variation between the winner in (15) and the alternate form with a retroflex can occur if the constraints *DELETE (trill) and *DISTANCE (back) are ranked very closely together and can reverse their ranking due to the random stochastic element in the constraint evaluation, recall the description of *DELETE (ɾ) and *DISTANCE (manner) for Norwegian retroflexion in section 4.1 above. The relevant tableau with the same example as in (15) is given in (16), where the markedness constraint *DISTANCE (back) is ranked high.

(16)

|girul|	*DISTANCE (back)	*DELETE (trill)
[girul] /girul/	*!	
☞ [giɽul] /giɽul/		*

Interestingly, this ranking does not trigger changes in other vowel contexts, since the remaining two vowels in Nyawaygi, [a] and [i], have a tongue back position that is not opposing the one of the apical rhotic, therefore the tongue back does not have to cover a distance and no violation of *DISTANCE (back) occurs, see the example /baːri/ 'to cry' in tableau (17).

(17)

|baːri|	*DISTANCE (back)	*DELETE (trill)
☞ [baːri] /baːri/		
[baːɽ̥i] /baːɽ̥i/		*!

It can be assumed that present learners of Nyawaygi acquire the constraint ranking as in tableau (16), since their output forms do not show any variation between retroflex flap and apical trill in back vowel context.

This example illustrated the diachronic acquisition of a retroflex allophone which is articulatorily motivated and perceptually acceptable.

4.3 Minto-Nenana retroflexion

The case of retroflexion in Minto-Nenana involves the change from secondarily rounded postalveolars to retroflexes with rhotic releases, recall the data in section 3.3. This process differs from the ones described above in two respects: firstly, to my knowledge it is attested in no other language (neither diachronically nor synchronically), and secondly, it does not involve an articulatory motivation, since no gestures are saved or articulatory distances shortened in the new form compared to the old one.[4] This second point of a missing articulatory simplification is the reason why the change in Minto-Nenana cannot be caused by variance in the production grammar. If a speaker acquired the underlying form of a laminal postalveolar fricative and affricate with secondary rounding and the respective articulatory realizations of these in her production grammar, no variation can occur if there is no articulatory motivation for it.

Instead, such a change must be assumed to have occurred at the point in acquisition where cues are associated with segmental categories. For a modelling of the (correct) construction of segmental categories in the perception grammar of Boersma's model, see Boersma et al. (2003). In the case of Minto-Nenana, some learners constructed cues that are perceptually very similar but not identical to the original cues. This development is illustrated in figure 5 exemplarily with the voiceless fricative (a) and the voiceless affricate (b).

(a)	ʃʷ	□	ʂ̥^{ɽ}
	low friction noise lowered CV transitions glide noise at CV transitions		lower friction noise lowered CV transitions friction noise at CV transitions

---

[4] In Hamann (2003c) I claim that the gesture of secondary rounding can be saved in the new output and therefore the process is articulatorily motivated. I do not follow this argument here because retroflexes are often optionally accompanied by secondary rounding. Furthermore, the new output requires a sequentiality of gestures (stop or fricative plus rhotic frication), whereas in the old form the rounding could be articulated simultaneously with the primary articulation, which indicates a more complex timing of gestures in the new form, contradicting an aritculatory simplification.

(b) t͡ʃʷ	□	$\text{ʈ}^{\text{ɽ}}$
stop closure		stop closure
fricative release		fricative release
low friction noise		lower friction noise
lowered CV transitions		lowered CV transitions
glide noise at CV transitions		friction noise at CV transitions

Figure 5 Cues of the adult categories, on the left, and the constructed child's cues, on the right, with the example of voiceless fricative (a) and voiceless affricate (b).

Both the fricative and the affricate show only few difference between the adult category and the newly acquired one: whereas the postalveolar segments on the left side of figure 5 are characterized by low friction noise and glide noise at the CV transitions, the retroflex segments on the right side have even lower friction noise and also friction noise at the CV transition. Since these are minor perceptual differences, the 'wrong' retroflex category was accepted by the speech community and spread until it eventually replaced the original postalveolar category.

5 Conclusion

The three diachronic processes of retroflexion that were illustrated here are based on the perceptual similarity between earlier and later forms, specifically the low F3 that characterizes retroflexes. Retroflexion in Norwegian and Nyawaygi involve articulatory optimizations that are licensed via this perceptual similarity. Retroflexion in Minto-Nenana, on the other hand, has no articulatory motivation, but is caused by a misacquisition of a phonological category in the learning process. Again, this misacquisition is accepted due to the perceptual similarity between early and late form.

Traditional phonological approaches with underlying articulatory representations cannot deal with such perceptual similarity. For this reason, the Functional Phonology approach by Boersma was applied, which operates with underlying perceptual features and separate articulation and perception grammar, and thus is able to give accounts for all three processes. Furthermore, the present article showed that articulatory markedness is irrelevant for an account of the diachronic development of retroflexes in Norwegian, Nyawaygi and Minto-Nenana.

References

Anderson, Viktoria B. & Ian Maddieson. 1994. Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics* 87: 131-162.

Anttila, Arto. 1995. Deriving variation from grammar: A study of Finnish genitives. Ms., *Stanford University.*

Anttila, Arto & Young-Mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104: 31-56.

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change (Opacity and globality in phonological change), University of Manchester: PhD. dissertation.

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

Bhat, D. N. S. 1974. Retroflexion and retraction. *Journal of Phonetics* 2: 233-237.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45-86.

Boersma, Paul. 2003. The odds of eternal optimization in OT. In: Eric Holt (ed.) *Optimality Theory and language change*, Dordrecht: Kluwer; 31-65.

Boersma, Paul, Paola Escudero & Rachel Hayes. 2003. Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In: *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona; pp. 1013-1016.

Butcher, Andrew. 1992. The phonetics of neutralisation: the case of Australian coronals. In: Jack Windson (ed.) *Studies in General and English Phonetics*, London: Routledge; 10-38.

Cook, Eung-Do & Keren Rice. 1989a. Introduction. In: Eung-Do Cook & Keren Rice (eds.) *Athapaskan Linguistics: Current perspectives on a language family*, Berlin: Mouton de Gruyter; 1-62.

Cook, Eung-Do & Keren D. Rice eds. 1989b. *Athapaskan Linguistics. Current Perspectives on a Language Family*. Berlin: Mouton de Gruyter.

Cooper, F. S., P. C. Delattre, A. M. Liberman, J. M. Borst & L. J. Gerstman. 1952. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America* 24: 597-606.

Dixon, R. M. W. 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1983. Nyawaygi. In: R. M. W. Dixon & Barry Blake (eds.) *Handbook of Australian Languages*, Amsterdam: Benjamins; 431-527.

Flemming, Edward. 2002. *Auditory Representations in Phonology*. London: Routledge.

Green, Antony Dubach. 1997. The prosodic structure of Irish, Scots Gaelic, and Manx, Cornell University: PhD dissertation.

Green, Antony Dubach. 2001. The Promotion of the Unmarked: Representing sound change in Optimality Theory. Ms., *University of Potsdam.*

Greenberg, Joseph H. 1966. *Language Universals*. 'S-Gravenhage: Mouton.

Greenberg, Joseph H. 1978. Some generalizations concerning initial and final consonant clusters. In: Joseph H. Greenberg (ed.) *Universals of Human Language. Volume 2: Phonology*, Stanford: Stanford University Press; 243-279.

Ham, W. H. 1998. A new approach to an old problem: gemination and constraint reranking in West Germanic. *Journal of Comparative Germanic Linguistics* 1: 225-262.

Hamann, Silke. 2003a. Is retroflexion licensed by VC cues only? Observations from Norwegian. *Nordlyd* 31.1: 63-77.

Hamann, Silke. 2003b. *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT.

Hamann, Silke. 2003c. Why not bend our tongue tips backwards! Three diachronic developments of retroflex segments. Talk presented at the *11th Manchester Phonology Meeting*, Manchester.

Haugen, Einar. 1982. *Scandinavian language structures: A comparative historical survey*. Tübingen: Niemeyer.

Jacobs, Haike. 1995. Optimality Theory and Sound Change. *Proceedings of the annual meeting of the North East Linguistic Society (NELS)* 2: 219-232.

Kari, James & Siri Tuttle. 1993. Robert McKennan's Tanana Vocabularies. Ms. In author's possession.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In: E. Bach & R. T. Harms (eds.) *Universals in Linguistic Theory*, New York: Holt; 170-202.

Krauss, Michael E. 1962. Minto Phonology. Ms., *Alaska Native Language Center*.

Krauss, Michael E. 1973. Na-Dene. In: T. Sebeok (ed.) *Linguistics in North America. (Current Trends in Linguistics, Vol. 10)*, The Hague: Mouton.

Krauss, Michael E. 1979. Na-Dene and Eskimo-Aleut. In: L. Campbell & M. Mithun (eds.) *The Languages of Native North America: An Historical and Comparative Assessment*, Austin: University of Texas Press; 803-901.

Krauss, Michael E. & Jeff Leer. 1981. Athabaskan, Eyak and Tlingit Sonorants. *Alaska Native Language Center Research Papers* 5.

Kristoffersen, Gjert. 2000. *The Phonology of Norwegian*. Oxford: Oxford University Press.

Ladefoged, Peter & Zongji Wu. 1984. Places of articulation: an investigation of Pekingese fricatives and affricates. *Journal of Phonetics* 12: 267-278.
Lass, Roger. 1984. *Phonology: an Introduction to Basic Concepts*. Cambridge: Cambridge University Press.
Liberman, A. M. 1957. Some results of research on speech perception. *Journal of the Acoustical Society of America* 29: 117-123.
Lindau, Mona. 1985. The story of /r/. In: Victoria Fromkin (ed.) *Phonetic Linguistics: Essays in Honour of Peter Ladefoged*, Orlando (Florida): Academic Press; 157-168.
Lindblom, Björn & Ian Maddieson. 1988. Phonetic universals in consonant systems. In: Larry Hyman & C. N. Li (eds.) *Language, Speech and Mind: Studies in Honor of Victoria A. Fromkin.*, London: Routledge; 62-78.
Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
Maddieson, Ian. 1986. The size and structure of phonological inventories: Analysis of UPSID. In: J. Ohala & J. Jaeger (eds.) *Experimental Phonology*, Orlando (Florida): Academic Press; 105-123.
McCarthy, John J. & Alan Prince. 1993. Prosodic morphology I: Constraint interaction and satisfaction. Ms., *University of Massachusetts and Rutgers University.*
McCarthy, John J. & Alan Prince. 1994. The emergence of the unmarked: optimality in prosodic morphology. *Proceedings of the annual meeting of the North East Linguistic Society (NELS)* 24: 333-379.
McDonough, Joyce & Keith Johnson. 1997. Tamil liquids: an investigation into the basis of the contrast among five liquids in a dialect of Tamil. *Journal of the International Phonetic Association* 27: 1-26.
Ohala, John J. 1993. The percetual basis of some sound patterns. In: B. A. Connell & Amalia Arvaniti (eds.) *Phonology and phonetic evidence: Papers in Laboratory Phonology IV*; 87-92.
Ohala, Manjari & John J. Ohala. 2001. Acoustic VC transitions correlate with degree of perceptual confusion of place contrast in Hindi. In: N. Grø & J. Rischel (eds.) *Travaux du cercel Linguisticque de Copenhague. Vol. XXXI.*, Copenhagen: Reitzel; 265-284.
Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Ms., *Rutgers University*. Brunswick.
Pruitt, J. S., Winifred Strange, Linda Polka & M. C. Aguilar. 1990. Effects of category knowledge and syllable truncation during auditory training on Americans' discrimination of Hindi retroflex-dental contrasts. *Journal of the Acoustical Society of America* 87: S72 (Abstract).
Puppel, Stanisl-aw, Jadwiga Nawrocka-Fisiak & Halina Krassowska. 1977. *A Handbook of Polish Pronunciation for English Learners*. Warszawa: Pan'stwowe Wydawnictwo Naukowe.
Recasens, Daniel. 1991. On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics* 19: 267-280.
Shadle, Christine, Andre Moulinier, C. U. Dobelke & Celia Scully. 1992. Ensemble averaging applied to the analysis of fricative consonants. In: *Proceedings of the International Conference on Spoken Language Processing*, University of Alberta; pp. 53-56.
Simonsen, Hanne G., Inger Moen & Steve Cowen. 2000. Retroflex consonants in Norwegian: are they really? Evidence from EMA and EPG. In: *Proceedings of the 5th Seminar on Speech Production*, Zeon, Germany; pp.?
Steriade, Donca. 1995. Positional neutralization. Ms., *University of California*. Los Angeles.
Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume & Keith Johnson (eds.) *Phonology: The Role of Speech Perception in Phonology*, San Diego (CA): Academic Press; 219-250.
Stevens, Kenneth & Samuel Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65: 81-106.
Stevens, Kenneth. 1998. *Acoustic Phonetics*. Cambridge (Massachusetts): MIT Press.
Svarny', O. & Kamil Zvelebil. 1955. Some remarks on the articulation of the 'cerebral' consonants in Indian languages, especially in Tamil. *Archiv Orientalni* 23: 374-407.
Sweet, Henry. 1874. *History of English sounds*. London: Trübner.
Tharp, George. 1972. The position of Tsetsaut among the Northern Athapaskans. *International Journal of American Linguistics* 38: 14-25.
Trudgill, Peter. 1999. *The dialects of England*. Oxford: Blackwell.

Tuttle, Siri. 1998. Metrical and Tonal Structures in Tanana Athabaskan, University of Washington: PhD dissertation.

Werker, Janet F., J. H. V. Gilbert, K. Humphrey & R. C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52: 349-353.

Zygis, Marzena & Silke Hamann. 2003. Perceptual and acoustic cues of Polish coronal fricatives. In: *Proceedings of the 15th International Conference of Phonetic Sciences*, Barcelona; pp. 395-398.