How do voiced retroflex stops evolve?
Evidence from typology and an articulatory study

Silke Hamann
Utrecht University, Utrecht, The Netherlands

Susanne Fuchs
Zentrum für Allgemeine Sprachwissenschaft, Berlin

The present article illustrates that the specific articulatory and aerodynamic requirements for voiced but not voiceless alveolar or dental stops can cause tongue tip retraction and tongue mid lowering and thus retroflexion of front coronals. This retroflexion is shown to have occurred diachronically in the three typologically unrelated languages Dhao (Malayo-Polynesian), Thulung (Sino-Tibetan), and Afar (East-Cushitic). In addition to the diachronic cases, we provide synchronic data for retroflexion from an articulatory study with four speakers of German, a language usually described as having alveolar stops. With these combined data we supply evidence that voiced retroflex stops (as the only retroflex segments in a language) did not necessarily emerge from implosives, as argued by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983). Instead, we propose that the voiced front coronal plosive /d/ is generally articulated in a way that favours retroflexion, that is, with a smaller and more retracted place of articulation and a lower tongue and jaw position than /t/.

1   Introduction

Retroflex segments are often understood as articulations that involve a bending backwards of the tongue tip (see, e.g., Trask 1996, p.308). This narrow definition excludes segments in a large number of languages that are traditionally described as retroflexes, such as the postalveolar fricative in Mandarin (see Ladefoged & Wu 1984). For this reason, the present study defines retroflexion as an articulation with the tongue tip (apical) or tongue underside (subapical or sublaminal) against the alveolar, postalveolar or palatal region, following Ladefoged & Maddieson (1996). This broader definition

includes segments such as the postalveolar fricatives in Russian and Polish, whose retroflex status is debatable (see the discussion in Hamann 2004). The tongue tip raising of retroflex articulations requires a flattening of the tongue middle, which co-occurs with a retraction of the tongue back (the retraction of the tongue back is argued to be a general property of retroflexes by Hamann 2002, 2003; but see Bhat 1974, and Flemming 2003).

The complexity of gestures involved in the articulation of retroflexes might be the reason why this segmental class occurs relatively seldom in the languages of the world. For instance, only an estimated 11 percent of all languages have a retroflex stop (Ladefoged & Bhaskararao 1983, p.292). Furthermore, retroflexes occur only in larger coronal inventories, no language is known to us in which retroflexes are the only coronals.[1]

In his thorough study on retroflexes, Bhat (1973) discusses several diachronic processes that introduced this articulatorily complex class into languages. He mentions assimilatory influences of adjacent back vowels, rhotics, and velar consonants, but also introduction of a single voiced retroflex /ɖ/ via voiced dental implosives (p.55). For the latter, Bhat refers to Greenberg (1970), though the explanation given by Greenberg (p.129) actually comes from Haudricourt (1950): Voiced dental implosives are quite often retracted, which can lead to a retroflex implosive and eventually to a pulmonic retroflex stop.

Ohala (1983, p.200) also describes a development of a voiced retroflex stop from a voiced apical implosive (also referring to Greenberg), and furthermore elaborates that this process has an aerodynamic cause: Voiced apical stops can maintain voicing longer if the tongue body is lowered during the closure, and since the tongue lowering goes together with a retraction of the tongue tip, it is argued to result in retroflexion. This aerodynamic explanation is obviously independent from implosion, though Ohala does not make this point explicit. We propose in the present article that *both* implosives and retroflexes can develop from voiced (not voiceless) front coronal stops, and argue that several factors (mostly also based in aerodynamics, but also articulatory and perceptual requirements) are responsible for this development.

The view taken by Haudricourt (1950), Greenberg (1970), Bhat (1973) and Ohala (1983) is depicted in (1), with implosion as the only possible development of retroflex voiced stops in a language with no other retroflexes (note that the intermediate step of a retroflex implosive is not mentioned by Bhat and Ohala).

---

[1] Maddieson (1984) lists Kota as having only one sibilant, namely a retroflex voiceless [ʂ], which can therefore be interpreted as a counterexample to the statement that retroflexes always occur with other coronals. Emeneau (1944), the original source for Maddieson's classification, however, describes this sound as [s], in free variation with [tʃ], which is realized as retroflex only adjacent to other retroflexes (see also Flemming 2003, p.354).

(1)   ɗ (> ɗ̢) > ɖ̢

Our alternative view is represented in (2), where both retroflexion (a) and implosion (b) can develop from voiced stops. This does not preclude the possibility that some retroflexes originated from apical implosives, as in (1).

(2) a)   d > ɖ̢
     b)   d > ɗ

Focus of the present study is the emergence of retroflex sounds from voiced stops proposed in (2a), though we come back to the development of implosives from plain voiced stops in sections 2.3 and 5 below. Evidence for our proposal in (2a) comes from diachronic developments of retroflexes in a number of languages where no intermediate stage of implosion is reported. Furthermore, we illustrate with articulatory data from German that there are general differences in place of articulation and tongue and jaw height between voiced and voiceless alveolar stops favouring retroflexion of /d/. Both the diachronic and the articulatory evidence support the phonetic naturalness of the process in (2a), which makes an intermediate stage of implosion redundant.

The present article is structured as follows. Section 2 elaborates on the articulatory and aerodynamic characteristics of voiced coronals, especially the similarities between plain stops, retroflexes and implosives. In section 3, we discus three typologically unrelated languages that have [ɖ] as only retroflex. Section 4 provides synchronic data from German, and section 5 is the conclusion.

2   Voiced coronal stops

To provide evidence for the claim that voiced but not voiceless front coronal stops are prone to develop into retroflexes and also into implosives, we first look at the articulatory differences between voiced and voiceless front coronal stops (§2.1), including possible explanations for this difference. We then compare the characteristics of voiced front coronals with those of retroflex stops (§2.2) and coronal implosives (§2.3). The last subsection (§2.4) discusses explanations and examples for developments of retroflexes via implosion.

In the following, we do not distinguish between dental and alveolar coronal stops but summarize them under the term 'front coronals'. Furthermore, we focus on segments in intervocalic position, for the following two reasons. First, we usually find fully voiced segments in this position (Keating 1984), which allows us to compare across languages without having to pay attention to the

actual realisation of the voicing contrast. And second, the intervocalic position is a location where all of the segmental types that we compare, that is, front coronals, retroflexes, and coronal implosives, can occur (note for instance that retroflex segments do not occur in initial position in a number of languages, see Steriade 2001; and Hamann, 2003, pp.114-118).

2.1 Front coronal voiced stops

Studies on a variety of languages have shown that there are systematic differences between the articulation of voiced and voiceless front coronal stops. /d/ is usually realized with a more posterior position of the tongue tip and thus a more backed place of articulation than its voiceless counterpart, see for instance the electropalatographic studies by Dixit (1990) on Hindi, Moen & Simonsen (1997) on English and Norwegian, and Farnetani (1989, 1990) on Italian. In all of these studies we can also observe a smaller amount of tongue palatal contact and more contextual variation for /d/ than for /t/. A further systematic difference lies in the active articulator: /t/ is often articulated with the tongue blade, whereas /d/ is usually produced with the tongue tip (see, e.g., the x-ray data by Dart 1991, 1998, on French and English), though this only holds for languages that have a single series of coronal stops. Some studies found a stronger tongue pressure against the palate during the closure of /t/ and deduce from this a higher tongue position for /t/ (e.g., Wakumoto, Masaki, Honda & Ohue 1998 and Fujimura, Tatsumi & Kagaya 1973 for Japanese). Others showed that /d/ is produced with a lower jaw position than /t/ (e.g., Fujimura & Miller, 1979 for American English; Dart, 1991, for French; and Mooshammer, Hoole & Geumann 2006, 2007 for German). A further observation is that voiced /d/ is usually shorter than its voiceless counterpart (Stevens, Keyser & Kawasaki 1986, p.432).

Several explanations have been proposed for the observed differences between voiced and voiceless front coronal stops. The first and most commonly given is the aerodynamic requirement for voicing. Vibration of the vocal folds is only possible when there is a pressure difference between the subglottal and the intraoral cavity. Such a transglottal pressure difference can easily be produced with an open vocal tract. However, during the production of plosives, the vocal tract is closed for a certain time, resulting automatically in an increase of intraoral pressure. In order to maintain voicing during oral closure, as required for thoroughly voiced stops, it is necessary to enlarge the oral cavity (either actively or passively). Mechanisms of cavity enlargement for /d/ are manifold and include for instance a change from tongue blade to tongue tip, a lowering of the tongue, the jaw or the larynx, and an extension of the cheeks (Perkell 1969; Bell-Berti 1975; Westbury 1983; for German see Fuchs, 2005). Recall from

section 1 that cavity enlargement was Ohala's explanation for diachronic changes from alveolar implosive to voiced retroflex stop.

A second explanation for the difference between /t/ and /d/ is also based on voicing requirements. Because the transglottal pressure difference can only be maintained for a certain time unless actively maintained (see the mechanisms of cavity enlargement discussed above), voiced stops have mostly a shorter duration than their voiceless counterparts, the latter having in principle no restriction on the length of their closure. The shorter duration of /d/ can then account for all other above-mentioned differences with /t/ in the following way. It has been argued that for coronal stops the tongue tip or blade is aiming at reaching a target somewhere above the constriction location (Fuchs, Perrier & Mooshammer 2001, 2006; Löfqvist & Gracco 2002), since no exact location is necessary compared to the precise positioning required for sibilants or trills. Voiced coronal stops cannot fully reach this target because they have only little time to do so, and this so-called target undershoot (Lindblom 1963) results in a lower tongue and jaw position and in a more variable articulation.

The third explanation discussed here is again grounded in aerodynamics. Voiceless stops have a greater oral pressure than voiced ones (both mean and peak pressure; see Ladefoged & Maddieson 1996, p.96), since the airflow is not arrested by the vibrating vocal folds. Consequently, they require a firmer closure at the place of articulation than voiced ones. Following Ladefoged & Maddieson (1996) we can argue that the articulatory characteristics of /t/ described above, which correlate with a more forceful articulation than for /d/, might be "an anticipation of this need to make a firmer seal" (p.96).

A last account for the articulatory difference between /t/ and /d/ proposed in the literature is that voiceless stops require a more salient burst than voiced ones. This prominent burst is an important perceptual cue to distinguish voiceless from voiced coronal stops (Lisker & Abramson 1964; Repp 1979). The higher intra-oral pressure required for such a salient burst can be achieved by a higher tongue and jaw position. Furthermore, the use of the lower teeth as a second noise source can enhance the strength of the burst and is also only possible with a high tongue and jaw position. With respect to the jaw, Mooshammer et al. (2006; 2007) found a high and stable jaw position for /t/ in German. For /d/, on the other hand, the jaw was positioned lower, which gives the tongue more freedom to move and to accommodate to the context.

Most of these explanations cannot be evaluated separately. Thus the less salient burst and the less forceful seal result both in a generally lower articulatory effort for /d/, and so does target undershoot. Only the mechanism of cavity enlargement predicts an additional active control of gestures for /d/. If the lowering of tongue and jaw were actively controlled then we would expect the voiced /d/ to show less contextual variation and to be more stable in its

articulation than /t/ (see Mooshammer et al., 2006, p.22, for a similar argumentation). This is, however, not what we find in the literature. Instead, we saw that /d/ shows a much higher variability, and hence the tongue and jaw position of /d/ are less tightly controlled than that of /t/. We can therefore exclude cavity enlargement as explanation for the difference between /t/ and /d/. The remaining three explanations can only indirectly account for the difference in place of articulation between /d/ and /t/, namely via the assumption that apical articulations are preferably alveolar and laminal ones preferably dental (Ladefoged & Maddieson 1996:20-21).

We will see in the following section that the difference between /t/ and /d/ in articulation and duration makes the voiced stop prone to change into a retroflex.

2.2   Retroflex voiced stops

Retroflexes are articulated with a raised and retracted tongue tip, that is, they are always apical or subapical, with a place of articulation between the alveolar and palatal region. The raising and retraction of the tongue tip requires a lowering of the tongue middle and a retraction of the tongue back (see introduction). Though tongue lowering usually goes together with jaw lowering, we could not find any explicit mentioning of a low jaw position for retroflexes in the literature. Retroflex segments seem also to be shorter than other consonants, see for instance Anderson & Maddieson's (1994) study on Tiwi coronal stops, where the closure duration of retroflex segments was the shortest of all coronal consonants.

Retroflex articulations in general are described as being strongly context-dependent and showing large variability due to vowel coarticulation (see Švarný´ & Zvelebil 1955; Ladefoged & Bhaskararao 1983; Dixit 1990; Dixit & Flege 1991; Krull, Lindblom, Shia & Fruchter 1995; Simonsen, Moen & Cowen 2000). Most of these studies show that retroflexes are articulated furthest back (and thus most retroflex-like) in /u/ context, and furthest front (i.e., most front coronal-like) in /i/ context. Phonological studies have shown that retroflexes often avoid /i/ context, since the two have antagonistic tongue gestures (Flemming, 2003; Hamann, 2003:94-107). The context of /u/, on the other hand, has been reported to cause retroflexion of front coronals (Bhat, 1973; Hamann, 2003:90-94), as /u/ has a similar lowered tongue middle and retracted (and raised) tongue back. The emergence of retroflexes in Australian languages is, for example, ascribed to backing of front coronals in /u/ context (Dixon 1980).

A difference between voiced and voiceless retroflex stop similar to that between voiced and voiceless front coronal stop discussed above is expected,

though we found only little work that was explicit on this point. Dixit (1990), for example, observed that the voiced retroflex stop has a narrower constriction than its voiceless counterpart, and a palatographic study by Khatiwada (2007) shows that the voiced retroflex stop in Nepalese is articulated further back and with more contextual variation than the voiceless one.

Apicality, lowered tongue middle, short duration, and strong contextual variation are characteristics that retroflex voiced stops share with the voiced front coronal stop /d/, see section 2.1 above. This strong similarity between a voiced front coronal stop and a voiced retroflex leads us to propose that the two can be considered endpoints on a continuum from no or few to a large amount of retroflex characteristics, supporting Dixit's (1990:190) observation that retroflexion is not so much a place of articulation than a descriptive term for a particular tongue shape. This tongue shape occurs sometimes together with a dental or alveolar place of articulation. This front-back retroflex continuum of voiced coronal stops supports our claim that a /d/ can develop into a /ɖ/ without an intermediate stage of implosion, by a slight shift along this continuum.

2.3 Implosives

Whereas voiced front coronal and voiced retroflex plosives differ from each other exclusively in their place of articulation, voiced coronal implosives are articulated quite differently. They can be defined by three successive articulatory stages, namely glottal closure (plus a closure along the supralaryngeal cavity), larynx lowering, which results in rarefaction of the air between the two closures, and an implosive release, where the pressure is equalized (Catford 1939). Implosive consonants are always stops and can be voiced and voiceless, but voiceless implosives are extremely rare in the languages of the world.

Though implosives are produced with an ingressive airstream, the voiced ones allow simultaneous pulmonic egressive airflow. According to Laver (1994:179), the egressive air is "not enough to overcome completely the rarefaction of the enclosed volume of air in the vocal tract caused by the descending larynx." Catford (1977a:75) proposes on the basis of cineradiographic films that there is no active pulmonic airflow in voiced implosives, and the airflow that causes the vocal fold vibration comes actually from the downwards movement of the larynx against a static pulmonic pressure.

Ladefoged (1964) describes three possibilities for producing implosive sounds, namely first the aforementioned larynx lowering with ingressive airflow at release, second a sound with laryngealized voicing, and third a preglottalized sound. These possibilities can be transcribed for instance for alveolars as [ɗ], [d̰]

and [ˀd], respectively. Ladefoged proposes that all three possibilities should be considered variants of one category, based on the following four arguments. First, the real implosive type of articulation often co-occurs with laryngealized voicing, as for instance in Hausa. Second, Ladefoged (1964:60) finds it difficult to consistently distinguish between the laryngealized and preglottalized variants. Third, some Mayan languages show positional variations of implosives, with the real implosive articulation in initial position, and preglottalized sounds intervocalically. And finally, no language has a phonemic contrast between any of these three, according to Ladefoged. This leads Ladefoged to summarize all three articulations under the category 'injective'. Clements & Osu (2003) use a similar cover-category, but employ the term 'nonexplosive stops'.

A summary of the three articulations as 'implosive' is questionable in the light of the fact that there are African languages contrasting two of the three articulatory possibilities for implosives listed by Ladefoged. Clements & Osu (2003) show in a phonetic study that the Niger-Congo language Ikwere (of the Igbo family) has a phonemic contrast between a bilabial voiced implosive and a bilabial voiced, glottalized implosive.[2] We therefore employ the term 'implosive' in the following to refer only to the real implosive articulation of this class, and not to preglottalized or laryngealized stops.

If we compare the characteristics of an implosive to that of a plain voiced stop articulated at the same place – coronal for our purposes – the two seem to differ in the movement of the larynx and the direction of the airflow, only: the implosive shows a lowering of the larynx and ingressive air at the release. Unfortunately, even the class of implosives that fall within the restricted definition employed here do not always display these two characteristics. Clements & Osu (2003) found that none of the Ikwere implosives is realized with larynx lowering, although these sounds show ingressive airflow. Similarly, Lex (2006) illustrates that the implosives in the Fouladou dialect of Fula, another branch of the Niger-Congo languages, do not always have an ingressive airflow (see also Ladefoged, 1964). Ordinary voiced stops, on the other hand, often can be accompanied by larynx lowering, for instance in English and French (Ewan & Krones 1974). These and similar findings lead Ladefoged (1964, 1971) and Ladefoged & Maddieson (1996, p.82) to suggest the difference between plain voiced stops and implosives is gradient, "lying primarily in the comparatively larger and more rapid descent of the glottis in implosive" (Ladefoged, 1971:27).

---

[2] Goyvaerts (1986) mentions a possible contrast between voiced implosives and preglottalized sounds in the East Nilo-Saharan language Lendu. Dimmendaal (1986) and Demolin (1988) argue against such a contrast since the phonetically preglottalized sounds in Lendu are phonemic sequences of glottal and plain stops.

From this we can conclude that larynx lowering and ingressive airflow are no reliable characteristics of implosives. Whether a sound in a language is categorized as (alveolar) implosive therefore depends very much on the definition of implosive employed by the linguist. For instance, all Chadic languages have the implosives /ɓ/ and /ɗ/ (see Schuh 2003). These are usually glottalized, which is the reason why they are often simply described as glottalized or laryngealized stops in the literature on Chadic, as pointed out by Clements & Rialland (2005:20).

Ladefoged's (1964 et seq.) idea that implosives without ingressive airflow form a gradient continuum with plain voiced stops, with no clear boundary between the two categories, is similar to the continuum we proposed for alveolar and retroflex articulations in section 2.2. Whereas the plain-retroflex continuum is one that differs in tongue shape, this plain-implosive continuum differs in amount and velocity of glottis lowering. The two are thus orthogonal to each other and create a two-dimensional space, including a gradient continuum from plain to implosive retroflex, but neglecting the dimension of ingressive airflow. We will come back to this proposed space in the general discussion in section 5.

2.4 Developments of retroflexes from implosives

On the affinity between retroflexes and implosives, Greenberg (1970:129) noted that an implosive corresponding to a non-implosive dental in a language is often "retracted to the alveolar or alveopalatal position and is consistently apical, often with accompanying retroflexion". Haudricourt (1950) explains that the negative air pressure (due to the larynx lowering) causes a vacuum which tends to suck in the mobile tongue tip. Such retracted implosives then tend to lose their glottalic feature, a development repeated in (3a).

(3) a) ɗ > ɗ̢ > ɖ̢   Haudricourt, Greenberg
     b) ɗ > ɖ̢   Bhat, Ohala

The proposal by Bhat (1973) and Ohala (1983) in (3b) does not include an intermediate retroflex implosive, and Ohala's explanation for the development in (3b) does not refer to the negative air pressure of implosives. Instead he proposes that retroflexion is caused by cavity enlargement, where the tongue tip is retracted due to a lowering of the tongue. However, we saw in section 2.1 above that an account based on cavity enlargement does not require an implosive in the initial stage.

Let us look at languages supporting the two assumptions in (3). Greenberg bases his proposed development of voiced retroflex stops primarily on Tucker and Bryan's (1966) description of the retroflex implosives in Moru-Madi, a

branch of East Central Sudanic languages of the Nilo-Saharan family. For these sounds, "the retroflex tongue position is in fact a more distinguishing feature than the manner of articulation, which hardly seems implosive at all" (Tucker & Bryan 1966:102). This indicates a variation between retroflex implosive and voiced retroflex stop at the time of description. However, Moru-Madi languages have an additional phonemic retroflex voiced stop (see Watson 1991, in general; Demolin & Goyvaerts 1986, for Madi; Andersen 1987, for Lulubo; and Bender 1992, for a reconstruction of the contrast in Proto-Central-Sudanic), which makes a realisation of the voiced implosive as retroflex and thus a neutralisation between the two phonemes unlikely (though not impossible).

The development in (3b) is better documented. It occurred, for instance, in the Gbe languages (e.g., Fon, Ewe, Maxi) of the Niger-Congo Kwa family (Bantu), see the comparative study by Stewart (1995). Interestingly, the change in Gbe was preceded by a change in Bantu, were the coronal implosive is usually assumed to be a reflex of Proto-Bantu *d (Clements & Rialland 2005, p.21, Guthrie 1967-1971).

In the following section, we provide evidence for diachronic changes from front coronal to retroflex stops from three unrelated language families. Together with developments of implosives from plain stops as just elaborated for Bantu this illustrates that implosion and retroflexion can both be independent developments, supplementing the proposals by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983).

3   Languages with retroflexed voiced stops only

The data for the diachronic development of retroflex voiced stops comes from three typologically unrelated languages or language groups, namely the Malayo-Polynesian Dhao (§3.1), the Sino-Tibetan language Thulung (§3.2), and the East Cushitic languages Afar, Somali and Rendille (§3.3).

3.1 Malayo-Polynesian: Dhao

Dhao, also called Ndao, Dao, Ndaonese or Ndaundau, is a Central Malayo-Polynesian language, subsumed under the Bima-Sumba subgroup (Gordon 2005).[3] It is spoken on Ndao, and partly on Rote and Timor; all three are islands in the Sabu Sea of Indonesia. Dhao has the coronal stops /t d ɖ ɗ/, where the

---

[3]   The subgroup of Bima-Sumba languages is based on the classification by Jonker (1896; 1903; see also Esser, 1938) and has been criticised for its lack of evidence in terms of shared innovations (see Ross 1995:83). Fox (2004:7-8) argues for a more fine-grained distinctions between the languages of Sumba, those of Bima and Manggarai, and a separate subgroup of Sabu and Ndao.

retroflex is released with frication (Grimes, 2006:4)[4]. The closely-related Sabu (or Sawu(nese), Hawu, Havu) is spoken on the neighbour-islands of Sawu. Sabu has implosives, but no retroflexes, and its coronal stops are /t d ɗ/. Ngad'a, a further Bima-Sumba language, is spoken on Westflores and has like Sabu only implosives but no retroflexes (Arndt 1933, Klamer 1998), and the same holds for its neighbouring languages Lio and Kambera (Baird 2002).

The retroflex in Dhao corresponds to a plain stop in cognate words of Sabu, and the plain voiced stop to a palatal implosive; full correspondences between Dhao and Sabu voiced coronal stops are given in (4) (from Grimes, 2006:8).

(4)   Dhao   Sabu
    a)   ɖ   d
    b)   d   ʄ
    c)   ɗ   ɗ

The retroflex stop in Dhao and the Sabu alveolar stop (both 4a) are assumed by Grimes (2006) to stem from a voiceless alveolar, retroflex or palatal stop in Proto-Malayo-Polynesian (PMP). Evidence for the reconstruction of a voiceless segment for these voiced sounds comes from the fact that the sounds in (4c) correspond to voiceless stops in neighbouring languages (Jonker, 1903:86). The exact place of articulation of the PMP sound is difficult to determine and depends to a large extent on what has been reconstructed for Proto-Austronesian. For the purpose of the present article we can summarize Grimes' assumption that Dhao developed voiced retroflex stops from voiceless coronal stops, and not from implosives. Whether this development went via an intermediate stage of voiced front coronal stop is open to speculation.

Interestingly, the alveolar implosives in Dhao and the neighbouring Sabu in (4c) are assumed to have developed from a retroflex or palatal voiced stop in PMP (see Grimes 2006), as depicted in (5). Most authors (e.g., Dempwolff, 1934; Dyen 1971; Ross, 1992) assume a voiced and a voiceless retroflex stop in PMP, whereas others (such as Wolff 1974, 1991) propose palatal stops instead.

(5)   a)   * ɖ > ɗ
      b)   * c > ɗ

If the change did take place as in (5a), then we would have a reversal of the general development in (1) assumed by Haudricourt (1950), Greenberg (1970), Bhat (1973) and Ohala (1983).

---

[4]   This sound might be a retroflex affricate, though we found no further indication for this in the literature.

3.2 Sino-Tibetan: Thulung

Thulung, also called Thulung(e) Rai (e.g. Lahaussois 2003), is a Sino-Tibetan language and belongs to the subgroup of Western Kiranti languages. It is mainly spoken in Eastern Nepal. Thulung has an extensive coronal inventory, with four laryngeal settings for dental plosives and affricates: /t tʰ d dʰ ts tsʰ dz dzʰ/ (Ebert, 1997:14). According to Ebert (1994, 2003), Thulung is the only Kiranti language with retroflex stops in addition to this dental series.[5] The voiced retroflex /ɖ/ is phonemic, since it forms minimal pairs with initial /d/ in native words. The voiceless retroflexes [ʈ ʈʰ] are marginal and do not contrast with other coronals (Ebert, 1994; Lahaussois 2003:1).

If we compare Thulung words having a voiced retroflex to cognates in neighbouring languages, we can see that other Western Kiranti languages (such as Dumi, Khaling, Jero) have a voiceless stop /t/, and the Eastern Kiranti languages (such as Camling, Bantawa, Yamphu) have a voiced stop /d/ in its place (Michailovsky 1994), see (6a).[6]

(6)	Western Kiranti (except Thulung)	Thulung	Eastern Kiranti
a)	t	ɖ	d
b)	d	d	t
c)	t	t	t

For the voiced /d/ in Thulung, we find the same phoneme in the other Western Kiranti languages, but a voiceless /t/ in the Eastern Kiranti languages, see (6b). Of importance for a historical reconstruction of Proto-Kiranti is furthermore that

---

[5] The discussion on Kiranti is restricted to initial consonants. Other Kiranti languages like Limbu and Camling have retroflex consonants in this position, but almost only in loanwords from Nepali (Ebert 2003:14; Driem 1987:27). The Western Kiranti language Jero seems to be a case like Thulung because it has the phoneme /ɖ/ in native words. However, Opgenort (2005:59) describes somewhat uninterpretable that its use instead of /d/ "seems to be generally determined by personal style or preference" (ibid.). He goes on to say that the retroflex flap [ɽ] is an allophone of /ɖ/ in intervocalic position, and is a common sound in native Jero words, indicating again that the postulation of a phoneme /ɖ/ is justified.
According to Ebert (2003:14), the Eastern Thulung language Athpare has no dental coronals, but retroflex segments instead. No further information on this language could be obtained.
Michailovsky (1994:766) lists Sunwar as Kiranti language with dentals and retroflexes. However, Sunwar is usually not considered a Kiranti language, but as belonging to the Kham-Magar-Chepang-Sunwari languages, which form together with the Kiranti languages the Mahakiranti branch of Himalayish (Gordon 2005).

[6] The Eastern-Kiranti Limbu has no voiced stops.

Thulung /t/ in (6c) merged with the cognates of Thulung /ɖ/ in the other Western Kiranti languages, and with the cognates of Thulung /d/ in the Eastern Kiranti languages (Opgenort, 2005). This intricate relationship led several scholars to reconstruct three sounds in Proto-Kiranti corresponding to the ones in (6a) - (c), namely *t for the uniformly voiceless stops in (6c), *d for the sounds in (6b), and a preglottalized *ʔt for the sounds in (6c) (see Starostin 1994, and Opgenort, 2005; Michailovsky 1994, assumes a glottalized segment at a later stage). Michailovsky (1994:770) points out that the reconstruction of a preglottalized segment is somewhat speculative since there is no direct evidence for it. Opgenort (2005:14) agrees, but proposes that the preglottalized consonant might go back to the Tibeto-Burman prefix *ʔə. None of these authors accounts for the change in voicing that has to have taken place, if one assumes the development *ʔt > ɖ. In any case, there is no indication that the reconstructed segment was realized as an implosive, nor did it give rise to an implosive in any Kiranti language.[7] We can therefore take Thulung as evidence for a further language in which a voiced retroflex stop developed directly from a front coronal stop without an intermediate stage of implosion.

3.3 East Cushitic: Afar, Somali, Rendille

East Cushitic languages belong to the Afro-Asiatic family and are spoken in Somalia, Ethiopia, Eritrea, and Kenia. A number of East-Cushitic languages are reported to have a voiced retroflex stop /ɖ/, namely Afar (Bliese 1981), Somali and Rendille (Sasse, 1979:25; Lloret 1995:69). The related languages Boni, Arbore and Elmolo have instead an alveolar implosive usually transcribed as /d'/, see Sasse (1979:25). Sasse (ibid.) also mentions Dasenech in this context. Tosco (2001), however, describes the Dasenech sound as a retroflex implosive, realized as a plain retroflex stop [ɖ] or flap [ɽ] intervocalically (p. 21). In Oromo, a further East-Cushitic language, the cognate sound is also realized as retroflex implosive (see Gragg 1976, on the Western dialect Wellega, and Stroomer 1987, on the Southern dialects Boraana, Orma and Waata). A summary of the correspondences between these languages is given in (7).

(7) a)   ɖ   Afar, Somali, Rendille
     b)   ɗ   Boni, Arbore, Elmolo
     c)   ɗ̢   Dasenech, Oromo (Western and Southern dialects)

---

[7]   Note that Opgenort (2005) proposes the existence of a preglottalized nasal *ʔn in Proto-Kiranti to account for the implosive /ɗ/ in Jero, which corresponds to plain nasals in all other Kiranti languages.

The sounds in (7) all stem from the same Proto-East Cushitic segment, which Sasse (1979:25) reconstructs as a voiced coronal stop *d' and describes as "glottalized or otherwise affected". Since this glottalized segment could be argued to have been an implosive (it resulted in implosives in neighbouring languages, and recall the discussion in section 2.3 on varied articulations and therefore inconsequent descriptions of implosives), the languages Afar, Somali and Rendille do not seem to provide strong evidence in favour of our argument that retroflexes did not necessarily develop from implosives.

It has to be mentioned, however, that Heine (1978) proposes a sub-classification of the Eastern Cushitic languages Somali, Rendille, and Boni as what he terms "Sam" languages (see also Tosco 2001), and reconstructs a Proto-Sam retroflex *ɖ which he assume to persisted into present-day Somali and Rendille but changed in Boni to an implosive /ɗ/. This reconstruction would, if correct, provide another example for the reverse development of a retroflex into an implosive, like the case of Dhao in (5).

The retroflex implosive in Dasenech (7c), which at present has a plain retroflex allophone [ɖ~ʈ] in intervocalic position (Tosco, 2001) that was not reported in earlier sources, provides an example for Haudricourt's (1950) assumption that plain retroflex voiced stops develop from retroflex implosives (3a).

To sum up, we illustrated with the examples of three typologically unrelated languages that diachronic developments of retroflex voiced stops do not necessarily proceed from alveolar or retroflex implosives. Furthermore, we saw two examples for a possible reverse development from a retroflex into an implosive, namely the change from Proto-Malayo-Polynese *ɖ to Dhao and Sabu /ɗ/, and from Proto-Sam *ɖ to Boni /ɗ/. For all the example developments discussed in this section, we have to keep in mind that we are dealing here with diachronic reconstructions of languages, which are usually based on the comparison of daughter languages and lack any kind of direct evidence. This is especially problematic for language families that exhibit a huge variety due to large areal spread and/or continuous contact with other language families (such as Austronesian).

While the languages presented up to now developed retroflex phonemes across several generations, the data on German in the following section differ in two ways: They are synchronic, and they illustrate allophonic retroflexion for one speaker (the other speakers in this study show allophonic backing). But again they provide evidence for the emergence of retroflexion from a voiced coronal stop without an intermediate stage of implosion.

4   A German case study

We chose German to provide us with synchronic data on the difference between front coronal voiced and voiceless stops and the affinity of /d/ to retroflexes for two reasons. First, it is a language without retroflex stops and where therefore the alveolars /t, d/ as the only coronal stop phonemes can considerably vary in their place of articulation (cf. the findings for French and English coronals by Dart 1998). Second, articulatory data on German in the form of Electromagnetic Articulography (EMA) and Electropalatography (EPG) was available from the study by Fuchs (2005), who looked at the realization of voicing in German obstruents. Data presented here are restricted to an intervocalic, unstressed position, because in this position a true voicing contrast is most likely for German. In initial position we find a contrast between plain and aspirated voiceless segments (Jessen 1998) and in final position a subtle contrast or none at all (due to final devoicing). The intervocalic position is also the one in which there seem to be no restrictions on the occurrence of plain stops, retroflexes and implosives.

We tested whether the voiced stop in German is realized in a way that favours retroflexion, that is, with a more retracted place of articulation, less palatal contacts, a lower tongue and lower jaw position, and with more contextual variation than /t/.

4.1 Methods

In order to test the above-mentioned differences we investigated tongue and jaw movements together with tongue-palate contact patterns by means of simultaneous EPG (Reading EPG3) and EMA recordings (AG 100, Carstens Medizinelektronik). Tongue tip (tt) movement was associated with the movement of a sensor placed midsagittally approximately 1 cm behind the tip. Tongue back (tb) movement was associated with a sensor that was placed at the posterior end of the tongue where it touches the soft palate. Since this sensor came loose during the recording session for 2 of the 4 subjects, we do not discuss it here. Two sensors, one for tongue mid (tm) and one for tongue dorsum (td), were placed in between and in equal distance to the tt and tb sensors. Jaw movement was associated with a sensor at the lower incisors. Two sensors served as reference points to compensate for helmet movements, one at the nasion and one at the upper incisors. Speech signals were recorded on Digital Audio Tape (DAT). Sampling frequencies were 16 kHz for the acoustic data, 100 Hz for EPG and 200 Hz for EMA data respectively.

Four German subjects were recorded, three male (Speakers 1-3) and one female (Speaker 4). The speech material consisted of nonsense words

"geC$_1$VC$_2$e" where C$_1$ and C$_2$ were either /t/ or /d/. The consonant C$_2$ occurred in an unstressed word medial position and the vowel preceding C$_2$ was always one of the stressed tense vowels /a, i, u/. We included different vowel context since we expected a retroflex-like articulation in /u/ context but not in the context of /i/ (recall the discussion in section 2.2). The target word was embedded in the carrier phrase *Ich habe geCVCe nicht Y erwähnt,* 'I said geCVCe not Y', with Y being another target word which is not the focus of this study here. Each sentence was repeated 10 times in a randomised order. We should point out that the measured tongue sensor signals are composed of both, the tongue and the jaw, since decomposition is not a straight forward process. For further details of the study, see Fuchs (2005).
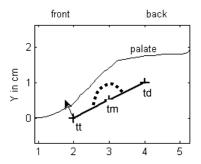
On the basis of the EMA data we labelled the consonantal target, defined as the highest vertical position of the tongue tip sensor in correspondence with tongue palate contacts in the alveolar part of the palate. For this point, the following three measures were carried out:
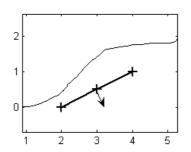
(8)   a)   the horizontal (x) position of the tongue tip,
       b)   the vertical (y) position of the tongue dorsum and of the jaw, and
       c)   the frequency of tongue palatal contacts over all repetitions.

Although jaw lowering (in 8b) has not been mentioned as a potential characteristic of retroflexes before, we assume that it goes hand in hand with the tongue dorsum lowering to allow more flexibility for the apical articulation (it may also be a requirement for tongue tip curling). By contrast, a high jaw position makes a retroflex tongue configuration very unlikely.

In addition to the consonantal target we also labelled the vowel target, defined as the lowest vertical position (or most backward position in /u/-context), corresponding approximately to the turning point for all tongue sensors.

Both vowel and consonant target are used for the measure of the tongue tip angle. This is a measure introduced by Tiede, Gracco, Shiller, Espy-Wilson & Boyce (2005) in their study on variations of American /r/ to distinguish retroflex from bunched varieties. The tongue tip angle is calculated using three successive sensor coils on the tongue starting at the tip. It is the angle between the line connecting the first (tt) and second sensor (tm) and the line connecting the second (tm) and third sensor (td), depicted in figure 1 in the left graph with the dotted line. If this angle is greater than 180 degrees the tongue has a bunched shape, if the angle is 180 degrees or lower the tongue has a retroflex shape.
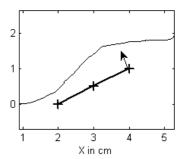
Figure 1: Schematic representation of Tiede et al.'s (2005) tongue tip angle for retroflex tongue configurations.

Figure 1 also shows that a retroflex tongue configuration according to this measure is not only possible with an upward movement of the tongue tip (graph on the left), but also with a lowering of the tongue mid (graph in the middle) and an upward movement of the tongue dorsum (graph on the right).

For the statistical analyses of the four measures we used SPSS (version 15.0).

4.2   Results

We discuss the results in the following order: horizontal position of the tongue tip in §4.2.1, frequency of tongue palatal contacts in §4.2.2, vertical position of the tongue dorsum and the jaw in §4.2.3, and the tongue tip angle in §4.2.4.

4.2.1   Retracted tongue tip position for /d/

Figure 2 displays the results for the horizontal position of the tongue tip at the consonantal target in the context of /a/ and /u/ based on EMA. It clearly shows that all speakers realize a significantly more retracted tongue tip position for /d/ in comparison to /t/ (for descriptive results and significance values see Appendix I). The differences are particularly pronounced for speaker 1 (up to 4 mm) and speaker 3 (up to 7 mm) whereas for speaker 2 and speaker 4 they are rather small (approximately 1 mm). The context of the front vowel /i/ was not included here, because in this context only speaker 3 had significant differences between /d/ and /t/.
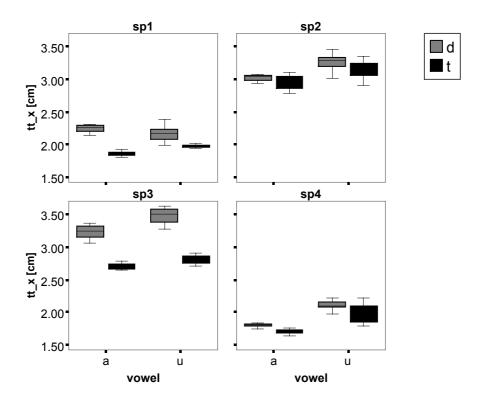
Figure 2: Boxplots with standard deviations for tongue tip horizontal target position for /d/ (grey) and /t/ (black) for the 4 speakers and /a, u/- contexts; lower values = more front articulation.

On the basis of the EMA data in figure 2, we can only specify the position of the tongue tip by its flesh point marker in the mid-sagittal plane, and conclude from it the place of articulation. Conclusions on the actual size of contact can only be gained from EPG data.

4.2.2 Area of contact for /d/

In figure 3, we see EPG frequency plots, which show the pattern of the tongue touching the palate at the consonantal target over all repetitions. The four columns in figure 3 below correspond to the four subjects, the four rows to /at/, /ad/, /ut/, and /ud/. The highest y-value corresponds to the most anterior row at the EPG palate and the lowest y-value to the most posterior row.
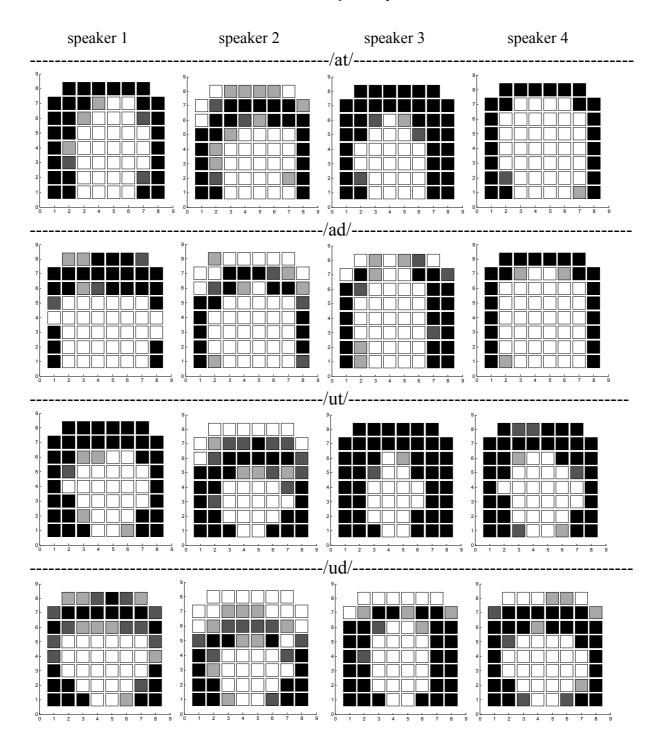
Figure 3: EPG frequency plots for all speakers (4 columns) in /a/-context (first 2 rows) and /u/-context (last 2 rows); /t/ = 1st and 3rd row, /d/ = 2nd and 4th row; black markers correspond to 76-100% tongue palatal contact with respect to all the subject's repetitions, dark grey markers to 51-75%, light grey to 26-50%, and white markers to 0-25%.

The EPG data in figure 3 shows that /d/ has generally a more retracted place of articulation than /t/, both in /a/- and /u/-context. The only exception is speaker 4, who shows equal amount of contact in /a/-context. This might be due to the fact that speaker 4 has a very fronted articulation in this context,

possibly with dental contact, which cannot be recorded with EPG. The percentage of contact over the whole palate for all speakers is significantly greater for /t/ than for /d/, and we can see that /d/ is often produced with less lateral contacts than /t/. Both findings can be interpreted as a more forceful articulation of /t/ (with a higher target position) and a difference in active articulator between the two (where the voiced stop is being articulated with the tongue tip and the voiceless one with the tongue blade). Furthermore, the contact for /d/ shows more variation than that for /t/ (see Appendix II for all descriptive statistics and significances).

Figure 4 displays the EPG frequency plots for the /i/-context. Although there are still subtle differences between /t/ (top row) and /d/ (bottom row), the overall amount of tongue palatal contact is very large for both, especially at the lateral margins of the palate.

Figure 4: Same as previous figure, but for /i/-context. Note that for speaker 2 only 3 repetitions of /t/ could be included since this speaker realized the relevant target word in most cases with a /d/.

4.2.3   Lowering of tongue dorsum and jaw for /d/

A lowered tongue dorsum is a typical property of retroflex segments, and often goes together with a lowering of the jaw, as discussed in section 2.2. To what extent can these properties also be found in /d/ compared to /t/? In a univariate ANOVA we took tongue dorsum y position and jaw y position as

dependent variables and phoneme (/d/ versus /t/) and vowel context (/a, i, u/) as factors (data were split by speaker). The descriptive statistics and significances are given in Appendix III.

For the vertical tongue dorsum position we found main effects of vowel for all speakers, and a main effect of phoneme for speakers 1 and 2. All speakers but speaker 3 show an interaction between the two factors: In /a/-context, /d/ is realized with a lower tongue dorsum position. In /i/-context, both consonants have a similar tongue position (except for speaker 1 who shows a slightly higher dorsum for /d/). In /u/-context, results vary speaker-dependently: speakers 1 and 2 show similar results for /d/ and /t/, speaker 3 shows a higher /t/ than /d/ and speaker 4 the reverse. These findings show that vertical tongue movement is influenced to a large degree by the vowel context: For the articulation of the high back vowel /u/ a raising of the tongue dorsum is necessary, and for the high front vowel /i/ the dorsum is raised along with the necessary raising of the tongue blade.
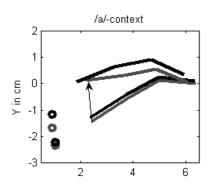
For the vertical jaw position, we found main effects of vowel for all four speakers, a main effect of phoneme for all but speaker 2, and an interaction of the two only for speaker 1.
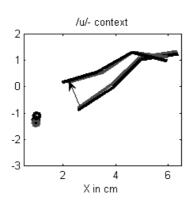/d/ is articulated with a lower jaw than /t/ for three of the four speakers. Considering the actual values, it becomes evident that although significant, jaw differences are often very subtle. The most pronounced differences are consistently found for speaker 1.

From these findings we can conclude that there are obviously speaker-dependent strategies in the use of the tongue and the jaw. Speaker 3 was the only one who did not show a significant tongue lowering for /d/ (in /u/-context), and speaker 2 the only one who did not show a significant jaw lowering for /d/. Thus, whereas some speakers show tongue lowering for /d/, others show jaw lowering, and some show both.

4.2.4   Retroflex tongue configuration for /d/

The tongue tip angle (Tiede et al. 2005) is a measure to separate a retroflex tongue configuration from a bunched one, independent of the actual phonological retroflexivity of the sound. Although our dataset does not include retroflex phonemes, speaker 1 shows consistently a retroflex tongue configuration for /d/ in back vowel context. All other speakers do not show such configurations. The following figure is therefore restricted to an illustration of our findings for speaker 1.
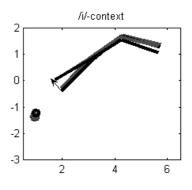
Figure 5: Averaged tongue contours for speaker 1 based on the connection of the 4 flesh point markers at the tongue; x-axis = horizontal movements in cm (left is anterior), y-axis = vertical movements in cm; lower lines = vowel targets, upper lines = consonantal target; little arrows mark the movement from vowel to consonantal targets; grey = /d/, black /t/; from left to right: /a, u, i/-context; empty circles correspond to jaw position.

In figure 5, an arrow indicates the movement from the vowel (lower line) to the consonant (upper line). In /a/-context (left plot) and /u/-context (middle plot), a retroflex tongue configuration can be seen for /d/. Here the tongue tip angle for the consonantal target is below 180 degrees. It is interesting to note that in these cases the tongue tip angle for the preceding vowel target is also below 180 degrees, thus retroflex, but to a far lesser extent (for the /a/-context this is not even visible). In /u/-context, the retroflex-like tongue configuration is also observable for the voiceless stop, and can thus be attributed to the vowel. For /a/-context, however, it is unique to the voiced stop.

In /i/-context (right plot) the tongue tip angle for /d/ is above 180 degrees, and we can see no retroflex tongue configuration in the average tongue contours. Here, the overall movement is very small and very close to the palate.

4.3 Discussion

Our data show that there is a systematic difference between /d/ and /t/ for all four speakers of German. This difference is mainly restricted to the context of /a/ and /u/, where /d/ has a smaller constriction and less lateral contacts (both can be interpreted as an apical articulation), shows more variation in the location of this constriction, and also has a more retracted place of articulation than /t/. Furthermore, three of the four speakers showed a small but significant difference in jaw position, with /d/ having a lower jaw, independent of context. These findings are in accordance with the literature on the difference between /d/ and /t/, recall the discussion in section 2.1, and indicate that the voiced alveolar stop in German is realized in a way that favours retroflexion.

In /i/-context, we could not observe significant differences between /d/ and /t/ in any of these parameters apart from jaw position. This influence of

vowel context coincides with previous observations that retroflex tongue configurations avoid /i/-context and that /u/ context leads to retroflexion, see the discussion in section 2.2.

The expected lowering of the tongue dorsum could only be found in the context of /a/ (for all but speaker 3). This is because the tongue dorsum plays an integral part in the articulation of non-low and back vowels, and if these vowels are adjacent to coronal consonants, they seem to influence the position of the dorsum to a large degree.

For one speaker (speaker 1) we actually found a tongue configuration for /d/ in /a/ and /u/-context that was distinctively retroflex, with a raised tongue tip, a lowered tongue mid, and a raised tongue back. We also found such a configuration for /t/, but only in /u/-context. We conclude from this that retroflexion in /u/-context is independent of the retroflexion of voiced coronal stops, since it influences both voiced and voiceless stop.

In sum, our data illustrate not only that German /d/ is articulated in a way that favours retroflexion, but that retroflexion is an acceptable articulation of /d/ in German, a language that does not have any other voiced coronal stops.

5   Conclusion

In this study we looked at the question whether voiced retroflex stops can develop from front coronal voiced stops. Our aim was to illustrate that this development is possible without an intermediate stage of implosion, contrary to what has been proposed by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983). In addition, we proposed and tested a possible phonetic motivation for this process, namely the articulatory affinity between voiced front coronal stops and voiced retroflex stops.

For these purposes, we provided data from sound changes that introduced voiced retroflex stops as single retroflex in a language from a front coronal stop. Our diachronic examples came from Central Malayo-Polynesian (Dhao), Sino-Tibetan (Thulung), and East Cushitic (Afar, Somali, Rendille). *d is the proto-segment proposed in the existing literature as corresponding to the present-day retroflexes for one language group (Central Malayo-Polynesian). For the other two groups a preglottalized stop is reported, either voiced (East Cushitic) or voiceless (Sino-Tibetan), though the latter has not been given any motivation for the change in voicing (or the drop in glottalization, for that matter).

Furthermore, we conducted an EPG and EMA experiment with four speakers of German to compare the articulation of voiced and voiceless coronal stops. We found that German /d/ shows a more retracted and more variable place of articulation, a smaller percentage of tongue palatal contact patterns, and a lower tongue and jaw position than its voiceless counterpart /t/, especially in

the context of low and back vowels. All these criteria are also used to distinguish retroflex from non-retroflex coronal articulations in languages like Norwegian, Hindi or Tiwi, and support our hypothesis that alveolar and retroflex articulations are similar to each other and can be said to form an articulatory continuum without a sharp boundary. One of our speakers actually produced retroflex [ɖ] as a realization of the voiced alveolar stop phoneme in low and back vowel context.

The articulatory data thus provides us with a phonetic explanation for the typological diachronic findings, an approach that has been forwarded in linguistic theory over the last decades (see, e.g., Ohala 1993 2005; Blevins 2004; 2007). Our study supplements the work by Haudricourt and following on the development of retroflexes from implosives. But whereas the explanations that were provided for the change from implosive to retroflex (such as cavity enlargement, see Ohala 1983) implied a strict direction of sound change (they could not account for the reverse process), the articulatory similarity we propose holds for processes in both directions.

It seems as if the process of retroflexion via implosion would benefit too from an explanation that does not imply a preferred direction, as there is evidence for reverse processes. We gave two potential examples of languages in which a retroflex might have become implosive. In Sabu and Dhao, /ɗ/ is likely to stem from a reconstructed *ɖ in Proto-Malayo-Polynesian (or Proto-Austronesian), and /ɗ/ in Boni might stem from *ɖ in Proto-Sam (as proposed by Heine 1978). An existing example for such a process is the development of Saramaccan, a creole language of Surinam, which has a voiced coronal implosive (Bakker, Smith & Veenstra 1995) which stems from the retroflex voiced stop in the lexical contributor language Fon and other closely related Gbe languages (Smith & Haabo 2007)[8].

All these processes can be acounted for by the articulatory-similarity space proposed in section 2.3, which is based on two gradual continua: One from plain stop to stop with rapidly and strongly lowered larynx (implosive), as proposed by Ladefoged (1964 et seq.), and one from plain tongue shape to retroflex tongue shape (independent of the actual place of articulation) that we proposed in section 2.2. The two continua are orthogonal to each other and create a two-dimensional space, including a gradient continuum from plain to implosive retroflex. Any change from one voiced coronal segment to another within this space is simply due to articulatory similarity (either on one of the two

---

[8]   Smith and Haabo (2007) find the circular diachronic development of Proto-Volta-Congo *ɗ to Proto-Gbe/Fon ɖ to Saramaccan ɗ "unexpected" (p.17) and propose that the implosive articulation of this sound has been continous. This contrasts with the general assumption that Proto-Gbe had a retroflex (Capo 1991).

dimensions or on both). The similarity does not imply a preferred direction of change. We have to be aware of the fact that ingressive airflow is not included in this space, though it is a defining criterion of implosives. Future research has to show whether the difference in airflow (from egressive to ingressive) forms a separate dimension and thus makes our similarity space three-dimensional, or whether it correlates with the existing dimension of plain to retroflex implosives.

6   Acknowledgements

7   References

Andersen, Torben (1987). An outline of Lulubo phonology. *Studies in African Linguistics* 18, 39-66.

Anderson, Victoria Balboa & Ian Maddieson (1994). Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics* 87, 131-162.

Arndt, Paul P. (1933). *Grammatik der Ngad'a-Sprache*. Bandoeng: A.C. Nix & Co.

Baird, Louise (2002). Kéo (Illustrations of the IPA). *Journal of the International Phonetic Association* 32, 93-97.

Bakker, Peter, Norval Smith & Tonjes Veenstra (1995). Saramaccan. In: Jacques Arends, Peter Muysken & Norval Smith (eds.) *Pidgins and creoles: an introduction*. Amsterdam: Benjamins.

Bell-Berti, Fredericka (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America* 57, 456-461.

Bender, M. Lionel (1992). Central Sudanic segmental and lexical reconstructions. *Afrikanische Arbeitspapiere* 29, 5-61.

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.

Blevins, Juliette (2004). *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Blevins, Juliette (2007). The importance of typology in explaining recurrent sound patterns. *Linguistic Typology* 11, 107-113.

Bliese, Loren F. (1981). *A generative grammar of Afar*. Dallas: Summer Institute of Linguistics.

Capo, Hounkpati B.C. (1991). *A comparative phonology of Gbe*. Berlin: Foris Publications.

Catford, John C. (1939). On the classification of stop consonants. *Le Maître Phonétique* 65, 2-5.

Clements, George N. & Sylvester Osu (2003). Explosives, implosives, and nonexplosives: the linguistic function of air pressure differences in stops. In: Carlos Gussenhoven & Natasha Warner (eds.) *Papers in Laboratory Phonology 7*. Berlin: Mouton de Gruyter, 299-350.

Clements, George N. & Annie Rialland (2005). *Africa as a phonological area*. Manuscript, Paris.

Dart, Sarah N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79, 1-155.

Dart, Sarah N. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics* 26, 71-94.

Demolin, Didier & Didier L. Goyvaerts (1986). Some aspects of Madi phonology and morphology. *Antwerp papers in Linguistics* 44, 89-103.

Demolin, Didier (1988). Some problems of phonological reconstruction in Central Sudanic. *Belgian Journal of Linguistics* 3, 53-96.

Dempwolff, Otto (1934). Vergleichende Lautlehre des Austronesische Wortschatzes, Band 1: Induktiver Aufbau einer indonesischen Ursprache. *Beihefte zur Zeitschrift für Eingeborenen-Sprachen* 15.

Dimmendaal, G.J. (1986). Language typology, comparative linguistics and injective consonants in Lendu. *Afrika und Übersee* 69, 161-192.

Dixit, R. Prakash (1990). Linguotectal contact patterns in the dental and retroflex stops of Hindi. *Journal of Phonetics* 18, 189-201.

Dixit, R. Prakash & James Flege (1991). Vowel context, rate and loudness effects of linguopalatal contact patterns in Hindi retroflex /ˇ /. *Journal of Phonetics* 19, 213-229.

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

Driem, George L. van (1987). *A grammar of Limbu*. Berlin: Mouton de Gruyter.

Dyen, Isidore (1971). The Austronesian languages and Proto-Austronesian. In: Thomas A. Sebeok (ed.) *Linguistics in Oceania*. The Hague: Mouton, 5-54.

Ebert, Karen (1994). *The structure of Kiranti languages: comparative grammar & texts*. Zürich: Universität Zürich.

Ebert, Karen (2003). Kiranti languages: An overview. In: Graham Thurgood & Randy L. LaPolla (eds.) *The Sino-Tibetan languages*. London: Routledge, 505-517.

Emeneau, M. B. (1944). *Kota texts. Part 1*. Berkeley & Los Angeles: University of California Press.

Esser, S. J. (1938). Talen. *Altas van Tropisch Nederland*. Amsterdam: Koninklijk Nederlandsch Aardrijkskundig Genootschap.

Ewan, William G. & R. Krones (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics* 2, 327-335.

Farnetani, Edda (1989). An articulatory study of "voicing" in Italian by means of dynamic palatography. Paper presented at the *Speech Research International Conference*. Budapest.

Farnetani, Edda (1990). V-C-V lingual coarticulation and its spatiotemporal domain. In: William J. Hardcastle & Alain Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer, 93-130.

Flemming, Edward (2003). The relationship between coronal place and vowel backness. *Phonology* 20, 335-373.

Fox, James J. (2004). Current Developments in Comparative Austronesian Studies. Paper presented at the *Symposium Austronesia*. Bali.

Fuchs, Susanne, Pascal Perrier & Christine Mooshammer (2001). The role of the palate in tongue kinematics: An experimental assessment in VC sequences from EPG and EMMA data. Paper presented at the *Eurospeech*. Aalborg.

Fuchs, Susanne (2005). Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS Working Papers in Linguistics* 41.

Fuchs, Susanne, Pascal Perrier, Christian Geng & Christine Mooshammer (2006). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In: Jonathan Harrington & Marija Tabain (eds.) *Towards a better understanding of speech production processes*. New York: Psychology Press, 149-164.

Fujimura, Osamu, I.F. Tatsumi & Ryohei Kagaya (1973). Computational processing of palatographic patterns. *Journal of Phonetics* 1, 47-54.

Fujimura, Osamu & J.E. Miller (1979). Mandible height and syllable-final tenseness. *Phonetica* 36, 263-272.

Gordon, Raymond G. Jr. (2005). *Ethnologue: Languages of the World*. Dallas: SIL International.

Goyvaerts, Didier L. (1986). Glottalized consonants: a new dimension? *Antwerp papers in Linguistics* 44, 105-113.

Gragg, Gene (1976). Oromo of Wellegga. In: M. Lionel Bender (ed.) *The Non-Semitic Languages of Ethiopia*. East Lansing, Michigan: Michigan State University Press, 166-221.

Greenberg, Joseph H. (1970). Some Generalizations concerning Glottalic Consonants, Especially Implosives *International Journal of American Linguistics* 36, 123-145.

Grimes, Charles E. (2006). Hawu and Dhao in eastern Indonesia: revisiting their relationship. Paper presented at the *10th International Conference on Austronesian Linguistics*. Puerto Princessa, Philippines.

Guthrie, Malcolm (1967-1971). *Comparative Bantu*. Farnborough, Hants: Gregg International.

Hamann, Silke (2002). Retroflexion and retraction revised. *ZAS Working Papers in Linguistics* 28, 13-25.

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

Hamann, Silke (2004). Retroflex fricatives in Slavic Languages. *Journal of the International Phonetic Association* 34, 53-67.

Haudricourt, André-Georges (1950). Les consonnes préglottalisées en Indochine. *Bulletin de la Société de Linguistique* 46, 172-182.

Heine, Bernd (1978). The Sam languages: a history of Rendille, Boni and Somali. *Afroasiatic Linguistics* 6.

Jonker, Johan Christoph Gerard (1896). Bimaneesche Spraakkunst. *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen (VBG)* 48.

Jonker, Johan Christoph Gerard (1903). Iets over de taal van Dao. *Album Kern (Opstellen geschreven ter eere van Dr. H. Kern)*. Leiden: Brill, 85-89.

Keating, Patricia A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286-319.

Khatiwada, Rajesh (2007). Nepalese retroflex stops: a static palatography study of inter- and intra-speaker variability. Paper presented at the *Interspeech*. Antwerp.

Klamer, Marian (1998). *A Grammar of Kambera*. Berlin: Mouton de Gruyter.

Krull, Diana, Björn Lindblom, B.-E. Shia & D. Fruchter (1995). Cross-linguistic aspects of coarticulation: an acoustic and electropalatographic study of dental and retroflex consonants. *Proceedings of the XIIIth International Congress of Phonetic Sciences (ICPhS95, Stockholm)* 3, 436-439.

Ladefoged, Peter (1964). *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey*. Cambridge: Cambridge University Press.

Ladefoged, Peter & Peri Bhaskararao (1983). Non-quantal aspects of consonant production: a study of retroflex consonants. *Journal of Phonetics* 11, 291-302.

Ladefoged, Peter & Zongji Wu (1984). Places of articulation: an investigation of Pekingese fricatives and affricates. *Journal of Phonetics* 12, 267-278.

Ladefoged, Peter & Ian Maddieson (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Lahaussois, Aimée (2003). Thulung Rai. *Himalayan Linguistics* 1, 1-25.

Lex, Gloria (2006). *Le Dialecte Peul du Fouladou (Casamance - Sénégal): Etude Phonétique et Phonologique*. München: Lincom.

Lindblom, Björn (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773-1781.

Lisker, Leigh & Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20, 384.

Lloret, Maria-Rosa (1995). Implosive consonants: their representation and sond change effects. *Belgian Journal of Linguistics* 9, 59-72.

Löfqvist, Anders & V.L. Gracco (2002). Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America* 111, 2811-2827.

Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Michailovsky, Boyd (1994). Manner vs. place of articulation in the Kiranti initial stops. In: H. Kitamura & et al (eds.) *Current Issues in Sino-Tibetan Linguistics*. Osaka, 766-772.

Moen, Inger & Hanne Gram Simonsen (1997). Effects of voicing on /t, d/ tongue/palate contact in English and Norwegian. Paper presented at the *5th European Conference on Speech Communication and Technology*. Rhodes.

Mooshammer, Christine, Philip Hoole & Anja Geumann (2006). Inter-articulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America* 120, 1028-1039.

Mooshammer, Christine, Philip Hoole & Anja Geumann (2007). Jaw and order. *Language and Speech* 50, 145-176.

Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In: P.F. MacNeilage (ed.) *The production of speech*. New York: Springer, 189-216.

Ohala, John J. (1993). The phonetics of sound change. In: Charles Jones (ed.) *Historical linguistics: problems and perspectives*. London: Longman, 237-278.

Ohala, John J. (2005). Phonetic explanations for sound patterns: Implications for grammars of competence. In: William J. Hardcastle & M. Beck (eds.) *A figure of speech: A festschrift for John Laver*. London: Erlbaum, 23-38.

Perkell, Joseph S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge: MIT Press.

Repp, Bruno (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech* 22, 173-189.

Ross, Malcolm (1992). The sounds of Proto-Austronesian: An outsider's view of the Formosan evidence. *Oceanic Linguistics* 31, 23-64.

Ross, Malcolm (1995). Some current issues in Austronesian linguistics. In: Darrell T. Tryon (ed.) *Comparative Austronesian dictionary: An introduction to Austronesian studies*. Berlin: Mouton de Gruyter, 45-120.

Sasse, Hans-Jürgen (1979). The consonant phonemes of Proto-East-Cushitic (PEC): A first approximation. *Afroasiatic Linguistics* 7, 1-67.

Schuh, Russell G. (2003). Chadic overview. In: M. Lionel Bender, Gábor Taka\cs & David L. Appleyard (eds.) *Selected comparative-historical Afrasian linguistic studies. In memory of Igor M. Diakonoff*. München: Lincom, 55-60.

Simonsen, Hanne G., Inger Moen & Steve Cowen (2000). Retroflex consonants in Norwegian: are they really? Evidence from EMA and EPG. Paper presented at the *5th Seminar on Speech Production*. Kloster Seeon, Germany.

Smith, Norval & Vinije Haabo (2007). The Saramaccan implosives: tools for linguistic archaeology? . *Journal of Pidgin and Creole Languages* 22.

Starostin, Sergei Anatolyevich (1994). Proto-Kiranti. Paper presented at the *Sino-Tibetan Conference*. Paris.

Steriade, Donca (2001). Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume & Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. San Diego (CA): Academic Press, 219-250.

Stevens, Kenneth, Samuel Keyser & Haruko Kawasaki (1986). Toward a phonetic and phonological theory of redundant features. In: Joseph S. Perkell & D. H. Klatt (eds.) *Invariance and Variability in Speech Processes*. Hillsdale: Erlbaum, 426-449.

Stewart, John M. (1995). Implosive, homorganic nasals and nasalized vowels in Volta-Congo. In: E. 'Nolue Emenanjo & Ozo-Mekuri Ndimele (eds.) *Issues in African languages and linguistics: Essays in honor of Kay Williamson*. Aba: National Institute for Nigerian Languages, 162-169.

Stroomer, Harry (1987). *A Comparative Study of Three Southern Oromo Dialects in Kenya*. Hamburg: Helmut Buske Verlag.

Švarny´, Oldřich & Kamil Zvelebil (1955). Some remarks on the articulation of the 'cerebral' consonants in Indian languages, especially in Tamil. *Archiv Orientalni* 23, 374-407.

Tiede, Mark K., Vincent L. Gracco, Douglas M. Shiller, Carol Espy-Wilson & Suzanne E. Boyce (2005). Perturbed palatal shape and North American English /r/ production. *Journal of the Acoustical Society of America* 117, 2568-2569.

Tosco, Mauro (2001). *The Dhaasanac Language: Grammar, Texts, Vocabulary of a Cushitic Language of Ethiopia*. Köln: Rüdiger Köppe.

Trask, Robert L. (1996). *A Dictionary of Phonetics and Phonology*. London: Routledge.

Tucker, A. N. & M. A. Bryan (1966). *Linguistic analyses, the non-Bantu languages of northeastern Africa*. Oxford: Oxford University Press.

Wakumoto, Masahiko, Shinobu Masaki, Kiyoshi Honda & Toshikazu Ohue (1998). A pressure sensitive palatography: application of new pressure sensitive sheet for measuring tongue-palatal contact pressure. *Proceedings of the 5rd International Conference on Spoken Language Processing*. Sydney, 3151-3154.

Watson, Richard L. (1991). Moru-Ma'di orthographies. In: M. Lionel Bender (ed.) *Proceedings of the Fourth Nilo-Saharan linguistics colloquium (Nilo-Saharan: Linguistic analysis and documentation Vol. 7)*. Hamburg: Helmut Buske Verlag, 273-282.

Westbury, John R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73, 1322-1336.

Wolff, John (1974). Proto-Austronesian *r and *d. *Oceanic Linguistics* 13, 77-121.

Wolff, John (1991). The Proto-Austronesian phoneme *t and the grouping of the Austronesian languages. *Pacific Linguistics (Series C)* 117, 535-549.

Appendix I: Descriptive statistics for horizontal position of tongue tip at the consonantal target

subject	vowel	/d/			/t/		
		N	Mean	Std. dev.	N	Mean	Std. dev.
speaker 1	a	10	2.22	0.12	10	1.86	0.04
	i	10	1.76	0.05	10	1.72	0.02
	u	10	2.17	0.13	10	1.98	0.03
speaker 2	a	10	3.03	0.08	10	2.95	0.11
	i	10	2.67	0.06	4	2.72	0.02
	u	10	3.26	0.14	9	3.13	0.15
speaker 3	a	10	3.23	0.10	10	2.72	0.05
	i	10	2.93	0.11	10	2.73	0.10
	u	10	3.48	0.13	10	2.80	0.06
speaker 4	a	10	1.80	0.03	9	1.70	0.06
	i	10	1.65	0.08	10	1.69	0.10
	u	10	2.12	0.10	10	2.00	0.14

subject	factor(s)	df	F	P
speaker 1	vowel	2	113.04	0.000
	phoneme	1	94.01	0.000
	vowel * phoneme	2	22.11	0.000
speaker 2	vowel	2	75.67	0.000
	phoneme	1	3.04	0.088
	vowel * phoneme	2	2.63	0.083
speaker 3	vowel	2	55.25	0.000
	phoneme	1	370.44	0.000
	vowel * phoneme	2	33.51	0.000
speaker 4	vowel	2	99.86	0.000
	phoneme	1	6.62	0.013
	vowel * phoneme	2	4.24	0.020

Appendix II: Descriptive statistics and univariate ANOVA for overall percent of tongue palatal contact patterns at the consonantal target

subject	vowel	/d/ Mean	/d/ Std. dev.	/t/ Mean	/t/ Std. dev.
speaker 1	a	40.32	4.30	50.81	4.45
	i	52.90	6.49	62.26	4.31
	u	44.84	13.14	59.03	3.82
speaker 2	a	28.23	11.91	45.48	6.62
	i	52.10	11.78	58.87	3.36
	u	38.55	7.03	49.28	7.04
speaker 3	a	40.48	6.24	62.10	3.25
	i	64.84	6.17	75.97	4.13
	u	50.65	4.70	68.23	14.43
speaker 4	a	37.90	1.90	38.71	2.42
	i	52.15	5.23	54.84	6.76
	u	50.97	6.04	56.77	6.49

subject	factor(s)	df	F	P
speaker 1	vowel	2	15.15	0.000
	phoneme	1	40.45	0.000
	vowel * phoneme	2	0.67	n.s.
speaker 2	vowel	2	15.64	0.000
	phoneme	1	19.68	0.000
	vowel * phoneme	2	1.37	n.s.
speaker 3	vowel	2	32.97	0.000
	phoneme	1	75.63	0.000
	vowel * phoneme	2	2.51	n.s.
speaker 4	vowel	2	55.38	0.000
	phoneme	1	5.10	0.028
	vowel * phoneme	2	1.14	n.s.

Appendix III: Descriptive statistics and univariate ANOVA for vertical position (in cm) of tongue dorsum (td) and jaw at the consonantal target

| tongue dorsum | | /d/ | | | /t/ | | |
subject	vowel	N	Mean	Std. dev.	N	Mean	Std. dev.
speaker 1	a	10	0.56	0.10	10	0.92	0.11
	i	10	1.67	0.04	10	1.56	0.08
	u	10	1.30	0.07	10	1.32	0.09
speaker 2	a	10	0.11	0.09	10	0.30	0.15
	i	10	0.66	0.08	4	0.68	0.15
	u	10	0.43	0.08	9	0.47	0.08
speaker 3	a	10	0.93	0.23	10	0.92	0.06
	i	10	1.47	0.11	10	1.50	0.07
	u	10	1.35	0.06	10	1.43	0.06
speaker 4	a	10	0.12	0.07	9	0.23	0.06
	i	10	1.02	0.11	10	0.98	0.04
	u	10	0.81	0.06	10	0.72	0.06

| jaw | | /d/ | | /t/ | |
subject	vowel	Mean	Std. dev.	Mean	Std. dev.
speaker 1	a	-1.67	0.09	-1.15	0.06
	i	-1.35	0.06	-1.22	0.06
	u	-1.36	0.09	-1.11	0.07
speaker 2	a	-1.36	0.03	-1.32	0.06
	i	-1.27	0.03	-1.26	0.02
	u	-1.23	0.03	-1.22	0.02
speaker 3	a	-0.94	0.03	-0.93	0.01
	i	-0.97	0.04	-0.93	0.02
	u	-0.93	0.03	-0.91	0.02
speaker 4	a	-1.10	0.05	-1.01	0.05
	i	-1.09	0.03	-1.05	0.03
	u	-0.98	0.02	-0.94	0.02

subject	factor(s)	df	F (td)	P (td)	F (jaw)	P (jaw)
speaker 1	vowel	2	555.05	0.000	31.04	0.000
	phoneme	1	18.18	0.000	256.12	0.000
	vowel * phoneme	2	41.00	0.000	37.95	0.000
speaker 2	vowel	2	78.71	0.000	54.25	0.000
	phoneme	1	7.64	0.008	3.04	n.s.
	vowel * phoneme	2	3.60	0.035	0.80	n.s.
speaker 3	vowel	2	134.16	0.000	5.70	0.006
	phoneme	1	1.44	n.s.	9.35	0.003
	vowel * phoneme	2	0.73	n.s.	1.19	n.s.
speaker 4	vowel	2	755.37	0.000	53.16	0.000
	phoneme	1	0.15	n.s.	41.43	0.000
	vowel * phoneme	2	11.23	0.000	3.01	n.s.