Retroflexion of Voiced Stops: Data from Dhao, Thulung, Afar and German

Silke Hamann[1], Susanne Fuchs[2]

[1] *Universität Düsseldorf*
[2] *Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin*

Key words

articulation (EMMA, EPG)

diachronic development

implosives

retroflex

Abstract

The present article illustrates that the specific articulatory requirements for voiced alveolar or dental stops can cause tongue tip retraction and tongue mid lowering and thus retroflexion of voiced 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) could have emerged from dental or alveolar voiced stops because the voiced front coronal plosive /d/ is generally articulated in a way that favors retroflexion, that is, with a smaller and more retracted place of articulation and a lower tongue and jaw position than /t/. The present proposal thereby supplements the observation made by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983) that retroflex voiced stops can emerge from voiced coronal implosives for articulatory and aerodynamic reasons.

Acknowledgments: We would like to thank Anders Löfqvist, Alexis Michaud, and Mark Tiede for helpful comments on the manuscript, and Marzena Żygis for remarks on an earlier version. We would also like to thank Jonathan Harrington, Phil Hoole, Claudia Kuzla, Marianne Pouplier, and further members of the Institut für Phonetik und Sprachverarbeitung München and the audience at the *Dritte Jahrestreffen der Phonetik und Phonologie* in Stuttgart in October 2007 for their remarks on our presentations of the data. For helping us with the calculation of the tongue tip angle, we thank Paul Boersma. We gratefully acknowledge a grant by the German Research Council (DFG; GWZ 4/8-P1) to Susanne Fuchs and a grant by the Dutch Science Foundation (NWO; GW 016.064.057) to Silke Hamann.

Address for correspondence. Silke Hamann, Universität Düsseldorf, Institut für Sprache und Information, building 23.21 room 04.48, Universitätsstr. 1, 40225 Düsseldorf, Germany; <hamann@phil-fak.uni-duesseldorf.de>

© The Authors, 2010. Reprints and permissions: www.sagepub.co.uk/journalsPermissions.nav

Language and Speech 0023-8309; Vol 53(2): 181–216; 357159; DOI:10.1177/0023830909357159
http://las.sagepub.com

1 Introduction

Over the last decades research on the phonetic explanations of diachronic processes has increased, drawing on typological data to support the claims made (see, e.g., Ohala, 1993, 2005; Blevins, 2004, 2007). The present study applies this approach to account for the diachronic emergence of retroflexes from voiced front coronal stops.

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 and 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 with retroflexes as 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 the 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 goes back to 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: "Retroflex stops are distinguished from nonretroflex primarily by having an enlarged oral cavity immediately behind the point of constriction" (p.200). Owing to this enlarged oral cavity, voicing can be maintained longer than in non-retroflex stops, hence retroflexion is a cavity enlarging strategy to maintain voicing.

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).

The development described by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983) is depicted in (1), with retroflexion emerging from voiced, implosive stops in languages with no other retroflexes (note that the intermediate step of a retroflex implosive is not explicitly mentioned by Bhat and Ohala).

(1) ɗ (> ᶑ) > ɖ

In the present article we describe an additional development of retroflex voiced stops from pulmonic egressive voiced front coronal stops, and argue that articulatory requirements are responsible for this process. This development is represented in (2).

(2) d > ɖ

The focus of the present study is the emergence of retroflex sounds from voiced stops proposed in (2), though we come back to the interaction of plain voiced stops with implosives and retroflexes in sections 2.3 and 6 below. Evidence for the process in (2) comes from diachronic developments of retroflexes in a number of languages. 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 favoring retroflexion of /d/. Both the diachronic and the articulatory evidence support the phonetic naturalness of the process in (2).

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 discuss three typologically unrelated languages that have [ɖ] as the only retroflex. Section 4 provides synchronic articulatory and acoustic data from German. In section 5 we elaborate how articulatory variation can cause diachronic change, and in section 6 we conclude.

2 Voiced coronal stops

To provide evidence for the claim that voiced but not voiceless front coronal stops are prone to develop into retroflexes, we first look at the articulatory differences between voiced and voiceless front coronal stops (section 2.1), including possible explanations for this difference. We then compare the characteristics of voiced front coronals with those of retroflex stops (section 2.2) and coronal implosives (section 2.3). The last section (section 2.4) discusses explanations and examples for developments of retroflexes via implosion, supporting the observation made by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983).

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 realization 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 are banned from initial position in a large 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 posterior place of articulation than its voiceless counterpart, see for instance the electropalatographic studies by Dixit (1990) on Hindi, Moen and 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 (e.g., Stevens, Keyser, & Kawasaki, 1986, p.432; unless /d/ is flapped, see Lisker & Price, 1979).

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 is Ohala's explanation for the diachronic change 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 often 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; Löfqvist & Gracco, 2002; Fuchs, Perrier, Geng, & Mooshammer, 2006), since no exact location is necessary compared to the precise positioning required for sibilants (to create a channel that directs the air onto the teeth) or trills (to allow for a rapid movement of the tongue tip caused by the airflow). 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) because 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 and 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 intraoral pressure required for such a salient burst can be achieved by a higher tongue and jaw position during the release phase. 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. (2007) found a high and stable jaw position for /t/ in German (significant difference for three of the five speakers in both normal and loud speech, for one speaker only in normal speech). For /d/ the jaw was positioned lower, giving the tongue more freedom to move and to accommodate to the context.

Most of these four explanations cannot be evaluated separately. Thus the less salient burst and the less forceful seal both result 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, 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 those of /t/. We can therefore exclude cavity enlargement as an 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, pp.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 and Maddieson's (1994) study on Tiwi coronal stops, where the closure duration of retroflex stops 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, pp.94–107). The context of /u/, on the other hand, has been reported to cause retroflexion of front coronals (Bhat, 1973; Hamann, 2003, pp.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 stops similar to that between voiced and voiceless front coronal stops discussed above is expected, though we found 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. Owing to the strong similarity between a voiced front coronal stop and a voiced retroflex, the two can be considered endpoints on a continuum from plain front stops to retroflexes with a large amount of retroflexion, as has been proposed by Ladefoged and Bhaskararao (1983, p.299). This continuum supports our claim that a /d/ can develop into a /ɖ/ without an intermediate stage of implosion (where the direction of airflow has to change, too), simply by a slight articulatory shift along this continuum.

2.3 Implosives

Voiced coronal implosives are articulated quite differently from voiced front coronal and voiced retroflex plosives. 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, p.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 (1977, p.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, p.60) states that it is 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 and 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 and 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.

When we compare the characteristics of an implosive to those 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 airflow 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 and 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 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 and 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[s]" (Ladefoged, 1971, p.27).

From this we can conclude that larynx lowering and ingressive airflow are not 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

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.

are often simply described as glottalized or laryngealized stops in the literature on Chadic, as pointed out by Clements and Rialland (2008, p.59).

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 proposed for alveolar and retroflex articulations in section 2.2. Whereas the plain–retroflex continuum is one that differs in place of articulation, 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 6.

2.4 Developments of retroflexes from implosives

On the affinity between retroflexes and implosives, Greenberg (1970, p.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." Such retracted implosives then tend to lose their glottalic feature, a development repeated in (3a). Haudricourt (1950) assumes a similar development, though his explanation seems not plausible to us: according to Haudricourt, the negative air pressure (due to the larynx lowering) causes a vacuum which tends to suck in the mobile tongue tip.

(3) a) ɗ > ᶑ > ɖ Haudricourt, Greenberg
b) ɗ > ɖ Bhat, Ohala

The descriptions by Bhat (1973) and Ohala (1983) in (3b) do 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 of implosives is caused by cavity enlargement, where the tongue tip is retracted due to a lowering of the tongue.

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, p.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 realization of the voiced implosive as pulmonic retroflex and thus a neutralization 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, where the coronal implosive is usually assumed to be a reflex of Proto-Bantu *d (Clements & Rialland, 2008, p.60; Guthrie, 1967–1971).

In the following section, we provide evidence for diachronic changes of 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 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 come 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 (section 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 retroflex is released with frication (Grimes, 2006, p.4).[4] The closely related Sabu (or Sawu(nese), Hawu, Havu) is spoken on the neighbor-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 neighboring 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, p.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 neighboring languages (Jonker, 1903, p.86). The exact place of articulation of the PMP sound is difficult to determine and depends to a large extent on what has

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

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

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 neighboring 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).

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 t^h d d^h ts ts^h dz dz^h/ (Ebert, 1997, p.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 $ʈ^h$] are marginal and do not contrast with other coronals (Ebert, 1994; Lahaussois, 2003, p.1).

If we compare Thulung words having a voiced retroflex to cognates in neighboring 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]

5 The discussion on Kiranti is restricted to initial consonants. Other Kiranti languages like Limbu and Camling also have retroflex consonants in this position, but almost only in loanwords from Nepali Driem, 1987, p.27; Ebert, 1997, p.14. The Western Kiranti language Jero seems to be a case like Thulung because it has the phoneme /ɖ/ in native words. However, Opgenort (2005, p.59) describes that its use instead of /d/ "seems to be generally determined by personal style or preference." 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 (1997, p.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, p.766) lists Sunwar as a 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.

(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 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)–(6c), namely *t for the uniformly voiceless stops in (6c), *d for the sounds in (6b), and a preglottalized *ʔt for the sounds in (6a) (see Starostin, 1994 (as reported in Opgenort, 2005) and Opgenort, 2005; Michailovsky, 1994, assumes a glottalized segment at a later stage). Michailovsky (1994, p.770) points out that the reconstruction of a preglottalized segment is somewhat speculative since there is no direct evidence for it. Opgenort (2005, p.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. Whether one reconstructs a *t or a *ʔt for the segments in (6a), 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 interstage 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 Kenya. A number of East Cushitic languages are reported to have a voiced retroflex stop /ɖ/, namely Afar (Bliese, 1981), Somali and Rendille (Sasse, 1979, p.25; Lloret, 1995, p.69). The related languages Boni, Arbore and Elmolo have instead an alveolar implosive usually transcribed as /d'/, see Sasse (1979, p.25). Sasse (1979, p.25) 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) ᶑ 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, p.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 neighboring 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 favor 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 assumes to have 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~ɽ] 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/.

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. How such variation can eventually lead to a sound change is elaborated in section 5.

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, therefore the alveolars /t, d/ are the only coronal stop phonemes and 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) were 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 seems to be no restrictions on the typological occurrence of plain stops, retroflexes and implosives.

We tested whether the voiced stop in German is realized in a way that favors retroflexion, that is, with a more retracted articulator and place of articulation, with less palatal contact, with a lower tongue and lower jaw position, and with more contextual variation than /t/. In addition, we measured in the acoustic signal whether /d/ has lower second and third formant transitions (as has been reported for retroflexes, see Hamann, 2003, pp.58–59 for an overview) and a shorter closure duration than /t/ (see discussion in section 2.2).

4.1 Methods

In order to test the above-mentioned articulatory 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 the first coil (placed midsagittally approximately 1 cm behind the tip). Tongue back (tb) movement was associated with the posterior coil (placed at the posterior end of the tongue where it touches the soft palate). Since this coil 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 below 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.

The occlusal plane was defined by a custom made t-bar which was inserted in the subject's mouth when clenching the teeth together. Two coils were glued midsagittally on the t-bar, one at the anterior part and one at the posterior part. The articulatory data are translated and rotated using the data from the occlusal plane in order to set the final origin and orientation of the coordinate system. This procedure was adapted from Hoole (1996). The mean of the tilt values and their standard variations of our data are reported in Fuchs (2005). They are all in a reliable range.

Four German subjects were recorded, three male (speakers 1–3) and one female (speaker 4). The speech material consisted of nonsense words [gəC_1VC_2ə] 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 contexts 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. Each sentence was repeated 10 times in a randomized order. The measured tongue sensor signals are composed of both the tongue and the jaw, since decomposition is not a straightforward process. For further details of the study, see Fuchs (2005).

On the basis of the EMA data we labeled for the consonant 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 highest vertical position for the consonant we also labeled the lowest vertical position (or most backward position in /u/ context) for the vowel.

For both the highest vertical tongue position and the lowest vertical tongue position we calculated the *tongue tip angle*. This is a measure introduced by Tiede, Gracco, Shiller, Espy-Wilson, and 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). We calculated first the angle of the line between tongue mid and tongue back, cf. (9), and of the line between tongue tip and tongue mid, cf. (10), separately.

(9) $\alpha_1 = 180/\pi * \text{arctan2}\ (x_{td} - x_{tm}, y_{td} - y_{tm})$

(10) $\alpha_2 = 180/\pi * \text{arctan2}\ (x_{tt} - x_{tm}, y_{tt} - y_{tm})$

The tongue tip angle is then calculated by subtracting α_2 from α_1. In cases where the result was negative, we added 360; in cases where the result was larger than 360 degrees, we subtracted 360. This angle is 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.[8]

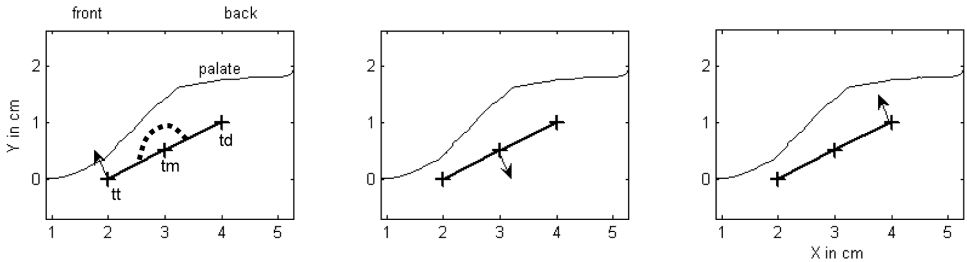
Figure 1 also shows that a retroflex tongue configuration according to this measure is possible not only with an upward movement of the tongue tip (graph on the left), but also with a lowering of the tongue mid (graph in the middle) and an upward movement of the tongue dorsum (graph on the right).

We restricted our articulatory measurements to two static points in time rather than including dynamic measures. The highest vertical position of the tongue tip is sufficient to determine the retroflex status of a consonant, and the lowest vertical position of the tongue tip during the articulation of the vowel is sufficient to provide information on possible co-articulatory effects. Though dynamic measures could provide insight into further differences between retroflex and front apical articulations, they would go beyond the scope of the present article.

8 Ladefoged and Bhaskararao (1983) use a similar measure, namely the angle between "the mean slope of the surface of the blade of the tongue and that of the front of the tongue" (p.296) in x-ray data by Hindi and Telugu speakers to estimate the degree of plosive retroflexion, but do not provide reference points for their definition of blade and front of the tongue. Simonsen et al. (2000, 2008) introduce an "r tip y" value to evaluate the degree of retroflexion in Norwegian sounds, which is based on the first three coils in EMA recordings. This measure is calculated as the distance of the tongue tip coil from the line drawn between the tongue mid and the tongue dorsum coil, and thus differs minimally from Tiede et al.'s tongue tip angle.

Figure 1

Schematic representation of Tiede et al.'s (2005) tongue tip angle for retroflex tongue configurations



In the acoustic comparison we measured for both /d/ and /t/ the values of the second and third formant at the end of the vowel (i.e., at the transition from vowel to stop), and also the closure duration of the plosives.[9]

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

4.2 Results

We discuss the results in the following order: the horizontal position of the tongue tip in §4.2.1, the frequency of tongue palatal contacts in section 4.2.2, the vertical position of the tongue dorsum and the jaw in section 4.2.3, the tongue tip angle in section 4.2.4, and the acoustic measurements in section 4.2.5.

For the analysis of the articulatory data we had to exclude 8 of the 240 tokens (4 speakers * 2 phonemes * 3 contexts * 10 repetitions) because speaker 2 realized 7 tokens and speaker 4 one token of /t/ as [ɾ] (intervocalic flapping). In the acoustic analysis we had to exclude a further 10 tokens because they were realized as approximants (2 tokens of /d/ by speaker 1, and 5 tokens of /d/ and 3 tokens of /t/ by speaker 2).

4.2.1 Retracted tongue tip position for /d/

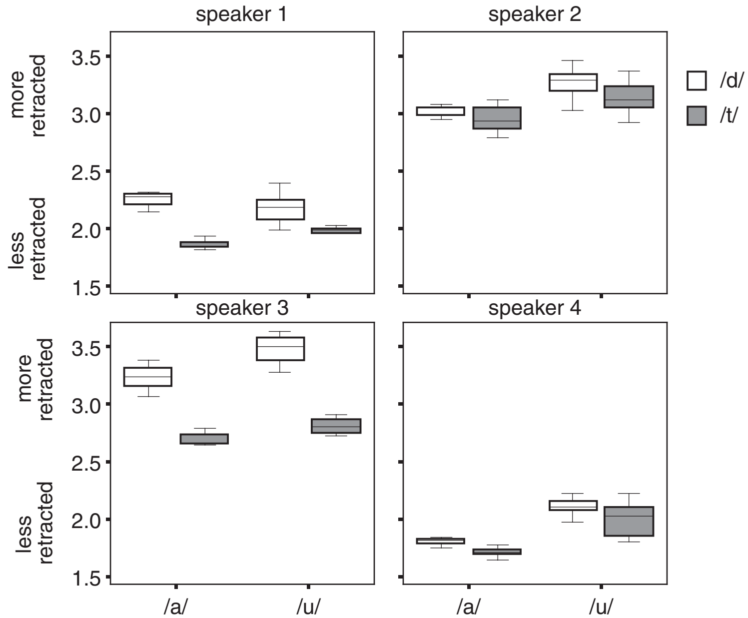
Figure 2 displays the results for the highest horizontal position of the tongue tip during the consonant in the context of /a/ and /u/ based on EMA, where higher values correspond to a more retracted place of articulation. It clearly shows that all speakers realize a significantly more retracted tongue 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/.

The EMA data in Figure 2 provide us with information on the position of the tongue tip and its place of articulation in the mid sagittal plane. However, the actual

9 Mark Tiede pointed out to us that the formant structure of the release bursts might yield interesting differences. Unfortunately, the bursts of the voiced plosives we recorded were extremely weak (probably due to their unstressed position) and did not allow for such a comparison.

Figure 2

Boxplots with standard deviations for highest horizontal position of the tongue tip for /d/ (white) and /t/ (gray) for the four speakers and /a u/ contexts; lower values indicate a more fronted articulation



amount of tongue palatal contact over the whole palate can only be gained from EPG data, as discussed in the following section.

4.2.2 Area of contact for /d/

In Figure 3, we see EPG frequency plots, which show palatal contact patterns at the time when the tongue tip reached the highest consonantal position, averaged over all repetitions. The four columns in Figure 3 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.

The EPG data in Figure 3 show that /d/ has generally a more retracted place of articulation than /t/, in both /a/ and /u/ context.[10]

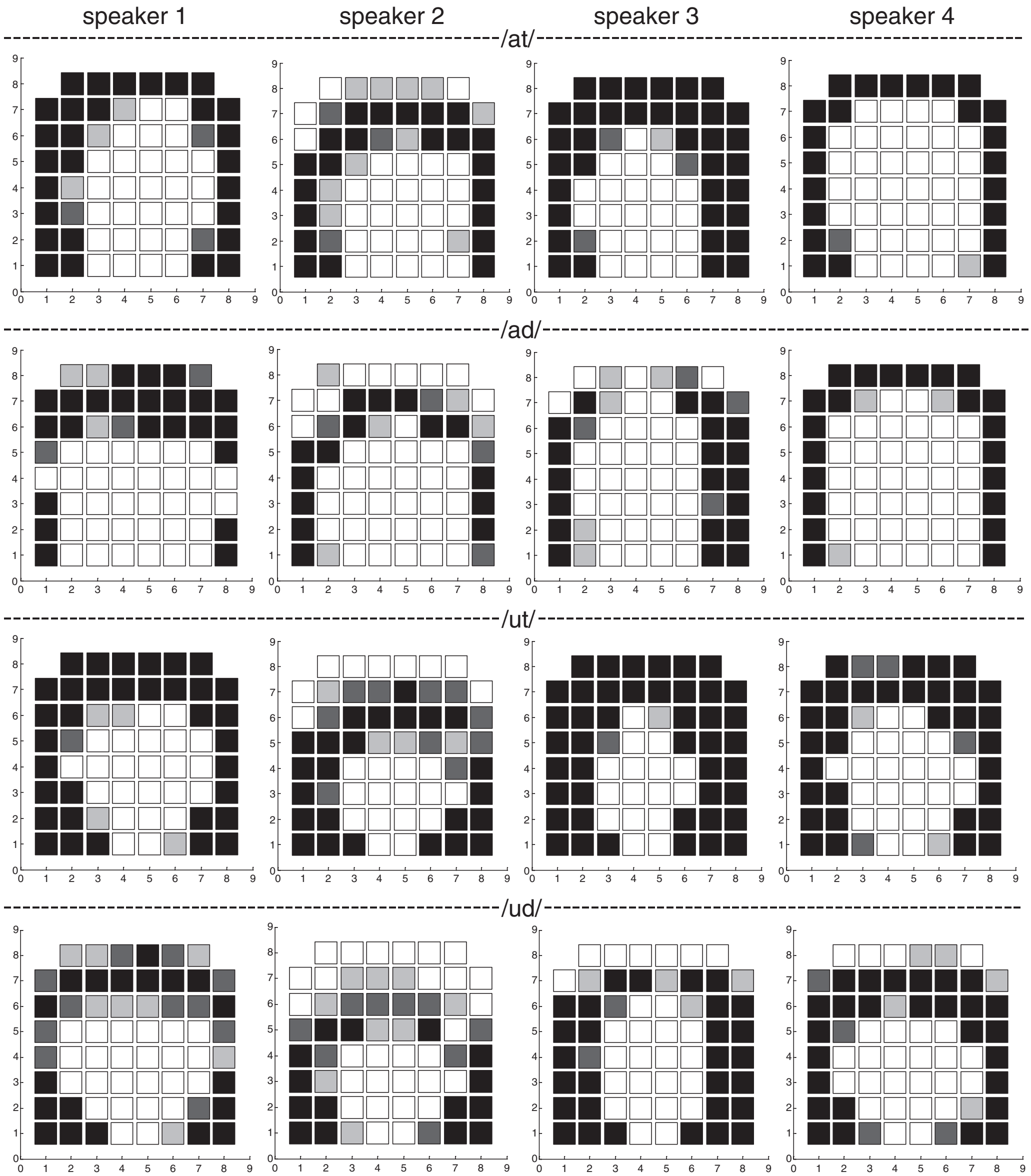
The percentage of contact over the whole palate for all speakers is significantly greater for /t/ than for /d/ (all descriptive statistics and significances are given in

10 For speaker 3 we can observe that there is no complete closure on the artificial palate for /ad/. This can be caused by three factors. First, the contact (and therefore a complete closure) might have been at the teeth, and therefore would not be reported with EPG, though this seems unlikely in the light of the EMA data. Second, this speaker might have realized the stop in /ad/ as approximant, though inspection of the acoustic recordings does not confirm this. And third, the tongue tip pellet made the articulation of a full closure impossible.

For speaker 4 we see an almost identical place of articulation for /d/ and /t/ in /a/ context.

Figure 3

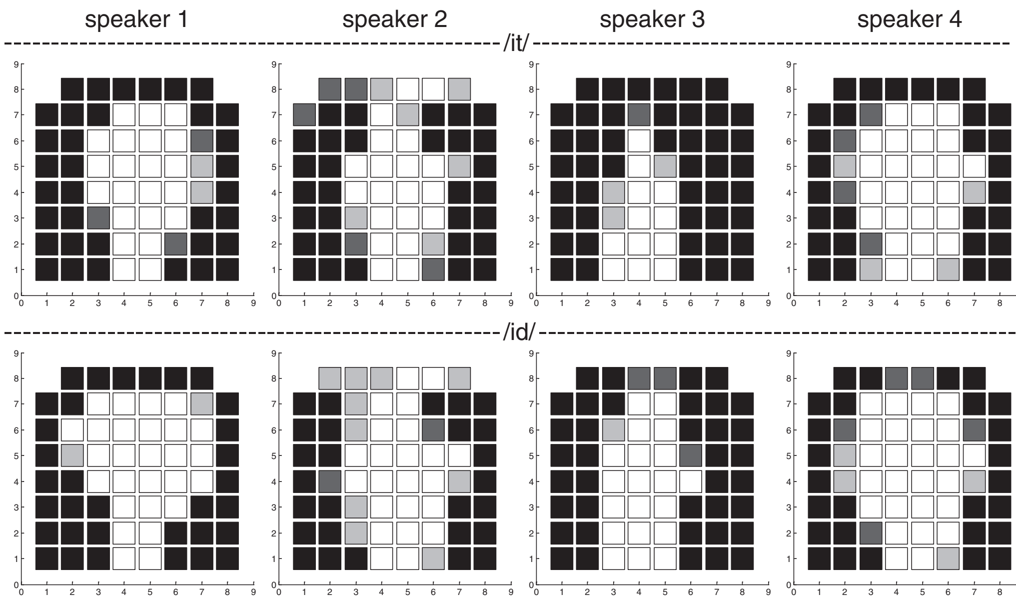
EPG frequency plots for all speakers (4 columns) in /a/ context (first 2 rows) and /u/ context (last 2 rows); /t/ = 1st and 3rd row, /d/ = 2nd and 4th row; black markers correspond to 76–100% tongue palatal contact with respect to all the subject's repetitions, dark gray markers to 51–75%, light gray to 26–50%, and white markers to 0–25%



Appendix II), 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/, and a difference

Figure 4

Same as Figure 3, but for /i/ context



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).

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.

The vowel /i/ thus exerts a larger co-articulatory influence on the following stop than the vowels /u/ and /a/.

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/? To answer this question, we took a univariate ANOVA with tongue dorsum *y* position and jaw *y* position as dependent variables and phoneme (/d/ versus /t/) and vowel context (/a, i, u/) as independent factors. Data were split by speaker (the descriptive statistics and significances are given in Appendix III). For the vertical tongue dorsum position we found a main effect of vowel context for all speakers, and a main effect of phoneme for speakers 1 and 2. All but speaker 3 show an interaction between the two factors: in /a/ context, /d/ is realized with a lower tongue dorsum position than /t/. In /i/ context, both consonants have a similar tongue position (except for speaker 1 who shows a slightly higher tongue 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 the vertical tongue position is to a large degree influenced 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 a main effect of vowel context 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 that can separate a retroflex tongue configuration from a bunched one, independent of the actual phonological retroflexion of the sound. The values we measured for the tongue tip angle of our four speakers are summarized in Figure 5, and the descriptive results and significance values are given in Appendix IV.

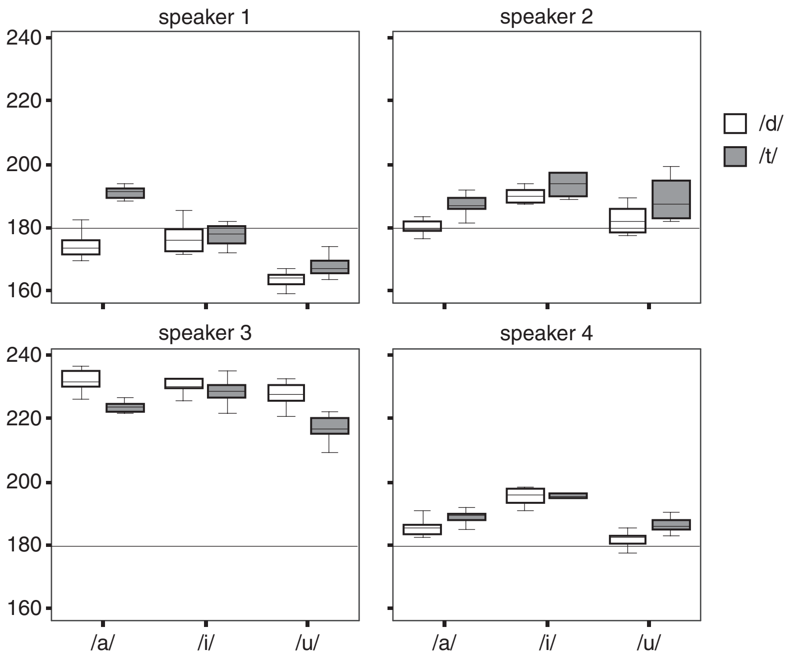
Figure 5 clearly shows that speaker 1 behaves differently from the rest in having a retroflex tongue tip angle for almost all tokens except those with /at/, with largest retroflexion for tokens with /u/. Speaker 3, on the other hand, is an exception in having very high (and thus no retroflex) tongue tip angles for all his tokens. Speakers 2 and 4 show a retroflex angle for some tokens with /ud/, and speaker 2 additionally for /ad/. The exact tongue tip angle thus seems to be speaker dependent, with some speakers generally using a more bunched tongue configuration (speaker 3), others a more retroflex tongue configuration (speaker 1).

For all speakers we found a dependence of tongue tip angle on the phoneme, with /d/ yielding lower values than /t/ for all but speaker 3. Additionally, speakers 1, 3 and 4 show an interaction of phoneme with vowel: lowest values for the tongue tip angle were achieved for the /ud/ tokens, apart from speaker 3, who had lowest values for the /ut/ tokens.

Since a retroflex tongue tip angle might already be articulated during the vowel, we included the calculation of tongue tip angle at the lowest tongue position for the preceding vowels. Again, we found that speaker 3 has very high values for the tongue tip angle and no retroflexion in any of the vowels. Speaker 2 also showed no retroflex tongue tip angle for the vowels. For speaker 4 we found retroflex tongue tip angles for most tokens of /a/, but this does not result in a retroflexion of the following stop. Only speaker 1 shows any co-articulatory effect of the vowel: he has a retroflex-like tongue configuration for the vowel /u/, which can explain the largest retroflexion in the following stops (both voiced and voiceless).

Figure 5

Boxplots with standard deviations for tongue tip angle for /d/ (white) and /t/ (gray) for the four speakers and /a i u/ contexts; black line indicates an angle of 180 degrees, values below the black line indicate a retroflex angle



In sum, there are large speaker-dependent differences in the tongue tip angle, but all but one speaker showed lower values for the voiced stop, for two speakers this was especially so in the context of /u/.

4.2.5 Acoustics of /t/ and /d/

To determine any retroflex-like acoustic property of the voiced compared to the voiceless stop, we measured the values of the second and third formant at the transition of the vowel into the following consonant. In addition, we measured the closure duration of the plosive. The descriptive statistics are included in Appendices V–VII for F2, F3 and duration, respectively.

Let us start with the second formant. Unsurprisingly, F2 was dependent on the preceding vowel for all four speakers, with lowest F2 values in /a/ context. For speakers 2 and 4, F2 also depended on the phoneme, with lower F2 values for a following /d/. Speaker 3 showed an interaction of vowel and phoneme: F2 values were lower for /d/ than for /t/, especially in /u/ context. Thus, apart from speaker 1 all showed a lower F2 for the voiced stop.

For the third formant, we observed a vowel dependent difference for all four speakers (highest F3 values for /i/) apart from speaker 1, who has similar F3 values for /a/. Speakers 2 and 3 furthermore show an interaction of vowel and phoneme: they both have much lower values for /d/ than for /t/ in the context of /u/. For the /ud/-items we can also observe more variation in the realization, see the large standard deviations compared to those for the /ut/-items in Appendix VI.

The acoustic closure duration was for all four speakers dependent on the phoneme, with shorter durations for /d/ than for /t/. For speaker 2 the closure duration was also affected by vowel context, with shorter duration in /i u/ than in /a/ context. Speaker 3 showed an interaction between vowel context and phoneme, with a longer /t/, especially in /a/ context.

Our data hence support the findings of earlier studies that voiced plosives have shorter closure durations than their voiceless counterparts. In addition, three of the four speakers show lower F2 values for /d/ than for /t/ (for one speaker this is restricted to /u/ context), and two speakers show lower F3 values for /ud/-tokens, indicating more retroflex-like acoustic characteristics.

4.3 Discussion

Our data show that there is a systematic articulatory 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), and 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 position, 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 favors retroflexion.

In /i/ context, we could observe significant differences between /d/ and /t/ only for jaw position. This influence of vowel context coincides with previous observations that retroflex tongue configurations avoid /i/ context, 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 speakers 2 and 4 we found a retroflex tongue tip angle for some tokens of /ud/, and for speaker 2 also for some tokens of /ad/. The first observation is in line with the fact that /u/ context can lead to retroflexion (see discussion in section 2.2). Speaker 1 showed a retroflex tongue tip angle for the majority of tokens, whereas speaker 3 showed no retroflex tongue tip angle at all.

In sum, our data illustrate that German /d/ is articulated in a way that favors retroflexion, and that retroflexion is an acceptable articulation of /d/ in German, despite the fact that German is traditionally described as having alveolar plosives (Wängler, 1974, pp.127f.; Ten Cate, Jordens, & van Lessen Kloeke, 1976, p.42; Kohler, 1995, p.159, the latter describes them as dental or alveolar).

In the acoustic analysis we saw that the voiced plosives have shorter closure durations than their voiceless counterparts. This has been described in the literature before for voiced plosives, and is in accordance with the shorter closure duration of retroflexes compared to front coronal plosives (cf., section 2.2).

The acoustic analysis further showed that three speakers have lower F2 values for /d/ than for /t/ (for one speaker this is restricted to /u/ context). Two speakers show lower F3 values for /ud/-tokens. Both a lower F2 and a lower F3 value are the acoustic correlates of a more retracted, that is, retroflex, place of articulation. For one speaker (speaker 1) we found a difference neither in F2 nor in F3 values between /d/ and /t/, which is not surprising since this speaker had retroflex articulations for almost all of his coronal plosive tokens. The acoustic data therefore support our articulatory findings and show some variation in the realization of the voiced stops, especially in /u/ context.

5 How can variation lead to sound change?

The German data in section 4 showed that there is variation in the articulation of the voiced coronal stop, including retroflex realizations. Since we present this articulatory variation in the context of languages that developed a retroflex from an alveolar voiced stop (section 3), we will elaborate in the following how such a variation can indeed lead to a sound change, without implying that German is on its way to developing a retroflex allophone or phoneme.

Let us first look at infants acquiring their native language(s). Newly acquiring infants have to construct the categories (phonemes) of their language on the basis of the auditory input they receive (for an illustration see, e.g., Maye & Gerken, 2000; Maye, Werker, & Gerken, 2002). This happens by focusing on how category tokens are distributed along the most invariable and robust auditory dimension such as first, second, or third formant, or duration. Variation in the input caused by different realizations of one category, like the variable realizations of /d/ we found in our data, yields an unreliable auditory dimension, such as for instance F3 in our case. The learning children will tend to pay little attention to such unreliable information. In a first step, thus, variation in the articulation leads to the weakening of the corresponding perceptual dimension as a cue for a category. We thus observe a development from an older generation, where for instance high F3 was a reliable cue to alveolar stops, to a younger generation, where F3 varies between high and low values and is not employed to cue alveolar stops. When starting to articulate themselves, the children of the younger generation will choose among the possible realizations corresponding to the variation along the perceptual dimension (F3) based on other factors, such as ease of articulation. It seems likely that in our example the learning children will opt for more retroflex articulations of /d/ in the context of /u/ for co-articulatory reasons.

Once the majority of (a group of) speakers have adapted a retroflex articulation in one context (for reasons why such an active change in the articulation of adolescents and adults should take place, see below), the input distribution changes, and is likely to show a more regular, predictable distribution: retroflex articulations occurring in /u/ context, alveolar ones elsewhere. This can then be interpreted by the newly acquiring child as regularity, resulting in an analysis in terms of two allophones.

Of course, acquiring infants are not the only (or even the main) initiators of sound change. We are all aware of how important our pronunciation is in determining our social and regional identity. Sociolinguistic factors are thus the driving force in sound changes initiated by adolescents and adults. A continuous updating of our pronunciation according to the input we receive is actually happening all the time and leads to small but noticeable changes in our sound system. A well-known example for this is the British queen's pronunciation, which altered over the last 50 years, showing signs of adjustment to the changes in the Received Pronunciation of English in her surroundings (Harrington, Palethorpe, & Watson, 2000).

6 Conclusion

In this study we looked at the question whether a voiced retroflex stop can develop from a front coronal voiced stop as single retroflex sound in a language. Our aim was to illustrate that this development occurs and is independent of the earlier observed development of retroflexes from implosives described by Haudricourt (1950), Greenberg (1970), Bhat (1973), and Ohala (1983). We proposed and tested a possible phonetic motivation for this process, namely the articulatory affinity between voiced front coronal stops and voiced retroflex stops.

Our diachronic examples of changes that introduced a voiced retroflex /ɖ/ via a front coronal stop are from Central Malayo-Polynesian (Dhao), Sino-Tibetan (Thulung), and East Cushitic (Afar, Somali, Rendille). For the Central Malayo-Polynesian language Dhao, the literature agrees that *d is the proto-segment corresponding to the present-day retroflex. For the other two language groups a preglottalized coronal stop is reported, either voiced (East Cushitic) or voiceless (Sino-Tibetan). We have to keep in mind that the diachronic descriptions of these languages are sparse, and that the proposed developments are often not motivated, like the drop of glottalization or the change in voicing. Further data are necessary to establish the reconstructed segments.

The phonetic explanation of the diachronic change was tested in an EPG and EMA experiment with four speakers of German. We found that German /d/ shows a more retracted 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 Mandarin, Norwegian, Hindi or Tiwi, where retroflexion is less pronounced than in for instance Dravidian languages. The common characteristics between voiced front stops and voiced retroflex stops support our hypothesis that voiced alveolar and retroflex articulations are similar to each other and can be said to form an articulatory continuum without a sharp boundary (see Ladefoged & Bhaskararao, 1983, p.299). Three of our four speakers actually produced retroflex [ɖ] tokens as a realization of the voiced alveolar stop phoneme in low and back vowel context. Our articulatory data thus support the phonetic explanation we proposed for the typological diachronic processes: the articulatory similarity between voiced front coronal stops and retroflex stops can account for the diachronic development of the former into the latter.

The acoustic analyses of the German data are in accordance with the articulatory findings. Tokens of /d/, especially in /u/ context, show lower F2 and/or lower F3 values in the transitions from vowel to consonant than tokens of the voiceless stop, indicating a more retracted, thus retroflex-like, articulation.

Our study supplements the work by Haudricourt (1950) and followers who describe the development of voiced retroflex stops 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) and a change in airflow mechanism, the articulatory similarity we propose here is not based on a change in airflow and holds for processes in both directions. 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 saw 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 in Boni /ɗ/ might stem from *ɖ in Proto-Sam (as proposed by Heine, 1978). An established 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).[11]

Changes from voiced coronal implosive stops to retroflexes and vice versa and those from pulmonic front coronal stops to retroflexes and vice versa can all be accounted for by the articulatory-similarity space proposed in section 2.3, which is based on two continua: one from a plain voiced stop to a voiced stop with rapidly and strongly lowered larynx (implosive), as proposed by Ladefoged (1964 et seq.), and one from a plain voiced stop to a voiced retroflex stop (based on Ladefoged & Bhaskararao, 1983, p.299). The two continua are orthogonal to each other and create a two-dimensional space, including a continuum from plain to implosive retroflex. Any change from one voiced coronal segment to another within this space is simply due to articulatory similarity (on either one or both dimensions). The similarity does not imply a preferred direction of change. 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 enlarges our proposed similarity space by a third dimension.

References

ANDERSEN, T. (1987). An outline of Lulubo phonology. *Studies in African Linguistics*, **18**(1), 39–66.

ANDERSON, V. B., & MADDIESON, I. (1994). Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics*, **87**, 131–162.

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

BAIRD, L. (2002). Kéo (Illustrations of the IPA). *Journal of the International Phonetic Association*, **32**(1), 93–97.

BAKKER, P., SMITH, N., & VEENSTRA, T. (1995). Saramaccan. In J. Arends, P. Muysken, & N. Smith (Eds.), *Pidgins and Creoles: An Introduction*. Amsterdam: Benjamins.

BELL-BERTI, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America*, **57**, 456–461.

11 Smith and Haabo (2007) find the circular diachronic development of Proto-Volta-Congo *ɗ to Proto-Gbe/Fon ɖ to Saramaccan ɗ "unexpected" (p.118) and propose that the implosive articulation of this sound has been continuous. This contrasts with the general assumption that Proto-Gbe had a (non-implosive) retroflex (Capo, 1991).

BENDER, M. L. (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, J. (2004). *Evolutionary Phonology*. Cambridge: Cambridge University Press.
BLEVINS, J. (2007). The importance of typology in explaining recurrent sound patterns. *Linguistic Typology*, **11**(1), 107–113.
BLIESE, L. F. (1981). *A Generative Grammar of Afar*. Dallas , TX: Summer Institute of Linguistics.
CAPO, H. B. C. (1991). *A Comparative Phonology of Gbe*. Berlin: Foris Publications.
CATFORD, J. C. (1939). On the classification of stop consonants. *Le Maître Phonétique*, **65**, 2–5.
CATFORD, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.
CLEMENTS, G. N., & OSU, S. (2003). Explosives, implosives, and nonexplosives: The linguistic function of air pressure differences in stops. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp.299–350). Berlin: Mouton de Gruyter.
CLEMENTS, G. N., & RIALLAND, A. (2008). Africa as a phonological area. In B. Heine & D. Nurse (Eds.), *A Linguistic Geography of Africa* (pp.36–85). Cambridge: Cambridge University Press.
DART, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, **79**, 1–155.
DART, S. N. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics*, **26**, 71–94.
DEMOLIN, D. (1988). Some problems of phonological reconstruction in Central Sudanic. *Belgian Journal of Linguistics*, **3**(Special issue on Phonological reconstruction ed. by M. Dominicy and J. Dror), 53–96.
DEMOLIN, D., & GOYVAERTS, D. L. (1986). Some aspects of Madi phonology and morphology. *Antwerp Papers in Linguistics*, **44**(Special issue on language and history in Central Africa ed. by D. L. Goyvaerts), 89–103.
DEMPWOLFF, O. (1934). Vergleichende Lautlehre des Austronesischen 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. P. (1990). Linguotectal contact patterns in the dental and retroflex stops of Hindi. *Journal of Phonetics*, **18**, 189–201.
DIXIT, R. P., & FLEGE, J. (1991). Vowel context, rate and loudness effects of linguopalatal contact patterns in Hindi retroflex /ʈ/. *Journal of Phonetics*, **19**(2), 213–229.
DIXON, R. M. W. (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.
DRIEM, G. L. V. (1987). *A Grammar of Limbu*. Berlin: Mouton de Gruyter.
DYEN, I. (1971). The Austronesian languages and Proto-Austronesian. In T. A. Sebeok (Ed.), *Linguistics in Oceania* (pp.5–54). The Hague: Mouton.
EBERT, K. (1994). *The Structure of Kiranti Languages: Comparative Grammar & Texts*. Zürich: Universität Zürich.
EBERT, K. (1997). *A Grammar of Athpare*. München: Lincom.
EBERT, K. (2003). Kiranti languages: An overview. In G. Thurgood & R. L. Lapolla (Eds.), *The Sino-Tibetan Languages* (pp.505–517). London: Routledge.
EMENEAU, M. B. (1944). *Kota Texts. Part 1*. Berkeley & Los Angeles: University of California Press.

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

EWAN, W. G., & KRONES, R. (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics*, **2**, 327–335.

FARNETANI, E. (1989). An articulatory study of "voicing" in Italian by means of dynamic palatography. *Proceedings of the Speech Research International Conference, Linguistic Institute of Hungarian Academy of Sciences, T. Szende, Budapest*, 395–398.

FARNETANI, E. (1990). V-C-V lingual coarticulation and its spatiotemporal domain. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp.93–130). Dordrecht: Kluwer.

FLEMMING, E. (2003). The relationship between coronal place and vowel backness. *Phonology*, **20**, 335–373.

FOX, J. J. (2004). Current Developments in Comparative Austronesian Studies. *Symposium Austronesia*, *Bali*. http://rspas.anu.edu.au/people/personal/foxxj_rspas/Comparative_Austronesian_Studies.pdf

FUCHS, S. (2005). Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS Working Papers in Linguistics*, **41**(8).

FUCHS, S., PERRIER, P., GENG, C., & MOOSHAMMER, C. (2006). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In J. Harrington & M. Tabain (Eds.), *Towards a Better Understanding of Speech Production Processes* (pp.149–164). New York: Psychology Press.

FUCHS, S., PERRIER, P., & MOOSHAMMER, C. (2001). The role of the palate in tongue kinematics: An experimental assessment in VC sequences from EPG and EMMA data. *Proceedings of Eurospeech 2001, Aalborg,* **2**, 1487–1490.

FUJIMURA, O., & MILLER, J. E. (1979). Mandible height and syllable-final tenseness. *Phonetica*, **36**, 263–272.

FUJIMURA, O., TATSUMI, I. F., & KAGAYA, R. (1973). Computational processing of palatographic patterns. *Journal of Phonetics*, **1**(1), 47–54.

GORDON, R. G. J. (2005). *Ethnologue: Languages of the World* (15th ed.). Dallas: SIL International.

GOYVAERTS, D. L. (1986). Glottalized consonants: A new dimension? *Antwerp Papers in Linguistics*, **44**(Special issue on language and history in Central Africa ed. by D. L. Goyvaerts), 105–113.

GRAGG, G. (1976). Oromo of Wellegga. In M. L. Bender (Ed.), *The Non-Semitic Languages of Ethiopia* (pp.166–221). East Lansing: Michigan State University Press.

GREENBERG, J. H. (1970). Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics*, **36**(2), 123–145.

GRIMES, C. E. (2006). Hawu and Dhao in eastern Indonesia: Revisiting their relationship. *10th International Conference on Austronesian Linguistics, Puerto Princessa*, *Philippines*. http://www.sil.org/asia/philippines/ical/papers/Grimes-Hawu_Dhao.pdf

GUTHRIE, M. (1967–1971). *Comparative Bantu*. Farnborough, Hants: Gregg International.

HAMANN, S. (2002). Retroflexion and retraction revised. *ZAS Working Papers in Linguistics*, **28**, 13–25.

HAMANN, S. (2003). *The Phonetics and Phonology of Retroflexes*. Utrecht: Utrecht Institute of Linguistics.

HAMANN, S. (2004). Retroflex fricatives in Slavic languages. *Journal of the International Phonetic Association*, **34**(1), 53–67.

HARRINGTON, J., PALETHORPE, S., & WATSON, C. I. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association*, **30**(1/2), 63–78.

HAUDRICOURT, A.-G. (1950). Les consonnes préglottalisées en Indochine. *Bulletin de la Société de Linguistique*, **46**(1), 172–182.

HEINE, B. (1978). The Sam languages: A history of Rendille, Boni and Somali. *Afroasiatic Linguistics*, **6**(2).

HOOLE, P. (1996). Issues in the acquisition, processing, reduction and parameterization of articulographic data. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, **34**, 158–173.

JESSEN, M. (1998). *Phonetics and Phonology of Tense and Lax Obstruents in German*. Amsterdam: Benjamins.

JONKER, J. C. G. (1896). Bimaneesche Spraakkunst. *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen (VBG)*, **48**(2).

JONKER, J. C. G. (1903). Iets over de taal van Dao. In *Album Kern (Opstellen geschreven ter eere van Dr. H. Kern)* (pp.85–89). Leiden: Brill.

KEATING, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, **60**(2), 286–319.

KHATIWADA, R. (2007). Nepalese retroflex stops: A static palatography study of inter- and intra-speaker variability. *Proceedings of the Interspeech 2007, Antwerp*, 1422–1425.

KLAMER, M. (1998). *A Grammar of Kambera*. Berlin: Mouton de Gruyter.

KOHLER, K. (1995). *Einführung in die Phonetik des Deutschen* (2nd ed.). Berlin: Erich Schmidt.

KRULL, D., LINDBLOM, B., SHIA, B.-E., & FRUCHTER, D. (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, P. (1964). *A Phonetic Study of West African Languages: An Auditory-instrumental Survey*. Cambridge: Cambridge University Press.

LADEFOGED, P. (1971). *Preliminaries to Linguistic Phonetics*. Chicago, IL: University of Chicago Press.

LADEFOGED, P., & BHASKARARAO, P. (1983). Non-quantal aspects of consonant production: A study of retroflex consonants. *Journal of Phonetics*, **11**, 291–302.

LADEFOGED, P., & MADDIESON, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

LADEFOGED, P., & WU, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, **12**, 267–278.

LAHAUSSOIS, A. (2003). Thulung Rai. *Himalayan Linguistics*, **1**, 1–25.

LAVER, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

LEX, G. (2006). *Le Dialecte Peul du Fouladou (Casamance—Sénégal): Etude Phonétique et Phonologique*. München: Lincom.

LINDBLOM, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, **35**, 1773–1781.

LISKER, L., & ABRAMSON, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, **20**(3), 384.

LISKER, L., & PRICE, P. J. (1979). Context-determined effects of varying closure duration. In J. J. Wolf & D. H. Klatt (Eds.), *Speech Communication Papers* (pp.45–48). New York: Acoustical Society of America.

LLORET, M.-R. (1995). Implosive consonants: Their representation and sound change effects. *Belgian Journal of Linguistics*, **9**, 59–72.

LÖFQVIST, A., & GRACCO, V. L. (2002). Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America*, **111**, 2811–2827.

MADDIESON, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

MAYE, J., & GERKEN, L. (2000). Learning phonemes without minimal pairs. *Proceedings of the 24th Annual Boston University Conference on Language Development*, 522–533.

MAYE, J., WERKER, J. F., & GERKEN, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**, B101–B111.

MICHAILOVSKY, B. (1994). Manner vs. place of articulation in the Kiranti initial stops. In H. Kitamura & E. Al (Eds.), *Current Issues in Sino-Tibetan Linguistics* (pp.766–772).

Osaka: Organizing Committee of the 26th International Conference on Sino-Tibetan Languages and Linguistics.

MOEN, I., & SIMONSEN, H. G. (1997). Effects of voicing on /t, d/ tongue/palate contact in English and Norwegian. *Proceedings of ESCA. Eurospeech 97, Rhodes*, 2399–2402.

MOOSHAMMER, C., HOOLE, P., & GEUMANN, A. (2006). Inter-articulator cohesion within coronal consonant production. *Journal of the Acoustical Society of America*, **120**(2), 1028–1039.

MOOSHAMMER, C., HOOLE, P., & GEUMANN, A. (2007). Jaw and order. *Language and Speech*, **50**, 145–176.

OHALA, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The Production of Speech* (pp.189–216). New York: Springer.

OHALA, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical Linguistics: Problems and Perspectives* (pp.237–278). London: Longman.

OHALA, J. J. (2005). Phonetic explanations for sound patterns: Implications for grammars of competence. In W. J. Hardcastle & M. Beck (Eds.), *A Figure of Speech: A Festschrift for John Laver* (pp.23–38). London: Erlbaum.

OPGENORT, J. R. (2005). *A Grammar of Jero: With a Historical Comparative Study of the Kiranti Languages*. Leiden: Brill.

PERKELL, J. S. (1969). *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Cambridge, MA: MIT Press.

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

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

ROSS, M. (1995). Some current issues in Austronesian linguistics. In D. T. Tryon (Ed.), *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies* (pp.45–120). Berlin: Mouton de Gruyter.

SASSE, H.-J. (1979). The consonant phonemes of Proto-East-Cushitic (PEC): A first approximation. *Afroasiatic Linguistics*, **7**, 1–67.

SCHUH, R. G. (2003). Chadic overview. In M. L. Bender, G. Takacs, & D. L. Appleyard (Eds.), *Selected Comparative-historical Afrasian Linguistic Studies. In Memory of Igor M. Diakonoff* (pp.55–60). München: Lincom.

SIMONSEN, H. G., MOEN, I., & COWEN, S. (2000). Retroflex consonants in Norwegian: Are they really? Evidence from EMA and EPG. *Proceedings of the 5th Seminar on Speech Production: Models and Data, Kloster Seeon, Germany*, 113–116.

SIMONSEN, H. G., MOEN, I., & COWEN, S. (2008). Norwegian retroflex stops in a cross linguistic perspective. *Journal of Phonetics*, **36**, 385–405.

SMITH, N., & HAABO, V. (2007). The Saramaccan implosives: Tools for linguistic archaeology? *Journal of Pidgin and Creole Languages*, **22**(1), 101–122.

STAROSTIN, S. A. (1994). The reconstruction of Proto-Kiranti. *27ème Congrès International sur les Langues et la Linguistique Sino-Tibébtaines*, *Paris*.

STERIADE, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (Eds.), *The Role of Speech Perception in Phonology* (pp.219–250). San Diego, CA: Academic Press.

STEVENS, K., KEYSER, S., & KAWASAKI, H. (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and Variability in Speech Processes* (pp.426–449). Hillsdale: Erlbaum.

STEWART, J. M. (1995). Implosive, homorganic nasals and nasalized vowels in Volta-Congo. In E. N. Emenanjo & O.-M. Ndimele (Eds.), *Issues in African Languages and Linguistics: Essays in Honor of Kay Williamson* (pp.162–169). Aba: National Institute for Nigerian Languages.

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

ŠVARNÝ, O., & ZVELEBIL, K. (1955). Some remarks on the articulation of the 'cerebral' consonants in Indian languages, especially in Tamil. *Archiv Orientalni*, **23**, 374–407.

TEN CATE, A. P., JORDENS, P., & VAN LESSEN KLOEKE, W. U. S. (1976). *Deutsche Phonetik: Laut- und Aussprachelehre für Niederländer*. Groningen: Wolters-Noordhoff.

TIEDE, M. K., GRACCO, V. L., SHILLER, D. M., ESPY-WILSON, C., & BOYCE, S. E. (2005). Perturbed palatal shape and North American English /r/ production. *Journal of the Acoustical Society of America*, **117**(4), 2568–2569.

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

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

TUCKER, A. N., & BRYAN, M. A. (1966). *Linguistic Analyses, the Non-Bantu Languages of Northeastern Africa*. Oxford: Oxford University Press.

WAKUMOTO, M., MASAKI, S., HONDA, K., & OHUE, T. (1998). A pressure sensitive palatography: Application of new pressure sensitive sheet for measuring tongue-palatal contact pressure. In *Proceedings of the 5th International Conference on Spoken Language Processing* (pp.3151–3154). Sydney.

WÄNGLER, H.-H. (1974). *Grundriss einer Phonetik des Deutschen* (3rd ed.). Marburg: N.G. Elwert Verlag.

WATSON, R. L. (1991). Moru-Ma'di orthographies. In M. L. Bender (Ed.), *Proceedings of the Fourth Nilo-Saharan Linguistics Colloquium (Nilo-Saharan: Linguistic Analysis and Documentation Vol. 7)* (pp.273–282). Hamburg: Helmut Buske Verlag.

WESTBURY, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America*, **73**(4), 1322–1336.

WOLFF, J. (1974). Proto-Austronesian *r and *d. *Oceanic Linguistics*, **13**, 77–121.

WOLFF, J. (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 (in cm) at the highest vertical position for the consonant

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
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	*Factors*	*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 highest vertical position for the consonant

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
Speaker 1	a	10	40.32	4.30	10	50.81	4.45
	i	10	52.90	6.49	10	62.26	4.31
	u	10	44.84	13.14	10	59.03	3.82
Speaker 2	a	10	28.23	11.91	10	45.48	6.62
	i	10	52.10	11.78	4	58.87	3.36
	u	10	38.55	7.03	9	49.28	7.04
Speaker 3	a	10	40.48	6.24	10	62.10	3.25
	i	10	64.84	6.17	10	75.97	4.13
	u	10	50.65	4.70	10	68.23	14.43
Speaker 4	a	10	37.90	1.90	9	38.71	2.42
	i	10	52.15	5.23	10	54.84	6.76
	u	10	50.97	6.04	10	56.77	6.49

Subject	*Factors*	*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	19.67	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 of tongue dorsum and jaw (both in cm) at the highest vertical position for the consonant

		Tongue dorsum				*Jaw*			
		/d/		/t/		/d/		/t/	
Subject	*Vowel*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Speaker 1	a	0.56	0.10	0.92	0.11	–1.67	0.09	–1.15	0.06
	i	1.67	0.04	1.56	0.08	–1.35	0.06	–1.22	0.06
	u	1.30	0.07	1.32	0.09	–1.36	0.09	–1.11	0.07
Speaker 2	a	0.11	0.09	0.30	0.15	–1.36	0.03	–1.32	0.06
	i	0.66	0.08	0.68	0.15	–1.27	0.03	–1.26	0.02
	u	0.43	0.08	0.47	0.08	–1.23	0.03	–1.22	0.02
Speaker 3	a	0.93	0.23	0.92	0.06	–0.94	0.03	–0.93	0.01
	i	1.47	0.11	1.50	0.07	–0.97	0.04	–0.93	0.02
	u	1.35	0.06	1.43	0.06	–0.93	0.03	–0.91	0.02
Speaker 4	a	0.12	0.07	0.23	0.06	–1.10	0.05	–1.01	0.05
	i	1.02	0.11	0.98	0.04	–1.09	0.03	–1.05	0.03
	u	0.81	0.06	0.72	0.06	–0.98	0.02	–0.94	0.02

			Tongue dorsum		*Jaw*	
Subject	*Factors*	*df*	*F*	*p*	*F*	*p*
Speaker 1	vowel	2	15.15	0.000	15.15	0.000
	phoneme	1	40.45	0.000	40.45	0.000
	vowel * phoneme	2	0.67	n.s.	0.67	n.s.
Speaker 2	vowel	2	15.64	0.000	15.64	0.000
	phoneme	1	19.67	0.000	19.67	0.000
	vowel * phoneme	2	1.37	n.s.	1.37	n.s.
Speaker 3	vowel	2	32.97	0.000	32.97	0.000
	phoneme	1	75.63	0.000	75.63	0.000
	vowel * phoneme	2	2.51	n.s.	2.51	n.s.
Speaker 4	vowel	2	55.38	0.000	55.38	0.000
	phoneme	1	5.10	0.028	5.10	0.028
	vowel * phoneme	2	1.14	n.s.	1.14	n.s.

Appendix IV: Descriptive statistics for the tongue tip angle (in degrees) at the highest vertical position for the consonant

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
Speaker 1	a	10	174	3.6	10	191	2.9
	i	10	177	4.5	10	178	3.1
	u	10	164	2.2	10	168	3.1
Speaker 2	a	10	180	2.0	10	187	4.1
	i	10	189	3.8	4	194	4.3
	u	10	183	4.4	9	190	6.6
Speaker 3	a	10	232	3.3	10	223	2.7
	i	10	231	3.4	10	228	3.7
	u	10	227	3.8	10	217	4.2
Speaker 4	a	10	186	2.6	9	189	2.1
	i	10	195	2.5	10	196	2.8
	u	10	182	2.4	10	186	2.3

Subject	*Factors*	*df*	*F*	*p*
Speaker 1	vowel	2	134.924	0.000
	phoneme	1	68.993	0.000
	vowel * phoneme	2	29.218	0.000
Speaker 2	vowel	2	12.587	0.000
	phoneme	1	21.903	0.000
	vowel * phoneme	2	0.408	0.667
Speaker 3	vowel	2	24.209	0.000
	phoneme	1	61.795	0.000
	vowel * phoneme	2	7.272	0.002
Speaker 4	vowel	2	114.779	0.000
	phoneme	1	16.199	0.000
	vowel * phoneme	2	3.435	0.040

Appendix V: Descriptive statistics for second formant values (in Hz) at the transition between vowel and consonant

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
Speaker 1	a	10	1427	92	10	1358	106
	i	10	1851	228	10	1946	147
	u	10	1817	370	10	1888	420
Speaker 2	a	10	1416	35	10	1520	249
	i	10	1910	48	4	2138	133
	u	10	1522	323	9	1941	183
Speaker 3	a	10	1323	301	10	1223	619
	i	10	2167	276	10	2216	359
	u	10	1479	485	10	2216	141
Speaker 4	a	10	1667	39	9	1764	228
	i	10	2271	362	10	2480	97
	u	10	1490	203	10	1682	241

Subject	*Factors*	*df*	*F*	*p*
Speaker 1	vowel	2	22.310	0.000
	phoneme	1	0.216	0.644
	vowel * phoneme	2	0.555	0.577
Speaker 2	vowel	2	26.928	0.000
	phoneme	1	16.646	0.000
	vowel * phoneme	2	2.598	0.087
Speaker 3	vowel	2	27.708	0.000
	phoneme	1	5.040	0.029
	vowel * phoneme	2	6.405	0.003
Speaker 4	vowel	2	73.056	0.000
	phoneme	1	8.277	0.006
	vowel * phoneme	2	0.361	0.699

Appendix VI: Descriptive statistics for third formant values (in Hz) at the transition between vowel and consonant

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
Speaker 1	a	10	2513	104	10	2462	298
	i	10	2406	437	10	2574	542
	u	10	2794	576	10	3087	258
Speaker 2	a	10	2688	70	10	2838	111
	i	10	2957	166	4	2978	121
	u	10	2304	522	9	2847	139
Speaker 3	a	10	2697	355	10	2806	721
	i	10	3229	390	10	3343	339
	u	10	2557	593	10	3351	334
Speaker 4	a	10	3179	88	9	3172	329
	i	10	3360	605	10	3377	371
	u	10	2743	345	10	2755	234

Subject	*Factors*	*df*	*F*	*p*
Speaker 1	vowel	2	8.201	0.001
	phoneme	1	1.641	0.206
	vowel * phoneme	2	0.930	0.401
Speaker 2	vowel	2	7.804	0.001
	phoneme	1	9.611	0.004
	vowel * phoneme	2	4.111	0.024
Speaker 3	vowel	2	6.345	0.003
	phoneme	1	7.529	0.008
	vowel * phoneme	2	3.388	0.041
Speaker 4	vowel	2	15.117	0.000
	phoneme	1	0.006	0.941
	vowel * phoneme	2	0.006	0.994

Appendix VII: Descriptive statistics for closure duration (in ms)

		/d/			/t/		
Subject	*Vowel*	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
Speaker 1	a	10	47	4	10	60	9
	i	10	55	14	10	68	15
	u	10	49	13	10	70	9
Speaker 2	a	10	43	7	10	53	7
	i	10	30	7	4	37	7
	u	10	32	7	9	31	4
Speaker 3	a	10	35	4	10	62	6
	i	10	35	7	10	56	6
	u	10	35	7	10	52	5
Speaker 4	a	10	48	6	9	56	9
	i	10	48	20	10	58	5
	u	10	51	5	10	58	9

Subject	*Factors*	*df*	*F*	*p*
Speaker 1	vowel	2	2.480	0.094
	phoneme	1	28.872	0.000
	vowel * phoneme	2	0.959	0.390
Speaker 2	vowel	2	31.000	0.000
	phoneme	1	7.177	0.011
	vowel * phoneme	2	3.015	0.061
Speaker 3	vowel	2	3.459	0.039
	phoneme	1	214.815	0.000
	vowel * phoneme	2	3.800	0.029
Speaker 4	vowel	2	0.322	0.726
	phoneme	1	9.145	0.004
	vowel * phoneme	2	0.046	0.955