# Supplementary file for “Pushes and pulls from below: anatomical variation, articulation and sound change” by Dan Dediu and Scott R. Moisik

## – Description of the *ArtiVarK* sample –

The *ArtiVarK project* is covered by the ethics approval 45659.091.14 (1 June 2015), Donders Center for Brain, Cognition and Behaviour, Nijmegen.

It contains data from 90 participants (35 female) from four umbrella ethnic backgrounds: “Chinese” (10 participants, 3 female, speakers of Sino-Tibetan languages), “North Indian” (15, 4 female, Indo-Aryan languages), “South Indian” (19, 3 female, Dravidian languages), and “European” (46, 25 female, mostly speakers of Dutch). The participants are generally young (mean age 25), highly educated, and most have very little prior formal training in phonetics (on a Likert scale from 1 = “no phonetic training” to 5 = “professional”, the median is 1.0 and the mean 1.7).

Our study comprised four sessions, the first three at the Donders Center: first, we introduced the participants to the study, they signed the informed consent forms, answered standardized questions concerning themselves, past orthodontic treatment and the languages they know, and received financial compensation at a standard rate and the reimbursement of travel expenses.

In the second session, we provided a standardized phonetic training (approximately 30 minutes) of producing potentially novel speech sounds (relevant here being the dental [ǀ] and the (post-)alveolar [!] clicks, and the standard North American [ɹ]) in the [a_a] phonetic context and with their sustained articulation.

Next, there was an MRI session (approximately 45 minutes) using a 1.5T MAGNETOM® Avanto Siemens scanner, of a structural T1 scan of the head and neck (T1 MPR NS PH8, TE=2.98ms, TR=2250ms, flip angle 15°, slice thickness 1mm, pixel spacing 1mm×1mm, FOV 256×256), followed by static sustained articulations in the [a_a] context (T2 TSE SAG P2 256, 60ms TE, 1020ms TR, 150° flip angle, 3mm slice thickness, 3.9mm interslice distance, 0.94mm×0.94mm pixel spacing, 256×256 FOV, and 10s acquisition time), ending with real-time articulation of five [a_a] and one “I said a_a for him” (TF2D15 retro iPAT, 1.04ms TE, 7.47ms TR, 10° flip angle, 5mm slice thickness, 2.45mm×3.38mm pixel spacing, 80×58 FOV, ≈6fps, and 15s acquisition time). Some participants could not take part in the MRI session for safety reasons.

Finally, we acquired high-resolution intraoral optical 3D scans of their hard palate using a TRIOS® 3shape system at the Department of Orthodontics and Craniofacial Biology, UMC Radboud, Nijmegen.

We designed a set of standardized and well-defined *landmarks* and *semi-landmarks*, inspired from biological anthropology and phonetics, that can be placed on such structural MRI and intraoral optical scans. From the 3D positions of these landmarks we can further derive *classical measures* (essentially angles and distances between such landmarks), or we can conduct *geometric morphometric* analyses that capture more global aspects of shape. These allow us to precisely quantify the metric variation between individuals in multiple aspects of the vocal tract.