

I. Annotation: PRAAT

Each target was hand-segmented by a careful examination of the waveform and the spectrogram, following the standard phonetic protocols for each segment. Four tiers were created- 'word', 'phone' and 'feature' and 'voicing' which were used according to the relevant phonetic category and acoustic feature, as described below.

Word-initial voiceless stop /t/: VOT

The boundary for the onset of the closure of the voiceless stop was placed at the offset of the visibly steady F2 of the preceding vowel (Turk et al., 2006). The entire steady duration of the vowel following the stop, identified by the onset/offset of the visually steady F2 of the vowel, was labelled as 'vow' in the 'phone' tier. The phone was labelled as 't' in the 'phone' tier. The entire duration of the vowel following the stop was coded as 'vow' in the 'phone' tier. For this phone category, the dependent variable was Voice Onset Time (VOT), for which two measurements were extracted separately: (1) the duration of the burst of the stop (from the segment labelled as 'burst' in the 'feature' tier; shown in Figure 1), and (2) duration of the aspiration of the burst (from the segment labelled as 'asp' in the 'feature' tier; shown in Figure 1). Finally, these two values were added to represent combined duration of the burst and aspiration which formed the dependent variable of Voice Onset Time (VOT) for this phone category. Please refer to Figure 1 for an example.

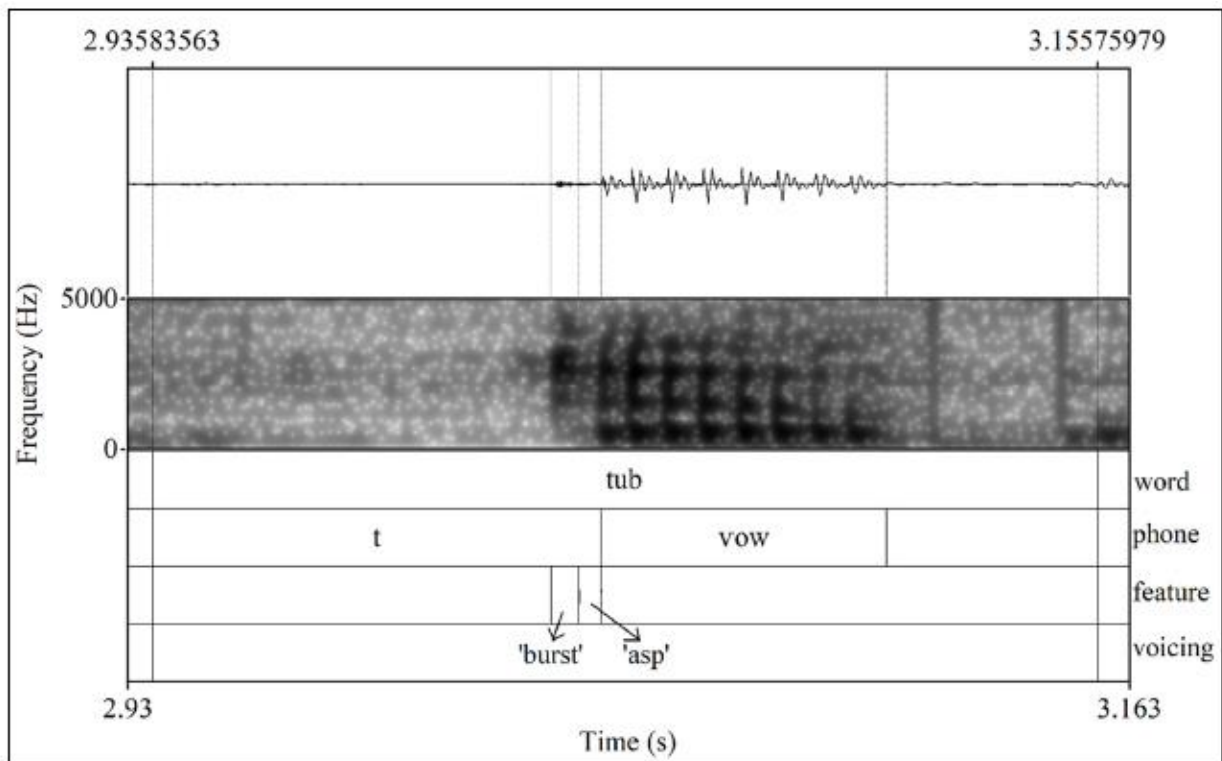


Figure S1. Annotation of word-initial /t/ in the word 'tub' as spoken by a male speaker from the control group 'Indian'

Word-initial voiced stop /d/: VOT

As suggested by Turk et al. (2012), the boundary for the onset of the closure of the voiced stop was placed at the offset of the visibly steady F2 of the preceding vowel. The entire steady duration of the vowel following the stop, identified by the onset/offset of the visually steady F2 of the vowel, was labelled as ‘vow’ in the ‘phone’ tier. Then the phone was labelled as ‘d’ in the ‘phone’ tier. After this, the segment from the onset of the release of the stop closure to the onset of voicing in the following vowel was labelled as ‘burst’ in the ‘feature’ tier. Please refer to Figure 2 for an example.

VOT was measured as the duration from the onset of the release of the stop closure to the onset of voicing in the following vowel (Cho & Ladefoged, 1999). For cases where there were multiple bursts, the onset of first burst was marked as the onset of release of the stop closure. This measure of VOT reflects the duration of the burst (Stuart-Smith et al., 2015) and this segment was labelled as ‘burst’ in the ‘feature’ tier in Figure 2.

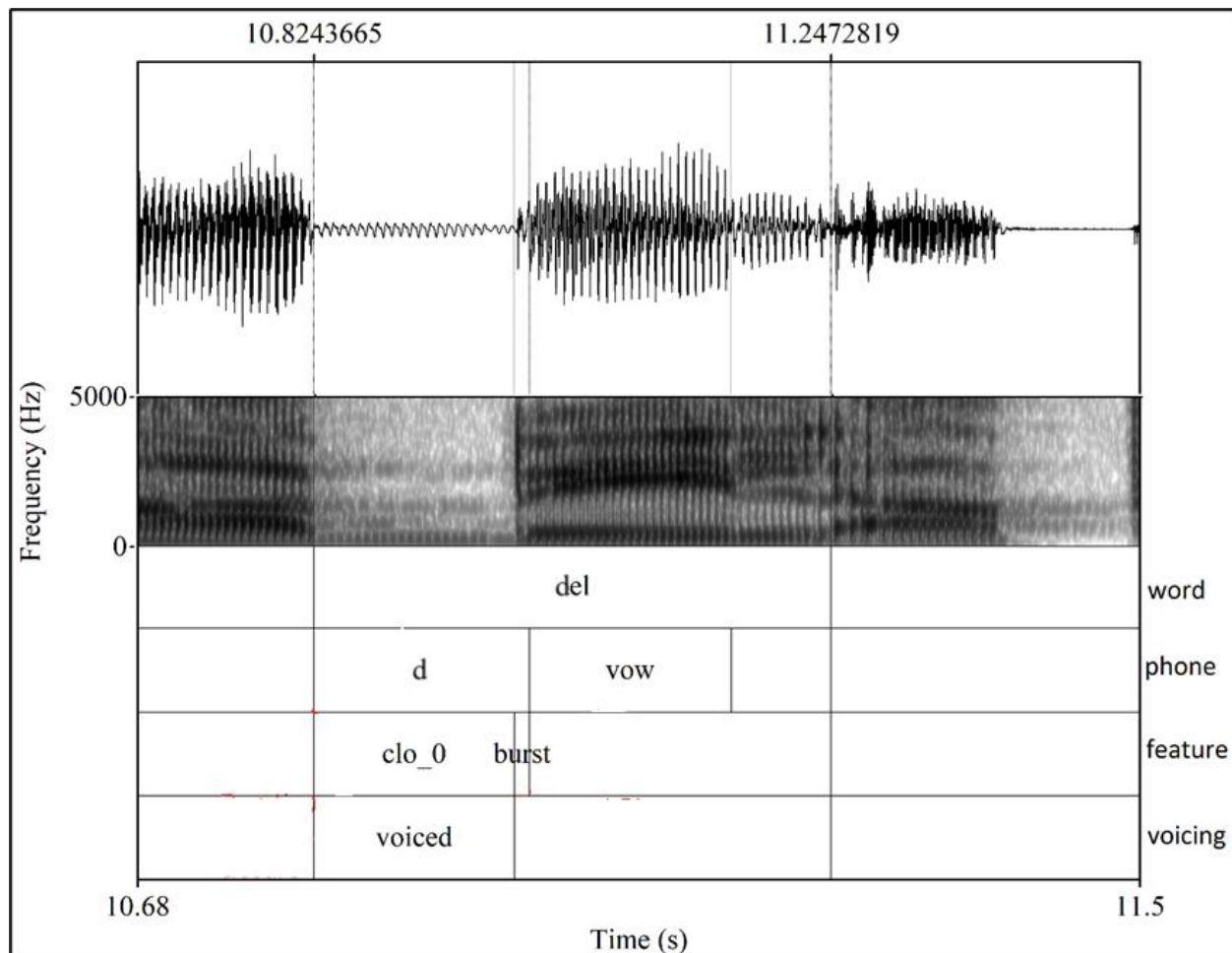


Figure S2. Annotation of word-initial /d/ in the Hindi word 'del' (meaning 'piece of rock') as spoken by a male speaker from the control group 'Indian'