Initial Geminates and Pertinacity

Joshua Booth & Aditi Lahiri, University of Oxford

Within Indo-European, perhaps the best-known historical gemination process is found in West Germanic (WGmc), where all consonants but /r/ were geminated when preceded by a light syllable and followed by /j/, e.g. OE *cynnes,* OHG *kunnes* 'race-GEN' (cf. Gothic *kunjis*). However, such processes are typologically common and a close parallel exists in Indo-Aryan (IA), where consonants were likewise geminated before the glides /j/ and /w/, e.g. Bengali [ʃo̪ːi] 'truth' (Sanskrit *satya*), [biʃːaʃ] 'belief' (Sanskrit *biswas*). This paper examines the historical pertinacity of the resulting geminates, drawing on theoretical evidence and a recent phonetic experiment conducted with speakers of modern Kolkata Bengali.

Unlike the majority of the WGmc dialects, which have lost WGmc's quantity contrast in consonants (but retained vowel length distinctions), Bengali has lost the vowel length contrast found in other IA languages, but preserved its geminate consonants, including those resulting from this gemination rule. There is no synchronic evidence in Bengali for the glides which originally triggered the gemination; rather, we suggest that they have been reanalysed as underlying geminates, with appropriately long closure duration (CD; the primary cue for length contrast, see Lahiri & Hankamer 1988; Hankamer et al. 1989).

Importantly, these [OBSTRUENT] + [w]/[j] clusters could occur both initially and medially (unlike WGmc, where they were medial or stem-final). However, modern Bengali allows very few word-initial clusters (largely [OBS]+[SONORANT], with a few words with initial [s]+[OBS] or [str] clusters). Accordingly, in line with typological tendencies, all *word-initial* geminates resulting from [OBS] + [w]/[j] clusters are pronounced as singletons, e.g. [ʃ]</ʃw/,[d̪]</dj/ etc. However, these clusters still survive in the orthography (as conjuncts, e.g. শ্ব ‹ʃw›, দ্ব ‹d̪w›, ত্য ‹t̪y›) and are pronounced as geminates medially: cf. শ্বাস ‹ʃwas› [ʃaːʃ] ~ আশ্বাস ‹aʃwas› [aʃːaʃ]. The hypothesis we explored was whether these clusters (e.g., <ʃw> or <d̪w>), which are geminates word medially, have been *reanalysed as underlying geminates everywhere, even though they always surface as singletons initially.* With word-initial geminates being crosslinguistically rare, this would be quite a 'nonesuch' phenomenon; yet, given that the medials have geminated, initial clusters warrant further examination.

We know from other languages that an unattached consonantal mora can be linked to a preceding coda and appear as a word-initial geminate (see Lahiri & Kraehenmann 2004; Kraehenmann & Lahiri 2008 for historical, articulatory and acoustic evidence for initial geminates Swiss German dialects). Acoustic analysis conducted in Kolkata revealed that such segments are indeed pronounced substantially longer in medial position than both singleton stops and the same clusters in absolute word-initial position. Most interestingly, however, when the initial clusters in words such as দ্বিতীয় [d̪it̪io] 'second' (CD 80ms) follow a vowel-final prefix, they surface as geminates, e.g. অদ্বিতীয় [ɔd̪ːit̪io] 'second to none' (CD 140ms), just as those appearing medially in simplex words, e.g. বিদ্যা [bid̪ːa] 'learning' (CD 168ms). This is similar to Swiss German initial geminates in a phrasal context.

We thus propose that Bengali orthographic [C+glide] clusters have been reanalysed as geminates *everywhere*; medially, they always appear as geminates, but in initial position, the unlinked mora is deleted. However, if a vowel-final prefix is added in complex forms, the underlying mora may be linked to the preceding syllable's coda and become visible. These geminates have remained remarkably pertinacious (particularly in comparison to WGmc). Once the context for the original gemination rule (i.e. the /w/ or /j/) was phonologically lost and no longer recoverable, these segments were reanalysed as *underlyingly* long. Even in word-initial position, there remains sufficient contextual evidence for learners to assign an underlying mora in their representations, which surfaces following a vowel-final syllable.