Sandhi and Syllables in Classical Sanskrit

Brett Kessler

Stanford University & Hewlett Packard Laboratories

Possibly the single most famous phonological fact about Classical Sanskrit[1] is that it has marvellously prolific sandhi. Virtually any sound segment at the end of a word is subject to assimilation to the first segment in the next word, regardless of syntactic phrasing. Some examples are shown in 1-3; but as seen in 4-6, the same changes appear not to apply word-internally, whether at morpheme junctures or not.[2]

(1) tat indriyam > tad indriyam 'that sense'
(2) tat manas > tan manaḥ 'that mind'
(3) tat cetas > tac cetaḥ 'that intellect'

(4) mahat- 'great' + -ā 'instr. sg.' > mahatā *not* *mahadā
(5) ātmā 'self' *not* *ānmā
(6) ad- 'eat' + -mi 'pres. 1 sg. act.' > admi *not* *anmi

In the sandhi charts at the end of this article is presented a synopsis of all these changes. Of course these phenomena are so striking and well known that the Sanskrit word *sandhi* has become the standard linguistic term for word-juncture phonology. In this paper, however, I come to the terminologically unfortunate conclusion that Sanskrit does not have sandhi.[3] The phonological

phenomena that people have associated with word junctures are in fact conditioned by syllable structure. I will show that this is not merely an alternative way of describing the same phenomena but in fact accounts for a wider range of facts more parsimoniously. These results should give support to those who would limit prosodic juncture rules, and refute those who claim that the Sanskrit phenomena are artifacts introduced into the language by the native linguists—leaving no clear winners, since these are largely the same set of people.

The fact that these changes occur between all words in a sentence but not within a word has made an impression on linguists over the years. Selkirk (1980) used Sanskrit to showcase her theory of prosodic phonology, but others were more doubtful about these facts. Whitney (1964:35), the preeminent American Sanskritist, decided that no language could have so many changes between words, and concluded that the native grammarians had artificially influenced the language.

More recently, linguists have been looking at the other end of the same problem: given that these changes do occur between words without regard to phrasing or any lexical property of the words, why don't they occur within words as well? As Allen (1972:25-26) noted, one would expect that assimilation should be more likely to take place between segments that are closer together, not the other way around. Kaisse (1985) considered the Sanskrit data to be a real problem for her theory, which held that word-juncture effects ought to key off of some special property of the word or syntax. Since the Sanskrit sandhi applies between all words, regardless of syntax, she adopted Whitney's solution of declaring the phonology to be partly artificial.

But in a paper exploring the domains of phonological rules, Rice (1990:307-311) made the important observation that Sanskrit voicing assimilation can be described as a syllable-coda effect rather than as a word-juncture effect. The core insight is that words emerge from the lexicon with their final consonants in a syllable coda. Since codas are well known to be particularly subject to neutralization and assimilation, one might surmise

that it is the syllable structure rather than the word juncture that is the operant condition. An obstruent never voices before a vowel word-internally because it will always be in an onset: word-internal consonants would normally be expected to syllabify with the following vowel, by a syllabification principle that is well-known to apply in most if not all languages.

(7a) ma-hat + -ā > *max. onset syll.* ma-ha-tā, *to which no sandhi rules apply*
(b) ma-hat > *word* ma-hat *with /t/ in coda at word level; so* ma-hat-ā-khyā-nam 'great narrative' > *sandhi* ma-had-ā-khyā-nam

   I would like to show that Rice's claim about voicing assimilation is correct, and should be applied to all external sandhi in Sanskrit. This would answer the concerns raised by Whitney and the others without impugning the accuracy of the ancient Indian linguists. I claim that the highlighted entries in the sandhi charts—sequences that behave differently when a word boundary falls in the middle of them—are all situations where the adjacent segments would be syllabified tautosyllabically as onsets word-internally.

(8a) ya-śas- 'glory' + -as '(gen.)' > ya-śa-sas > *sandhi* ya-śa-saḥ
(b) ya-śas a-sti 'glory is' = ya-śas-as-ti > *sandhi* ya-śo-sti

   To make this claim convincing, I must show that if there are any word-internal heterosyllabic clusters, they are subject to external sandhi changes, since at least one of their elements is in a syllable coda. For this purpose I shall attempt a precise definition of the Sanskrit syllable structure. With this information, it will be seen that there are no word-internal clusters that fail to undergo external sandhi. This will constitute a sort of negative evidence in favor of the syllable-based theory of sandhi, in that it will serve to explain why certain sequences are not found inside a word. I shall then

adduce positive evidence that external sandhi rules are needed to account for changes of stems inside inflected and derived words.

Onsets

The most conservative way to begin to account for onsets is to enumerate the clusters found at the beginning of words. The following clusters appear at the beginning of words listed in [Monier-Williams (1899)](). Forms in parentheses are doubtful; it is not clear that words beginning with such clusters appear in any Classical Sanskrit texts.

(9) (kn), (km), (ky), kr, kry, kl, kv, kṣ, (kṣṇ), kṣm, kṣv, khy, (gn), gr, gl, ghn, ghr, cy, (chy), jñ, (jm), jy, (jr), jv, (jhy), (ṭv), (tm), ty, tr, try, tv, ts, dy, dr, dv, dvy, dhm, dhy, dhr, dhv, (dhvr), ny, nr, (nv), (pn), py, pr, pl, ps, br, (bl), (bhy), bhr, (bhl), mn, (my), mr, ml, (yv), (lp), (ly), (lv), vy, vr, vl, śc, ścy, śn, śm, śy, śr, śl, śv, (ṣṭ), ṣṭh, (ṣv), sk, skh, st, sty, str, stry, sth, sn, sp, spr, sph, sphy, sm, sy, sr, sv, hn, (hm), hy, hr, hl, hv

Using the sonority hierarchy proposed by Clements (1990) —obstruent < nasal < liquid < glide—it is clear that Sanskrit onsets allow segments to be followed by any segment of greater sonority.[4] However this rule does not in itself account for a large class of onsets. Several clusters begin with two obstruents, and some onsets have the glide /v/ appearing before liquids, which should be less sonorous than /v/:[5]

(10) śc, ṣṭh, sk, skh, st, sth, sp, sph
(11) kṣ, ts, ps
(12) vr, vl, mn

Following the terminology of [Cairns and Feinstein (1982)](), these initial segments belong to a PREMARGIN, which in the case of Sanskrit is at most one segment which precedes the head of the

onset.[6] This construct serves as a locus of sonority sequencing violations, and also explains why three-letter sequences of segments that are legal pairwise might be illegal as a triplet, e.g., that there are no */sts-/, */tsp-/, etc.

(13a) a [+cont] obstruent before a voiceless [-cont] obstruent:
  śc, sk(h), st(h), sp(h)
(b) a [-cont] obstruent before a voiceless [+cont] obstruent:
  kṣ, ts, ps
(c) a [+lab] sonorant before a [-lab] sonorant: vr, vl, mn

There is some morphological evidence that this one-segment restriction is an active rule of Sanskrit phonology and not just a historical survival from the parent language. The *s*-aorists normally insert an /s/ between the stem and the personal ending (*acchait-s-ma* 'we cut' , *anai-s-ta* 'ye led'), but this /s/ is not found if it would appear between two oral stops (*acchait-ta* 'ye cut').[7] It would appear that syllabification attempts to build up the onset /tst/ on the basis of pairwise licensing (/t/ is permitted before /s/, /s/ is permitted before /t/), then notes the single-premargin violation and deletes the /s/, possibly because it is the locus of the sonority sequencing violation. A more oblique piece of evidence for the premargin and the /s/-deletion rule can be seen in verb reduplication, where /s/-stop clusters behave anomalously: unlike all other circumstances, including /s/-sonorant clusters, where the initial segment is copied, /s/-stop clusters apparently copy the stop: *ta-sthau* 'stood' vs. *sa-smāra* 'remembered' . This was adduced by [Kiparsky (1979:434-435)](#) as evidence for structure within the onset, with reduplication copying the metrically weaker (less sonorous) part of the onset. It is also possible that the entire margin is copied at a deep level (**sta-sthau*), which would agree with the Gothic evidence he presents (*stai-staut* 'pushed' , even though the usual reduplication in that language is also uniconsonantal, cf. *gai-grot* 'wept'). But since the margin is almost always uniconsonantal, Sanskrit imposed a shape-invariant

Ca- pattern on reduplication. Again, the first repair strategy when a cluster competes for that single slot is to delete the /s/, leaving the attested *ta-sthau*. This is congruent with the historical account suggested by [Osthoff (1882)](#).

   So the correct formulation of the onset would appear to be that consonants may sequence in ascending sonority (or glide may follow glide), which sequence is optionally preceded by a premargin. The exact internal organization of the onset really isn't important to my thesis, as long as one accepts that the description results in the correct syllable divisions in the context of maximum-onset syllabification. But precisely because it is important that we agree on the syllabification, it is worthwhile to consider [Steriade's](#) proposal (1982), by which the grammar prefers to keep /s/ out of onsets. Her analysis was driven by the theoretical concern that these /s-/ clusters violate sonority sequencing, and was supported by the facts about reduplication and *s*-aorists. In reduplication, /s/ would be extrasyllabic and therefore invisible to the copying. In the *s*-aorist, the /s/ would not be able to syllabify in an onset before stops, nor in a coda after stops (see next section), so it would disappear.

   Part of the problem with that argumentation is that it is motivated by a universal claim about sonority sequencing in syllables, but the Sanskrit facts are Indo-Aryan innovations that do not generalize to other languages. In several Indo-European languages, including the closely related Avestan, one always reduplicates the initial /s/, and in Gothic of course the entire margin is copied [(Osthoff 1882)](#). This also brings to mind Old English alliteration, which normally matches words that have the same single initial consonant, but which in the case of /s/-stop clusters requires both segments to match [(Bright 1957:230)](#); this would be very odd if the /s/ is not even part of the syllable. Furthermore, the same facts that rule out *acchaitsta* should, if universal, rule out Eng. *capstan*. Since languages have so much variation in how /s/ is treated, much if not all of the simplicity Steriade gains by not having to give structure-

sensitive rules (e.g., that reduplication copies the margin, not the edge, of the onset; or that /s/ itself is specifically liable to deletion) is lost by having to give language-specific conditions about how extrasyllabic /s/ is treated (it would end up getting licensed postlexically at the left margin of a word, but deleted elsewhere), and about the relative ordering of /s/ deletion and Bartholomae's Law (*araudhsta > araudhta > arauddha*, [Whitney 1964:315](#).[8]

   Furthermore, it will be seen in the section on Morphological Evidence that the behaviour of stem-final sibilants in the morphology directly patterns with the predictions of my syllabification theory. If the ending is such that the sibilant can form an onset with it, the sibilant remains a sibilant, perhaps with some place assimilation; if the ending cannot form an onset, the sibilant is neutralized in ways paralleling the treatment at the end of utterances. For example, the stem *ās-* 'sit' retains the /s/ in forms like *āste* 'sits' and *āsmahe* 'we sit', but makes it discontinuant in *āddhve* 'ye sit'. This agrees with the fact that there are words that begin with /st/ and /sm/ but not /sdh/, and can most directly be explained by differences in syllabification.

   Finally but perhaps most crucially, it will be seen that the sandhi theory itself constitutes evidence for the syllabification of /s/. If I am correct in stating that sandhi effects are triggered by syllable structure, then sandhi such as /-s p-/ > /-ḥ p-/, contrasted with the persistence of word-internal /sp/ sequences, is to be explained by the fact that word-final /s/ are in syllable codas, but word-internally form onset clusters with /p/. If however /s/ before /p/ were always in a syllable coda, then the two cases could not be distinguished, and one would have to resort to invoking word boundaries in sandhi. I would argue that the evidence that all other sandhi is syllable-conditioned, as well as the indecisiveness of the evidence against /s/-stop onsets, should lead to the conclusion that the overall picture is simpler if one accepts these onsets.

Codas

   An inventory of word-final codas would be a good deal harder to compile by thumbing through a dictionary, but fortunately the 'doctrine of permitted finals' is a well-established part of Sanskrit grammar. [Whitney (1964:49-53)](#) teaches that for all practical purposes, a word on the surface can end in one of the sequences /(r)k/, /ṅ/, /(r)ṭ/, /(r)t/, /n/, /(r)p/, /m/, /l/, or /ḥ/. An examination of the sandhi tables suggests that a few small changes need to be made to this inventory to account for deeper levels. Although it is traditional to describe sandhi in terms of the surface forms the input words have in isolation, most generative linguists agree that in some cases one needs to assume somewhat different forms feeding sandhi; these are added in parentheses in the charts. So it would appear from the evidence of word-final consonants that prior to the postlexical phonology, Sanskrit syllable codas can end in virtually any consonant other than a palatal, or /ṣ/, or a voiced or aspirated obstruent. Examination of the inventory of onsets suggests that the restrictions on palatals and laryngeal features apply to any stop not followed by a tautosyllabic sonorant. Some bisegmental clusters of descending sonority are allowed (glide before any non-glide, non-obstruent before obstruent).
   That the lexical phonology handles most or all of the neutralizations needed to satisfy these coda restrictions is apparent from morpheme- or class-dependent variations. The fact that much reduction must be lexical even though the assimilation must be postlexical reinforces the view most recently championed by [Cho (1990)](#) that neutralization and assimilation are two separate processes.

(14a) Monomorphemic -rT: suhārd 'good-hearted' > suhārt
(b) Bimorphemic: a-bibhar-t 'he carried' > abibhar > abibha**ḥ**
(15a) Palatal > velar: vaṇij 'merchant' > vaṇik
(b) Palatal > retroflex: parivrāj 'mendicant' > parivrā**ṭ**

(16a) /ṣ/ > /ṭ/: dvi**ṣ**-ḍhve ʻye hate (mid.)ʼ > dvi**ṭ**-ḍhve > dvi**ḍḍ**hve

(b) /ṣ/ > /k/ before verbal /s/: dvi**ṣ**-se ʻthou hatest (mid.)ʼ > dvik**ṣ**e

Word-internal clusters

It turns out to be the case that these simple observations about word margins fully account for Sanskritʼs lexical syllable structure word-internally as well. To verify this, I wrote a program to extract all word-internal consonant clusters from a body of on-line Sanskrit texts.[9] Of these, the following are analyzable as onset clusters according to the criteria just discussed. Note that they fill in many apparent gaps where the onset criteria generated clusters that do not show up word-initially.

(17) Word-internal onsets: kṇ, kn, km, kmy, ky, kṣṇ, kṣmy, kṣy, kś, gṇ, gn, gny, gm, gmy, gy, gry, gv, ghny, ghy, chy, jm, jr, ṭhy, ḍy, ḍr, ḍv, ḍhy, ḍhr, ḍhv, ṇy, ṇv, tn, tny, tnv, tm, tmy, tvy, tsn, tsy, tsv, thn, thy, thv, dry, dhn, dhny, dhry, nv, pn, pny, pś, psy, by, bhn, bhṇ, bhy, bhv, mṇ, my, mv, yy, ly, lv, vṇ, vn, śp, śvy, ṣk, ṣkr, ṣṭ, ṣṭy, ṣṭv, ṣṭr, ṣṇy, ṣp, ṣpr, ṣph, ṣm, ṣy, ṣr, ṣv, ṣvy, stv [Plus those found word-initially]

The remaining consonant sequences are analyzable as codas followed by onsets. The following inventory lists those which are simply sequences of permitted codas and permitted onsets.

(18) Word-internal deep codas plus onsets, sandhi applying vacuously: k-k; k-c; k-t(h); k-p; ṅ-k; ṅk-t; ṅ-g; ñ-c; ñ-ch; ñ-j; ṭ-k; ṭ-ṭ; ṇ-ṭ(h); ṇ-ḍ; ṇ-ṇ; ṇ-m; t-k; t-t(h); t-p(h); t-sth; n-t(h); n-d; n-dh; n-n; n-m; p-t; p-p; m-p(h); m-b(h); m-m; r-k(h); r-g(h); r-c(h); r-j; r-ḍ; r-t(h); rt-t; r-d(h);

r–n; r–p(h); r–b(h); r–m; r–ś; r–ṣ; r–s; l–k; l–g; l–p; l–ph; l–b; l–m; l–l; s–s [Omitting extensions of onsets, such as /n–dr/ given /n–d/]

The theory that external sandhi applies at all syllable boundaries implies that the sandhi rules should apply to these, though vacuously. And such turns out to be the case. The only apparent problem is the case of /s–s/, where one would normally expect /ḥs/. This happens when an ending is added to an *s*-stem verb: *ās–* 'sit' + *–se* '(pres. mid. 2nd sg.)' > *āsse*, not **āḥse*. The lexicon seems to turn the two /s/ in these verb forms into a geminate, which would protect the first /s/ from coda disarticulation because of geminate integrity. It is not surprising that the treatment of the coda would be exceptional here, since verbs usually have special coda neutralizations for sibilants, especially before /s/, as illustrated in 16b.

Finally, here is the list of word-internal clusters for which sandhi changes must be assumed.

(19) gg < k–g; gd(h) < k–d(h); gbh < k–bh; cc(h) < t–c(h); jj < t–j; ḍg < ṭ–g; ḍḍh < ṭ–dh; ḍbh < ṭ–bh; dg(h) < t–g(h); dd(h) < t–d(h); db(h) < t–b(h); bj < p–j; bd(h) < p–d(h); ṃś < m–ś; ṃṣ < m–ṣ; ṃs < m–s; ḥś < s–ś; ḥs < s–s

So word-internal consonant clusters can safely be analyzed as being susceptible to the same sandhi rules as are consonants that meet at word junctures. In fact, one is better off using that analysis, since otherwise one would need extra rules to account for the word-internal assimilations seen above. This is a nice result, and the mechanisms needed to account for this theory are ordinary, everyday linguistic processes. Maximum-onset syllabification applies within words, but not between words. Then consonants remaining in codas are neutralized in ways that are partly sensitive to the morpheme or its class. Immediately before emerging from the lexicon, words ending in /r/ devoice it, and words ending in /n/ add an /s/.[10] Postlexically, certain

assimilations affect elements in syllable codas, regardless of their position within the word. Since these processes are so common; since this analysis frees us from having to attribute the cause of assimilation to phonological or syntactic separation (word boundaries); and since this accounts for the distribution of word-internal clusters as well, I would judge that all modern linguistic descriptions that describe Sanskrit external sandhi as being a word-juncture phenomenon are unnecessarily complex.

Lexical exceptions

There are a few morphs that take different forms depending on what sound is adjacent in the sentence. *sas* 'that' and its extension *eṣas* 'this' drop the /s/ before consonants in Classical Sanskrit: *sa putraḥ* 'that son', not **saḥ putraḥ*. On the other hand, the dual endings *-ai* (surface *-e*), *-ū*, *-ī*, and perhaps a few interjections, add a glide before vowels, resulting ultimately in hiatus (glide deletion and pause being the sources of hiatus in Sanskrit): *phalaiy admi > phale admi* 'I eat two pieces of fruit', not **phaledmi*. The word *aum* drops its /a/ after words ending in /-ă/: *tava um > tavom* 'thine *om*', not **tavaum;* this no doubt originated in an effort to ensure that the mystic syllable always surfaces as /om/. The words *ā* 'to' and *mā* 'not' add a /t/ before /ch/ in those dialects that have geminate /ch/ after all vowels word-internally, but word-initially normally a single consonant after long vowels and a geminate after short ones; this /t/ extension produces a geminate /ch/ as sandhi, and must have originated at a point when *ā* and *mā* were considered to form a prosodic word with the following verb, therefore causing gemination despite their long vowel: *āt chādaya* 'cover!' > *ācchādaya.*

But these are fundamentally different from the sandhi I have been discussing. The rule of *sas/sa* alternation is not a phonological property of /s/ at word boundaries, but a lexical property of the word *sas* itself, which happens to be sensitive to its phonological environment. It is a matter of allomorphy,

exactly comparable to the alternation *an*/*a* in English, where no one would claim there is a general synchronic rule deleting /n/ before consonants at word boundaries. Thus these effects—the only actual word-juncture effects in Sanskrit—satisfy [Kaisse's (1985)](#) requirement that they should key off of some lexical or syntactic property of the environment.

Morphological evidence

So far the evidence has mainly been negative: I have shown only that the inventory of word-internal clusters does not contradict the syllable-based theory of sandhi. In this section I address whether there is direct evidence that external sandhi rules come into play in the morphology.

One answer is that the traditional grammar has always recognized a large overlap between the rules of external and internal sandhi. Many of the unhighlighted entries in the sandhi charts also apply word internally. Among these are the rules that [Selkirk](#) analyzed as domain-span rules (1980:125) and include Glide Formation, Vowel Contraction, Assimilation of /m/, Anusvara (disarticulation) of /m/, and Obstruent Voicing (before other obstruents, which case she must separate from their voicing word-finally). The outlook I am advocating obviates the need to distinguish external rules from rules that are both external and internal: the distinction is a matter of syllable boundaries. It is particularly easy to see that sandhi rules affecting word-final vowels before other vowels should also find the same syllabic conditions when vowels come into contact word-internally, and indeed there are many examples of Glide Formation and Vowel Contraction in the morphology: *kāma-i* 'love (loc.)' > *kāme*; *kāma-as* 'loves (nom.)' > *kāmāḥ*; *āsya-ī* 'two mouths' > *āsye*; *gati-ā* 'gait (instr.)' > *gatyā*; *nadī-ā* 'river (instr.)' > *nadyā*; *tanū-e* 'body (dat.)' > *tanve*; *jyā-istham* 'old-est' > *jyeṣṭham*; *juhu-āthai* 'you two sacrifice (middle)' > *juhvāthe*; *abhava-i* 'I was (imperf. mid.)' > *abhave*. Of course, when a vowel precedes a consonant, there is also a syllable boundary

just as there is between words, but since the sandhi rules call for no change in this case, the congruence of internal and external sandhi is not particularly noticeable: *kāma-bhyas* '(dat./abl. pl.)' > *kāmabhyaḥ*.

   The situation is slightly obscured for stem-final consonants because the most common case in the morphology is for that consonant to form an onset with the ending. This of course bleeds the sandhi rules, which are coda-sensitive, and has led to the false generalization that most sandhi rules for consonants as the prior elements are inoperative word-internally. Thus consonant-final noun stems are unchanged before a vowel ending: *manas-ā > manasā,* not **manaā.* It turns out that the only noun endings beginning with a consonant are *-bhyām* '(instr., dat., abl. dual)' , *-bhis* '(instr. pl.)' , *-bhyas* '(dat., abl. pl.)' , and *-su* '(loc. pl.)' . /bh/ cannot be the second element of any onset cluster, so a preceding consonant will always end up in a coda, leaving the appropriate environment for sandhi changes: *yaśas-bhis* 'fame (instr. pl.)' > *yaśobhiḥ.* /s/ is the second element of an onset only after a tenuis (/k/, /p/, /t/, /ṭ/), so most of the time it too will close a syllable *yaśas-su > yaśaḥsu;* and the tenuis + /s/ sequence does not undergo any change at word boundaries, so this case is indistinguishable from external sandhi. These facts led to the false conclusion that these suffixes themselves are specially marked as causing word-boundary sandhi changes, and so they are called PADA suffixes (Sanskrit for 'word'). It is of course troubling that inflectional endings that have no morphological unity should be given the status of separate words, especially since the suffixes otherwise behave as part of the word: the suffix *-su* is subject to the rule of Ruki, the /s/ becoming retroflex after sounds such as /k/ (*vāc-su* 'voices' > *vāk-su > vākṣu*), which change in the classical language never occurs between words; and stem-final /r/ does not become /ḥ/ before *-su* (*dhur-su* 'yokes' > *dhūrṣu*), though it would before another word. In this new account, the suffixes do not need to be given any special marking, nor do they need to insert a dubious word boundary. The behaviour is

completely predictable from the phonology of the syllable structure.

Verbal inflections are more probative, since suffixes beginning with consonants are not all such as occasion syllable boundaries. For example, with the verb *dviṣ-* 'hate' one finds the /ṣ/ surviving intact in inflections such as these present active forms: *dviṣvaḥ* '(1st dual)', *dviṣmaḥ* '(1st pl.)', *dviṣṭhaḥ* '(2nd dual)', *dviṣṭaḥ* '(3rd dual)', *dviṣanti* '(3rd pl.)' —that is, in forms where it can form an onset cluster with the following segment. But voiced obstruents cannot be part of onset clusters, so before suffixes beginning with /dh/, one does not get the expected **dviṣḍhve* '(2nd pl. middle)' but rather *dviḍḍhve*.[11] Similarly, before /s/ suffixes one gets not **dviṣse* '(2nd sg. mid.)' but *dvikṣe* (with Ruki), because onsets cannot be built from multiple fricatives either. Of course most of these segments rejected from the onset cannot occur in codas either,[12] and so they are subject to certain neutralizations. As mentioned earlier, these neutralizations are partly idiosyncratic, though generally drawn from a very small set of possibilities. In the case of *dviṣ-,* the coda neutralization to /ṭ/ (whence /ḍ/ by normal voicing assimilation) is regular, a simple dropping of +cont. The change to /k/ in *dvikṣe* is a subregularity, normal for continuants before verbal endings in /s-/.

The above example makes it clear that it is not simply the case that consonantal endings occasion external sandhi. But one could conclude that verbal suffixes in /dh-/ and /s-/ should simply be listed among the pada suffixes. However, verb stems ending in /h/, a placeless voiced aspirate (phonetically, a breathy *h*), pattern somewhat differently. *duh-* 'milk', which is marked for conversion to velar in a coda (*go-dhuk*[13] 'cow-milker (nom. sg.)'), forms most inflections analogously to *dviṣ-*: *duhvaḥ, duhmaḥ, duhanti,* and before *s-* suffixes one sees the expected conversion to /k/: *dhukṣe*. But before *t(h)-* suffixes the situation is not analogous to *dviṣ-*, because /h/, being voiced, cannot form an onset cluster with obstruents. The effect is

obscured a bit by Bartholomae's Law, but the conversion of /h/ to a velar shows that coda neutralizations are taking place: *dugdhaḥ* for the 2nd and 3rd person dual. Before /t(h)/, an /h/ remains in a coda. In that position it must give up its laryngeal features, which dock on the following dental, and the /h/ must become articulated, which in this case means conversion to a velar.

The same principles are broadly true of derivation as well. For example, *prāc-* 'forward', which converts to a velar before the so-called pada suffixes, does so also before the comparative ending *-tara-*, since /ct/ is an impossible onset sequence: *prāktaram*. However, the so-called TADDHITA suffixes—those which are appended to derived stems rather than to bare roots—are a more complicated story. Those beginning with vowels behave as expected, inducing no change in stem-final consonants, but those beginning with consonants have a marked tendency to combine by the rules of external sandhi. Thus *manas-in* 'having a mind' > *manasin*, not *manain, but *manas-mayam* 'mental' > *manomayam*. By and large it is clear that these are being combined by the rules of compounding (a final /r/ changes to /s/ in words like *prātastanam* 'matutinal', which is strictly a word-margin effect, cf. *kartanam* 'excision'), but the fact that only consonantal endings induce external sandhi raises the suspicion that syllabification is involved. Conceivably this treatment originated in a tendency to apply empty-onset filling rather than maximum-onset syllabification as the syllabification principle at the lexical stratum at which taddhita suffixes are adjoined. Since this would almost always give the same results as compounding, this could easily have led to a restructuring by which the common consonantal suffixes were reanalyzed as bound words that require compounding.

Conclusion

It is curious that this theory was not proposed and accepted long ago. It is simpler and more explanatory than an account

invoking word boundaries, pada suffixes, and many exceptions in conjugation, and certainly the ancients knew about syllables. But the doctrine of word boundaries has been taught for millennia by the ancient Indian linguists, by the neogrammarians, and by the structuralists, down through the modern generativist prosodic phonologists, the only exception I know of being Rice's observation on voicing assimilation. I suspect that the greatest problem has been a decided ambivalence toward syllables as part of linguistic theory; they have largely been ignored in favour of linear segmental models, at least up until the introduction of autosegmental frameworks in the past couple of decades. But even if one were inclined to pay attention to syllables, Sanskrit does have a few features that can lead one astray. Most notably, the Sanskrit poetic metre, like that of Latin and Greek, is based on the pattern of heavy and light syllables, where a syllable is heavy if its vowel is long or followed by two or more consonants. The usual way of accounting for this is to provide that when two consonants follow a vowel, at least a short one, the first consonant syllabify with the preceding vowel, so that one can say that closed syllables are heavy. This is of course at first blush completely contrary to my claim that syllables are built word-internally only by maximizing the onset in accordance with the general onset templates. Nowadays one can explain these conflicts by rule ordering—perhaps a postlexical rule readjusts syllables after the sandhi rules take effect—or by ambisyllabicity, but the neogrammarians and early structuralists had no such devices.

I submit that current techniques of non-linear and metrical phonology have made it possible to offer a significantly more concise and explanatory account of Sanskrit external sandhi, one which does not require reference to word boundaries. If these results are accepted, Sanskrit will have to cease being adduced as the exemplar of a language with a rich sandhi in the strict sense of the word, and cease being faulted for being phonologically unnatural.

I hesitate, however, to extrapolate universal claims about prosodic phonology just because it turns out not to be directly

relevant to Classical Sanskrit sandhi: as far as I know, this language is not the linchpin in anybody's theory. I have offered these observations because of the prominent role the Sanskrit phenomena have played in the history of linguistics, from traditional discussions of sandhi on down through Selkirk's famous 'Sanskrit revisited' paper and beyond. It is noteworthy, however, that prosodic juncture rules seem to be rather rare, and rejecting the Sanskrit data as evidence of word juncture rules makes the data even thinner. I agree with Rice (1990) that it might be worthwhile looking more deeply into other cases of alleged prosodic juncture rules and examining how many of them might, like those of Sanskrit, be analyzable as caused by other conditions such as syllable structure.

Sandhi charts

The chief rule of external sandhi seen in the vowel chart is Glide Deletion (disregarded by Selkirk, 1980:113, but cf. Allen, 1972:38). With initial /a-/ is also seen the a-Deletion rule of Selkirk (1980:125), which Allen (1972:39-43) treats rather as a wholesale monophthongization of the sequences /aya/, /ava/. In the consonant charts are seen Final Voicing of obstruents (Selkirk 1980:115) and Stop To Nasal (p. 118). The change /as a/ > /o/ is analogous to a-Deletion for final vowels. Concentrating on Vedic, Selkirk ignores the dropping or nasalization of /s/ in final /ns/, and the disarticulation of /s/ before non-coronal voiceless stops. She gives additional word-juncture rules for /as/ > /o/ and /s/ deletion (for final /s/), but these cases could be treated as prosodically unconditioned developments of the /z/ (otherwise not found in Sanskrit) expected from Final Voicing; it behaves analogously to the glides, which are licensed only after /a/. She also treats r-Deletion as a word-juncture rule (for final /s/ or /r/ before /r/), but two /r/ never come together in Sanskrit, so this too could be an utterance-span rule. In the approach advocated in this paper, there are two word-level lexical rules-finals in /n/ add /s/, and final /r/ devoices. The rest of 'sandhi' is seamlessly integrated into

several dozen prosodically unconstrained postlexical rules.

-ă	-ĭ	-ŭ	-ṛ	-e (-ay)	-ai (-āy)	-o (-av)	-au (-āv)	Init
ā	ya	va	ra	e	ā a	o	āva	a-
ā	yā	vā	rā	a ā	ā ā	a ā	āvā	ā-
e	ī	vi	ri	a i	ā i	a i	āvi	i-
e	ī	vī	rī	a ī	ā ī	a ī	āvī	ī-
o	yu	ū	ru	a u	ā u	a u	āvu	u-
o	yū	ū	rū	a ū	ā ū	a ū	āvū	ū-
ar	yṛ	vṛ	ṝ	a ṛ	ā ṛ	a ṛ	āvṛ	ṛ-
ai	ye	ve	re	a e	ā e	a e	āve	e- (ay-)
ai	yai	vai	rai	a ai	ā ai	a ai	āvau	ai- (āy-)
au	yo	vo	ro	a o	ā o	a o	āvo	o- (av-)
au	yau	vau	rau	a au	ā au	a au	āvau	au- (āv-)

Chart 1: Sandhi of final vowels. Parentheses show underlying forms, boldface shows outcomes different from those seen word-internally. /e/ and /o/ are always long. /v/ is best considered a glide (a labiodental approximant). Short /a/ is -low on the surface, but throughout the phonology behaves as +low.

Final vowels plus initial consonants undergo no change, except that many manuscripts alternate /ch/ after long vowels and /cch/ after short ones; word-internally, many of the same MSS have /cch/ after all vowels.

-k	-ṅ (-ṅṅ)	-ṭ	-t	-n (-ns)	-p	-m	-ḥ (-s/-r̥̊)	-aḥ (-as)	-āḥ (-ās)	Init
ga	ṅa **ṅṅ**a[1]	**ḍ**a	da	na nna[1]	ba	ma	ra	o	ā a	a-
gV	ṅV **ṅṅ**V[1]	**ḍ**V	dV	nV nnV[1]	bV	mV	rV	a V	ā V	V-

kk*h*	ṅk*h*	ṭk*h*	tk*h*	nk*h*	pk*h*	ṅk*h*	**ḥ**k*h* xk*h*[2]	aḥk*h* axk*h*[2]	āḥk*h* āxk*h*[2]	k*h*–
gg*h*	ṅg*h*	ḍg*h*	dg*h*	ng*h*	bg*h*	ṅg*h*	rg*h*	og*h*	āg*h*	g*h*–
kc*h*	ṅc*h*	ṭc*h*	cc*h*	**ṃ**śc*h*	pch	ñc*h*	śc*h*	aśc*h*	āśc*h*	c*h*–
gj*h*	ṅj*h*	ḍj*h*	jj*h*	ñj*h* nj*h*	bj*h*	ñj*h*	rj*h*	oj*h*	āj*h*	j*h*–
kṭ*h*	ṅṭ*h*	ṭṭ*h*	ṭṭ*h*	**ṃ**ṣṭ*h*	pṭ*h*	ṇṭ*h*	ṣṭ*h*	aṣṭ*h*	āṣṭ*h*	ṭ*h*–
gḍ*h*	ṅḍ*h*	ḍḍ*h*	ḍḍ*h*	ṇḍ*h*	bḍ*h*	ṇḍ*h*	rḍ*h*	oḍ*h*	āḍ*h*	ḍ*h*–
kt*h*	ṅt*h*	ṭt*h*	tt*h*	**ṃ**st*h*	pt*h*	nt*h*	st*h*[2]	ast*h*[2]	āst*h*[2]	t*h*–
gd*h*	ṅd*h*	ḍd*h*	dd*h*	nd*h*	bd*h*	nd*h*	rd*h*	od*h*	ād*h*	d*h*–
ṅn gn	ṅn	**ṇ**n **ḍ**n	nn dn	nn	mn bn	nn	r n	oÌn	ān	n–
kp*h*	ṅp*h*	ṭp*h*	tp*h*	np*h*	pp*h*	mp*h*	**ḥ**p*h* **ɸ**p*h*[2]	aḥp*h* aɸp*h*[2]	āḥp*h* āɸp*h*[2]	p*h*–
gb*h*	ṅb*h*	ḍb*h*	db*h*	nb*h*	bb*h*	mb*h*	rb*h*	ob*h*	āb*h*	b*h*–
ṅm gm	ṅm	**ṇ**m **ḍ**m	nm dm	nm	mm bm	mm	r m	om	ām	m–
gy	ṅy	**ḍ**y	dy	ny	by	**ṃ**y	ry	oy	āy	y–
gr	ṅr	**ḍ**r	dr	nr	br	**ṃ**r	(V̄)r	or	ār	r–
gl	ṅl	**ḍ**l	ll	lÕl nl	bl	**ṃ**l Ĩl	rl	ol	āl	l–
gv	ṅv	**ḍ**v	dv	nv	bv	**ṃ**v	rv	ov	āv	v–
kś	ṅś ṅkś	ṭś	cch ch cś	ñch ñś	pś	ṃś	ḥś[3] śś	aḥś[3] aśś	āḥś[3] āśś	ś–
kṣ	ṅṣ ṅkṣ	ṭṣ	tṣ	nṣ nt**ṣ**	pṣ	ṃṣ	ḥṣ[3] ṣṣ	aḥṣ[3] aṣṣ	āḥṣ[3] āṣṣ	ṣ–

ks	ṅs ṅks	ṭs ṭts	ts	ns nts	ps	ṃs	**ḥ**s[3] ss	a**ḥ**s[3] ass	ā**ḥ**s[3] āss	s–
ggh g h	ṅh	**ḍḍ**h **ḍ** h	ddh d h	nh	bbh b h	**ṃ**h[4]	rh	oh	āh	h–

Chart 2: Sandhi of final consonants. Multiple entries usually show options in free variation. The rare final /l/ neither suffers nor causes changes. /ḥ/ and /ṃ/ are disarticulated continuant and nasal, other inferior dots denote retroflexes. /ṅ/ is a velar nasal. /r̥̊/ is a voiceless retroflex tap. /h/ is breathy-voiced as an independent segment; after a stop it marks aspiration. /ś/ and /ñ/ are palatal; /ĩ/ is nasal. [1]The nasal doubles only after short vowels. [2]The fricative is optionally /-ḥ/ if the initial stop is followed by a sibilant. [3]/ḥ/ is omissible if the sibilant is followed by an oral stop. [4]The /-m/ may assimilate to a (sonorant) consonant following the /h-/.

References

Allen, W. Sidney. 1972. Sandhi: the theoretical, phonetic, and historical bases of word-junction in Sanskrit. (Janua linguarum, ser. min., 17). 2nd printing of 1962 ed. ' s-Gravenhage: Mouton.

Bright, James W. 1957. Bright's Anglo-Saxon reader, rev. and enlarged by James K. Hulbert. New York: Henry Holt.

Cairns, Charles E. & Mark H. Feinstein. 1982. Markedness and the theory of syllable structure. LI 13.193-225.

Cho, Young-Mee Yu. 1990. Parameters of consonantal assimilation. Doctoral dissertation, Stanford.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. Papers in laboratory phonology, John Kingston and Mary E. Beckman, eds., 283-333. Cambridge:Cambridge Univ. Press.

Kaisse, Ellen M. 1985. Connected speech: the interaction of syntax and phonology. Orlando: Academic Press.

Kiparsky, Paul. 1979. Metrical structure assignment is cyclic. LI 10.421-441.

Monier-Williams, Monier. 1899. A Sanskrit-English dictionary. Oxford: Oxford Univ. Press.

Osthoff, H. 1882. Zur Reduplicationslehre. Beiträge zur Geschichte der deutschen Sprache und Literatur. 8.540-567.

Rice, Keren D. 1990. Predicting rule domains in the phrasal phonology. The phonology-syntax connection, ed. by Sharon Inkelas and Draga Zec, 289-312. Chicago: Univ. of Chicago Press.

Rice, Keren D. 1992. On deriving sonority: a structural account of sonority relationships. Phonology 9.61-99.

Selkirk, Elisabeth O. 1980. Prosodic domains in phonology: Sanskrit revisited. Juncture, M. Aronoff, 107-129. Saratoga, Calif.: Anma Libri.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral dissertation, MIT.

Whitney, William Dwight. 1964. Sanskrit grammar. Reprint of 2nd ed., 1889. Cambridge, Mass.: Harvard Univ. Press.

Footnotes

[1]By 'Classical Sanskrit' is here meant the language that flourished in the first millenium of the common era, as distinguished from the language of the earlier Vedic literature. It is essentially the language described by Pāṇini, although I will here ignore variant forms unused in the later literature, as well as the pitch accent, which was lost.

[2]For ease of comparison with almost all previous studies, I use the standard Indological transliteration system for Sanskrit. The symbols are explained in connection with the sandhi charts at the end of this paper.

[3]In Sanskrit studies, one uses the word *sandhi* to refer to all phonological changes at morpheme boundaries, then distinguishes INTERNAL SANDHI—within-word changes—from EXTERNAL SANDHI—between-word changes. The latter is of course what we usually mean

nowadays by *sandhi* simpliciter. I accept the existence of certain internal sandhi rules, but will argue that the so-called rules of external sandhi really have nothing to do with word juncture. Nevertheless, in order to avoid too great a break with established terminology, I shall use the term *external sandhi* loosely, to refer to those changes that are conventionally described as being conditioned by word juncture.

[4] Cho (1990:202-203) has argued that onset clusters must not be homorganic (see also Rice, 1992, for a claim that this restriction is nearly universal). In particular Cho would disagree with my position that /dhn/ or /pm/ could form an onset in words like *budhnaḥ* 'bottom' or *pāpmā* 'evil' , and her case is bolstered by the fact that such sequences do not occur at the beginning of words. By making these into absolute and pervasive restrictions on onset templates, however, Cho is forced to treat many neutralizations as being disjunctively conditioned by word margins and by obstruent clustering, since, for example, /dh/ loses its laryngeal features at the end of a word and before obstruents, but not internally before nasals. I would rather say that these heterorganicity requirements are simply morpheme structure constraints, or word margin constraints, or at most onset structure constraints that apply only in the base lexicon and are relaxed at later levels.

[5] This could be taken for evidence that the /v/ is an obstruent at the time of lexical syllabification (Cho 1990:201-204), which could be supported by the fact that this /v/ has a labiodental pronunciation. But these onsets date back to Proto-Indo-European, where the labial is clearly a glide (note the comparatively recent survival of /wr-/ onsets in English). It would appear rather that it is the labiality of the /w/ that somehow makes a difference. It is noteworthy, for example, that Sanskrit morphemes can begin with /mr/ (*mriyate* 'dies'), /ml/ (*mlecchaḥ* 'foreigner'), and /mn/ (the root *mnā* 'note'), but not */nr/, */nl/, */nm/.

[6] In their system, a branching onset like /try/ is divided

into a margin /t/ and an adjunct /ry/. In cases like those under consideration here, the margin can itself branch into a premargin and a margin core, so that the onset /stry/ would be bracketed [[[s][t]][ry]]. I prescind from whether quite that much structure is needed, while agreeing that the premargin definitely has some special status.

[7]Whitney (1964:315). It must be noted however that this /s/ presumably also disappears after short vowels and before /dh-/ endings, which is not analogous to how /s/-final stems behave, so the possibility of morphological idiosyncrasy must not be ruled out.

[8]Bartholomae's Law provides that when a [+voiced +spread] Laryngeal node delinks from a coda, it will dock on an immediately following dental stop (Whitney 1964:55). The scope of this law is strictly lexical, so the deletion of the /s/ cannot simply be a postlexical development.

[9]Since these are blind enumerations of clusters, it should be kept in mind that they constitute a mixture of underlying clusters and clusters derived from morphological processes. Some of these clusters will also result from compounding and prefixing, which with respect to syllabification and sandhi behave like separate words. The point of this section is not to precisely enumerate word-internal clusters, but to show that there are no heterosyllabic clusters that resist the rules of external sandhi.

[10]This adding of an /s/ is obviously not a universal tendency, but is peculiar to the history of Sanskrit. In pre-Classical (Vedic) times, the ending /-n/ was rare but /-ns/ very common. The /s/ of the latter disappeared in many sandhi environments, yielding results identical to that for /-n/. This situation was simplified in Classical Sanskrit so that there was no distinction between the terminations /-n/ and /-ns/. The controversy as to whether the /s/ is underlying or added by sandhi (see Cho 1990:71, who sides with Schein and Steriade in considering it underlying against Odden, who says it is added in sandhi) may

partly reflect some unclarity about whether the discussants are referring to the Vedic or the Classical language.

[11]Retroflection spreads to a following dental stop.

[12]Although /s/ can apparently occur in codas at some level (its sandhi treatment at the end of the word strongly suggests that it emerges from the lexicon as /s/, and the treatment of /s/ at the end of noun stems is totally analogous), verb stems consistently reject it from codas. Perhaps verb inflection happens in an earlier, more restrictive, stratum than that of nouns.

[13]The alternation of the initial /d/~/dh/ is interesting but distracts from the matter at hand. The general rule appears to be that when aspiration is delinked from a voiced consonant at the end of a root, it can dock on a voiced consonant at the beginning of the root (Whitney 1964:53). Bartholomae's Law, however, which transfers the laryngeal features to a following dental stop, takes precedence (p. 55).