Ambisyllabicity in the language of the Rigveda

Brett Kessler

February 11, 1998

1 Introduction.

Vedic Sanskrit provides strong evidence for the phenomenon of ambisyllabicity, the idea that the consonantal interlude between two vowels can belong simultaneously to both of the adjacent syllables.[1] Despite several studies purporting to demonstrate ambisyllabicity in modern Germanic languages, including Kahn's classical description of consonantal allophony in English (1976), the existence of ambisyllabicity has been seriously questioned. In this paper I shall support the existence of ambisyllabicity, arguing that in the language of the Rigveda, the oldest extensive literature in Sanskrit, the first consonant of an intervocalic onset cluster is ambisyllabic, at least after a short vowel. Since accent is not a conditioning factor, this is a rather different environment from the stress-sensitive phenomena described in Germanic, and so the data are not susceptible to Kiparsky's (1979) foot-based reanalysis. It will also be seen that the Sanskrit data are not easily tractable by Selkirk's (1982) resyllabification-based alternative to ambisyllabicity, either.

Among those who support ambisyllabicity, the details of its representation are a matter of some debate. Borowsky *et al.* (1984) conjectured that there is no distinction between geminate and ambisyllabic consonants: it is a matter of a language's phonetic implementation to determine whether consonants shared by adjacent syllables will be pronounced as short (ambisyllabically, as in English) or as long (as geminates, as in Italian). On the other hand, Kang (1991) argued that the two phenomena contrast in Korean, and so the two must have different structures. It turns out that Vedic Sanskrit has both ambisyllabicity and long consonants, but they are in complementary distribution. This argues for Borowsky *et al.*'s conjecture, although with a more fine-grained phonetic implementation than a single language-wide setting.

2 Background.

2.1 The reality of ambisyllabicity.

I am unaware of just how far back the concept of ambisyllabicity goes, but it has been a recurring idea ever since phoneticians discovered in this century that there is no reliable phonetic correlate to the idea of syllable boundary. Hermann (1923:3) was willing to entertain the possibility that syllable boundaries could occur inside certain segments rather than between them, and Kuryłowicz (1948) developed an elaborate theory of syllable structure that relied heavily on ambisyllabicity. In his view (simplifying somewhat), intervocalic clusters of two or more consonants break down into a coda (sounds admissible at the end of a word, or as part of a geminate cluster) and an onset (sounds admissible at the beginning of a word), with all but the last consonant having the possibility of ambisyllabic overlap if both conditions obtained. Thus in Latin *monstrum*, the *str* would be in the onset of the second syllable, since words can begin with *str*, and the *ns* would be in the coda of the first syllable, since words can end in *ns*. The *s* overlaps ambisyllabically.

Linguists of the pregenerative period however rarely adduced much internal evidence for their claims about syllable structure, and the evidence that was given centred almost entirely on syllable weight as determined by the metre. The situation changed in the 1970s, as phonologists began discovering many kinds of phenomena that were syllabically conditioned. For ambisyllabicity, a major breakthrough came with Kahn (1976), who showed how such allophonic processes as aspiration, glottalization, and tapping (reduction of alveolar stops to [ɾ]) of stops in American English can be felicitously described in terms of syllable structure. According to Kahn, English basically syllabifies in accordance with the maximum onset principle, but a few subsequent adjustments can make a single consonant be shared between two syllables. One adjustment is that an open syllable will get a coda by sharing the first consonant of a following unstressed syllable, as the /t/ in *city*. Another is that a word beginning with a vowel will get an onset by sharing the last consonant of the preceding word, so that the /t/ in *let Ann* is ambisyllabic as well. Gussenhoven (1986) has emended and extended Kahn's analysis to account for perhaps a dozen allophonies in American and British English. A compelling feature of this sort of analysis is that ambisyllabic segments are not only subject to rules of their own, such as tapping, which in American English applies only to ambisyllabic alveolar stops, but also may participate in rules that apply to onsets and also those that apply to codas. For example, in *alley*, the first vowel is an /æ/, which occurs only in stressed, closed syllables, which places the /l/ in the coda; but the lateral is clear (non-velarized), which is the allophone taken by /l/ in onsets.

Convincing cases have also been made for other modern Germanic languages. Vennemann (1982:280–281) argued that a German word like *Roggen* /ʀɔgən/ 'rye' must have an ambisyllabic /g/, since /ɔ/ otherwise appears only in closed syllables, but all stops are voiceless when not in an onset. Borowsky *et al.* (1984) stated the case for gemination in Danish, where stressed syllables must be heavy, and so a syllable ending with a stressed short vowel will ambisyllabically share the initial consonant of the next syllable. This consonant is now eligible to bear a stød (glottal stop) or make a preceding /a/ grave [ɑ], which properties are otherwise features of coda consonants. Hulst (1985) argued that all full (non-schwa) vowels in Netherlandic must be in heavy syllables. If a full short vowel appears at the end of a syllable, a following single consonant ambisyllabically spreads to fill the preceding coda.

Several attempts have been made to account for the same phenomena without ambisyllabicity. Kiparsky (1979:432–440) proferred a foot-based approach, attempting to show that the consonantal allophony of English can be determined by first constructing metrical feet beginning at the left edge of the word and at all stressed syllables, then weakening consonants that follow a [-cons] and are not foot-initial. Selkirk (1982:360–379) more directly addressed the paradoxical status of segments that seem to have both onset-like and coda-like properties when she provided that consonants may move from an onset into the preceding coda during the course of a derivation, picking up properties conditioned by the respective syllable component or conditioning the appropriate phonological changes as they move about. Gussenhoven (1986) demonstrated that both of these solutions were inadequate on empirical grounds. But it is worthwhile to ask why people object to ambisyllabicity. A major objection is that it violates proper bracketing, or specifically, the prosodic hierarchy, which teaches that elements at one prosodic level are properly included in a single parent construct (Selkirk 1982:355). To this one may well ask why segments should be considered part of the prosodic hierarchy; or why geminates, which autosegmental phonologists agree are typically single melodic constituents shared by two syllables, are not an equally big problem; or on exactly what grounds improper bracketing should be excluded from linguistics (Lass 1985:249). A somewhat more troublesome challenge is Picard's objection (1984) that ambisyllabicity has no phonetic grounding, and so amounts to nothing more than an unverifiable formalism the phonologist can make use of. Here the appropriate safeguard would be to make sure that prospective ambisyllabic segments actually have properties both of onsets and of codas. Some phonologists have inverted the requirement, finding ambisyllabicity where segments have properties devolving *neither* from their status as onsets *nor* from their status as codas. For example, Suzuki (1985:102-104) claims to have proved that Old English had ambisyllabicity, because /h/ disappeared between sonorants but not next to obstruents or word-marginally. Certainly, /h/ could have been ambisyllabic between sonorants and as such susceptible to a special rule (like Kahn's tapping of ambisyllabic /t/ in American English), but arguing that a segment does something that is not a necessary result of being in either an onset or a coda does not seem like a very convincing proof that it must in fact be in both. In the same way, Kang's argument for ambisyllabicity in Korean (1991:90-94) hinges on the assertion that intervocalic [r] from underlying /l/ (/tali/ → [tari] 'leg') must be ambisyllabic, because not all onset /l/ or all coda /l/ change to [r] (contrast /kamlam/ → [kamnam] 'olive', /məlli/ → [məlli] 'far'). Or that [β] must be ambisyllabic because it only appears between vowels ([tɨtsaβa] 'hear (humble completive)'), not in other codas or onsets. Not every change has to be conditioned exclusively by syllable structure.

Admittedly, there are a few other objections to ambisyllabicity that are harder to dismiss. Picard also complains (1984:45–46) that different phonologists have significantly different criteria for ambisyllabicity in even a single language such as English. Kahn (1976), as has been seen, provides for ambisyllabicity at the left edge of unstressed syllables; Hooper (1978:189) provides for it at the right edge of stressed syllables; Anderson and Jones (1974) make stress irrelevant. There is also the sticky problem that not even Gussenhoven's recent analysis (1986) fully accounts for all the allophony data, which, to be fair, are certainly complicated by a good deal of variability, even for a single speaker. Finally, it is suspicious that most findings of ambisyllabicity have been restricted to the modern Germanic languages. Arguably, these cases all share in common the loss of geminates in a system where stressed syllables are necessarily bimoric, and all are languages that areally share many features. The case of ambisyllabicity would be bolstered by showing that it is present in a more distantly related language, with no contact to English, and under a different set of historical circumstances. Fortunately Sanskrit is an answer to this last objection.

2.2 The representation of ambisyllabicity.

Ever since it became recognized that long (geminate) consonants are best thought of as a single consonant shared by two adjacent syllables, the question arose of how this differs from ambisyllabicity. Perhaps the majority view has been that it does not differ at all. Borowsky *et al.* (1984) suggested that languages have either true geminates, or ambisyllabicity, or neither, but not both. Therefore the same representation can be used for both phenomena; whether the shared consonant is pronounced long or short is a matter of phonetic implementation. Borowsky *et al.* provided positive evidence that ambisyllabic consonants in Danish behave like geminates, since they are immune to lenition in a word like *kappa* (/kapa/, not */kaba/); immunity to coda lenitions is a well-known property of geminates. Hulst (1985:61) pointed out that Dutch ambisyllabic consonants are immune to devoicing, another property credibly attributable to geminate inalterability. On the other hand, people who believe that there is contrasting ambisyllabicity and gemination in the same language need to draw a structural distinction. Suzuki (1985:106) criticizes Lass for having to add an extra CV tier so that ambisyllabicity (CC aligns with C) can be distinguished from gemination (CC aligns with CC), but it is not clear what Suzuki's alternative is. Kang (1991:92), working within Moraic Theory, has both types of consonants linked to two moras, including the first mora of the posterior syllable; but whereas ambisyllabic consonants are linked to the only mora of the prior syllable, geminate consonants are linked to the second mora of the prior syllable, a mora which is assigned to the geminate consonant lexically. But while such proposals are certainly worth entertaining, I am not convinced of the presence of ambisyllabicity in either Old English or Korean, for the reasons stated above in section 2, so the economical possibility that gemination and ambisyllabicity have the same representation needs to remain open.

2.3 Prior proposals for Sanskrit syllabification.

Many conflicting ideas have been put forward concerning the syllabification of Vedic Sanskrit. Varma (1929:61–86) presents the statements of the Prātiśākhyas, ancient Indian treatises devoted to the correct pronunciation of the various Vedas. All agree that consonants at the edge of words go with their adjacent vowels, and that a single consonant between vowels goes with the following syllable. But they disagree on how to treat clusters of consonants. Most of the Prātiśākhyas state that the consonants split up (Allen 1953:82 cites Atharvaprātiśākhya 1.55–57 and Vājasaneyiprātiśākhya 1.99–106). But the R̥kprātiśākhya, the Prātiśākhya that was written for the R̥gveda, states that it is also acceptable for both consonants to go with the following syllable (1.15, apud Varma 1929:62). Finally, the Taittirīyaprātiśākhya draws a distinction: most consonants are split into two syllables, but they are kept together if the second is an approximant (21.7, apud Varma, p. 75), or if the sequence is a stop followed by a fricative (21.0, apud Varma, p. 71).[2] Table 1 gives some examples of how the various schools would syllabify some of the words found in the first hymn of the R̥gveda.[3] It should be kept in mind that the discussions in those treatises are very cursory, and that the authors may not have intended these precise outcomes. The statement of the R̥kprātiśākhya that either syllabification is permitted may originally have meant that the alternation is determined in part by some unspecified property of the word, not necessarily entirely by the whim of the speaker. More importantly, the authors of the Prātiśākhyas are only concerned with fixing the pronunciation of a liturgical text composed centuries before their time. They did not put their definition of the syllable to any phonological or grammatical use. If one is inclined to believe that syllables are ultimately phonological or that there can be differences between phonetic and phonemic syllables, then the statements of these phonetic treatises should be approached with caution.

************* Insert Table 1 about here **********

Thus already in ancient India there were two fundamentally different approaches to syllabification of intervocalic clusters. The first I shall refer to as the cluster-splitting approach (CS). It held that all clusters split up, so that the first consonant is in the prior syllable, and the last consonant is in the succeeding syllable. The second approach, that of the Taittirīyaprātiśākhya, may be referred to as onset-maximization (OM), which decrees that as many consonants syllabify with the second vowel as is compatible with general constraints on the structure of the onset. Different linguists have formulated different onset structure constraints. According to the theory of the Taittirīyaprātiśākhya, those constraints are that a consonant may be followed by an approximant, or that a stop may be followed by a fricative.

The same two classes of approaches have continued into more recent times. Most Indo-Europeanists subscribe to the CS theory, that all clusters split in Vedic Sanskrit (Hermann 1923:261, Kuryłowicz 1948:193, Vaux 1992). In recent years, theoretical linguists have published several proposals taking an OM approach, that clusters syllabify together in an onset if they can (Steriade 1982, Rice 1990:309-311, Cho 1990:204; Kessler 1993 argued this for Classical Sanskrit, a later variety of the language). They differ among themselves only in certain small details concerning what the possible onset clusters are. Both sides of the debate have strong evidence on their side. In the following sections I shall briefly present the evidence that I am aware of for each analysis, and then show that ambisyllabicity would account for both sets of data.

3 Data from the Rigveda.

3.1 Cluster splitting.

3.1.1 Metre.

By far the most important evidence is that of the Vedic poetic metre. Vedic verse works on the same principles as that of Ancient Greek: it is based on alternations of heavy and light syllables. A syllable with a long vowel or a diphthong is always heavy, but a syllable with a short vowel may be either light or heavy. The distinction in the latter case depends on how many consonants come between that short vowel and the next vowel, ignoring word boundaries: if at most one, then it is light; if two or more, then it is heavy. For example, in the gāyatrī metre, the last three syllables but one are regularly light, heavy, light. The first line in RV 1.1.5 is *agnír hótā kavíkratuḥ* 'Agni the wise invoker', so the second syllable of the last word is heavy. The obvious conclusion is that the first of multiple intervocalic consonants is in the coda of the syllable containing the prior vowel, and that a single intervocalic consonant is never in the coda (Hermann 1923:257).

3.1.2 Sievers's Law.

The authors of the Prātiśākhyas already noticed that the received text of the R̥gveda scans very badly unless many instances of the glides *y* and *v* are treated as syllabic. Thus in RV 2.6.7, the line *dūtó jányeva mítryaḥ* has only seven syllables even though the verse is written in the eight-syllable gāyatrī metre. This can be rectified if one reads the last word, *mítryaḥ* 'befriended', as *mítriaḥ*. Sievers's Law (Collinge 1985:159–174, Seebold 1972) gives one condition for this alternation: the vocalic reading is more likely if the preceding syllable is heavy. This includes the case where that syllable has a short vowel followed by two consonants, such as *tr* in this example. Again, this suggests that the first of two consonants closes a syllable, as predicted by CS (Hermann 1923:257-260).

3.1.3 Perfect union vowel.

In the perfect stem of the verb, inflectional endings that begin with a consonant are often preceded by the vowel *i*. If the stem ends in a consonant, the rule for the language of the Ṛgveda is that the vowel is added only if the root has a long vowel or ends in any two consonants (Whitney 1889:287). For example, before the suffix *-ma*: *ūcimá* 'spoke' (RV 1.161.1), *paptima* 'flew' (RV 9.107.20), but *jaganma* 'went' (RV 4.16.18). The principle appears to be that the *i* is syncopated after a light syllable. This explanation entails that the first of any two consonants closes the syllable and makes it heavy, as the CS theory predicts.

3.1.4 Reduplicated aorist.

One way of forming the aorist of a verb root involves reduplication using the template C$\breve{\bar{\text{V}}}$-, usually preceded by a temporal augment *a-* (Whitney 1889:309–310). For the vast majority of verbs, the reduplicated aorist stem ends in a short vowel plus a single consonant. Since these aorists almost always add an *a* suffix (*vaśa-*, *syada-*), this makes the stem light, by either the CS or the OM theory. For these light stems (but not for the rarer heavy stems), the vowel of the prefix is short when the stem begins with two consonants (*ádidyutat* 'shone', RV 6.11.4), and long when it is followed by only one consonant (*anīnaśat* 'drove away', RV 10.162.2). This makes sense if the actual requirement is for the prefix to be heavy, in which case a basic short vowel is lengthened only if the syllable is not already heavy because it is closed by the first of two consonants. The CS theory is validated in that any two consonants will do: all versions of the OM theory, including that of the Taittirīyaprātiśākhya, would syllabify *á-di-dyutat*, which would undesirably leave the prefix light (Vaux 1992:295).

3.1.5 Reduplicated intensive.

One way of forming an intensive or frequentative stem in Sanskrit is to reduplicate the stem with the template CaC$\breve{\bar{\text{ı}}}$-. Short *i* is used before any two consonants (*kanikradat* 'roared', RV 5.83.9, from the root *krand*), long *ī* before a single consonant (*ganīganti* 'approach', RV 6.75.3) (Whitney 1889:364). Again the requirement must be for that syllable to be heavy, whether by being in a closed syllable or by having a long vowel. Since the first of any two consonant pairs is sufficient to close the syllable (giving *ka-nik-ra-dat* rather than the *ka-ni-kra-dat* that any OM theory would propose), the CS theory wins.

3.1.6 Summary of the data for CS syllabification.

The language of the Ṛgveda shows many circumstances where description is simplified if one can invoke the distinction between heavy and light syllables, where a heavy syllable contains either a long vowel or a coda consonant. In all such cases I am aware of, any sequence of two consonants parses in such a way that the first one closes the prior syllable, confirming the CS hypothesis over against the OM hypothesis, which says that at least some combinations of consonants go together to form an onset for the following syllable. On the other hand, a single intervocalic consonant is not enough to make the preceding syllable heavy, so it must regularly syllabify with the following syllable.

There are however some weaknesses in this demonstration. The first is that the data are probative only for the first consonant that appears after short vowels. If the vowel is long, the syllable is going to be heavy, regardless of how the following consonant or cluster syllabifies. Furthermore, a syllable closed by one consonant is indistinguishable in weight from one closed by more than one consonant: a cluster of two intervocalic consonants could conceivably both go with the prior syllable. This does not seem likely, however, since a single intervocalic does clearly go with the posterior vowel: one may generalize that the last segment of a cluster always goes with the posterior vowel.

The second weakness is that only the metre is an exceptionless phenomenon that is unambiguously phonological in the Vedic language. The other phenomena all are lexical phenomena which one could argue are inherited from an earlier stage of the language where perhaps other syllabification rules applied. Occasional exceptions highlight this possibility. Thus one finds the perfect *véttha* 'know' (RV 10.15.13) despite the rule that other heavy perfect stems ending in a consonant add *i* before a consonantal ending. And one finds *davidhāva* 'we shook' (RV 1.140.6) despite the rule that other intensives have long *ī* before single consonants (*dh* is a symbol for a single consonant, an aspirated stop). And Sievers's Law in the Ṛgveda is riddled with exceptions, morphological conditioning, and plain optionality: Hock (1980) considers it to be a set of variable rules in the Labovian sense (but see Kiparsky 1972 for a bold claim that it is actually exceptionless), and certainly many words can appear either with glides or with vowels, apparently ad libitem (or metri causa). Nevertheless the other rules are close to exceptionless, and at any rate I know of no situations where an alternation sensitive to syllable weight systematically fails to treat any set of two-consonant clusters as if they fail to make position, even universally favoured onset clusters such as *tr*.

3.2 Onset maximization.

There are equally good arguments that Vedic Sanskrit maximizes onset clusters.

3.2.1 Word margins.

Perhaps the first documented approach to defining syllable boundaries was to look at the beginnings and ends of words. According to Allen (1953:82), Herodian taught that clusters that begin words also begin syllables within words. Thus he not only was in favour of Onset Maximization, but he also provided a useful test as to what might be found in a maximized onset. Similar approaches have been used more recently (Anderson and Jones 1974:4, Kahn 1976:39–40), although some have cautioned that word margins may have extra restrictions that do not apply to word-internal syllables, or that they conversely may contain appendices admitting segments not licensed by word-internal syllables (Booij 1983:251–252, Hayes 1980:148). Fortunately Sanskrit turns out not to have appendices (Kuryłowicz 1948:214), and phonological word margins have almost no additional restrictions over syllable margins. (The exception is that *r* is devoiced at the end of a word, though not within compound words.)

Word beginnings. Looking through the text of the R̥gveda, it appears that words can begin with any basic consonant with the exception of some of the retroflex or aspirated stops: *k*, *kh*, *g*, *gh*, *c*, *ch*, *j*, *ṭ*, *ṇ*, *t*, *d*, *dh*, *n*, *p*, *ph*, *b*, *bh*, *m*, *y*, *r*, *l*, *v*, *ś*, *ṣ*, *s*, *h*. (Underlyingly, words do not begin with retroflex stops or *ṇ*, and very few begin with *ṣ*; these appear on the surface courtesy of retroflection rules, which can apply across words within a phonological phrase.) More importantly, there are many clusters. It appears that in general an obstruent or *h* can be followed by a sonorant (I find *kr*, *kl*, *kv*, *khy*, *gn*, *gm*, *gr*, *gl*, *ghn*, *ghr*, *chn*, *chm*, *chy*, *chr*, *chl*, *chv*, *jñ*, *jm*, *jy*, *jr*, *tm*, *ty*, *tr*, *try*, *tv*, *dy*, *dr*, *drv*, *dv*, *dhm*, *dhy*, *dhr*, *dhv*, *ny*, *nv*, *py*, *pr*, *pl*, *br*, *bhr*, *my*, *mr*, *ml*, *vr*, *vy*, *śn*, *śm*, *śy*, *śr*, *śl*, *śv*, *ṣn*, *ṣm*, *sn*, *sm*, *sy*, *sr*, *sv*, *hn*, *hr*, *hl*, *hv*, *hy*). (The sequences with *ch-* are sandhi forms of underlying *ś-* after a dental stop or nasal.) A voiceless stop can be preceded by a sibilant, which must be homorganic in the case of coronals (*śc*, *ṣk*, *ṣṭ*, *ṣṭh*, *ṣph*, *sk*, *st*, *str*, *sth*, *sp*, *sph*). A voiceless unaspirated stop may also be followed by a sibilant (*kṣ*, *kṣm*, *kṣv*, *ts*, *ps*). The inventory of consonants that can follow a sonorant is more idiosyncratic: *ny*, *nv*, *my*, *mr*, *ml*, *vr*, *vy*. But clearly only sonorants may follow a sonorant, and only labial sonorants can be followed by a liquid.

There are some gaps in the generalized pattern. The sound *l* never follows a dental, which is a quite common restriction cross-linguistically (Clements 1990:316). The near absence of palatals and of voiceless aspirates in clusters can be referred to their fairly recent origin, where they developed from other sounds, primarily before vocoids. Dental stops are not found before *n* and labials are not found before *v*. Certainly homorganic clusters are universally discouraged in onsets, but these may just be accidental gaps, since Vedic does have homorganic *jñ*, *sn*, and *jy*.

The presence of word-initial clusters makes for a certain presumption against the CS theory. All things being equal, one would expect word-internal onsets and word-initial onsets to follow the same rules; strict CS is normally associated with a language lacking word-initial clusters. To harmonize these clusters with CS, one would have to assume a large set of word-initial appendices, even though word-initial appendices are not very common cross-linguistically, and these clusters look like typologically quite normal onset sequences. Or one might hypothesize a very strong preference for building codas, even though onsets are strongly favoured cross-linguistically.

Word endings. An inventory of word endings will not in itself supply an argument for OM syllabification, but it is useful preparation for one. Superficially, there seems to be little restriction on the inventory of consonants that can appear at the end of a word: I find *ḥ*, *ṃ*, *k*, *g*, *ṅ*, *c*, *j*, *ñ*, *ṭ*, *ḍ*, *ṇ*, *t*, *d*, *n*, *p*, *b*, *m*, *y*, *r*, *l*, *v*, *ś*, *ṣ*, *s*. The only sounds obviously missing are the aspirated stops and *h*. As for clusters, *r* is found preceding final *k*, *g*, *t*, *d*; and the disarticulated nasal (perhaps to be considered a nasalisation of the preceding vowel) precedes sibilants, *ḥ* and *r*. But Sanskrit is famous for its system of sandhi, where word-final sounds undergo neutralizations and assimilations, which are represented in the text. So for example any of the final *c*, *j*, *t*, *d*, *n*, *l* may actually come from any of final *t*, *th*, *d*, *dh*, depending on the initial segment of the next word.

As for the neutralizations, obstruents lose all inherent laryngeal features of voice and aspiration at the end of a word (e.g., *t*, *th*, *d*, *dh* all become *t*), palatals become either velar or retroflex (i.e., a [distributed] feature is lost: *c*, *ch*, *j* → *k* or *ṭ*), and fricatives other than *s* lose their [cont] feature, becoming stops (*ṣ*, *ś* → *k* or *ṭ*). It is important to note that this neutralization cannot be purely phonetic in nature: *j* becomes *k* in some words, *ṭ* in others. This has led Cho (1990:78–79) to argue that there are really two different *j*, a conclusion for which there is an historical basis but no other synchronic evidence: all *j* are otherwise indistinguishable. Furthermore *ś* and *h* also share this property, and following Cho's approach would require setting up three artificial phonemes to account for all the alternations phonologically. I would prefer to merely recognize that whereas the phonology constrains the set of segments that can appear word-finally, the lexeme itself has a role in determining exactly which segment is chosen. Thus neutralization cannot be left to some utterance-level postlexical process, but must take place at some level where the lexical information in stems is accessible.

It would appear that after the neutralizations, the last consonant of a word must be drawn from the set *k*, *ṭ*, *t*, *p*, *m*, *y*, *r*, *v*, *s*. As for clusters, *ns* is very frequent, *ṅs* tolerably frequent, and *rt* and *rk* rather rare (*ns* itself is in part a reduction of *nts*, and *ṅs* comes from underlying *nks*.) It should be remembered that this inventory is somewhat abstract, because the actual surface forms will be conditioned to a large extent by assimilation to the next segment, and in pausa further neutralizations occur. These are the subject of the next section.

Sandhi. The facts of sandhi are treated comprehensively in all of the grammars, such as Macdonell (1916:20–39) and Whitney (1889:34–81), so I shall merely summarize the facts in Table 2.

************ Insert Table 2 about here *****************

The most striking fact about consonant sandhi is that with very few exceptions, it is a word-final consonant that is affected, not a word-initial one. In contrast to all the changes listed in the table for the final consonant, the only thing that happens to onsets is an occasional timing phenomenon: a stop may be inserted between nasal and sibilant, or an *h* after a stop may become the homorganic (but aspirated) stop, or an *ś* after a *t* (including excrescent *t* after *n*) may become *ch*. These are all caused by a slight delay in releasing the oral closure after a stop, and were all considered optional changes by the Prātiśākhyas. As Cho (1990) notes, this broad pattern of neutralization and assimilation is well documented for the ends of words, and indeed for the ends of syllables in general.

3.2.2 Sandhi generalized.

So far I have been giving a fairly traditional account of sandhi, which describes word-final[4] changes as a special case, called external sandhi. This is necessary within the CS theory. If *t* voices to *d* word-finally when before *r*, but fails to do so within a word like *atra*, then if *atra* is syllabified *at-ra*, there is nothing else to attribute the difference to than to the end of the word. But there are several drawbacks to such a theory. The first is that it seems curious that there would be an assimilation that is enabled by distance (Allen 1962:25–26). One might expect a neutralization to become more likely at the end of higher prosodic categories, but for an assimilation one would expect the opposite: if it occurs across word boundaries, the unmarked case would be to expect it to occur within word boundaries, as well. Kaisse (1985) also found such rampant word-juncture effects problematic for her theory of sandhi, under which such effects should be lexically or syntactically conditioned. Secondly, many of the neutralizations discussed above, as well as the assimilatory changes called for in the external sandhi chart, do occur word-internally. This is particularly striking when the same process appears to have a disjunctive condition. Obstruents, for example, assimilate in voice to a following obstruent regardless of position (*mahat-* 'big' plus *-bhis* instr. pl. → *mahádbhiḥ*, RV 1.72.9). But before a sonorant, they do so only at the end of a word (as opposed to word-internal *mahatā́* instr. sg., RV 1.32.5). Similarly *m* assimilates fully to stops everywhere, but to glides and *l* only at word boundaries.

Whitney already pointed out (1889:38) that the major difference between rules of internal and external sandhi, if one abstracts away from the fact that in word juncture one must account for consonant combinations that accidentally do not occur in the morphology, is that sonorants voice preceding obstruents at word boundaries, but not word-internally. This is a strong hint that syllable structure is involved, since one expects obstruents to be able to form onset clusters with sonorants. Rice (1990:307–311) drew the correct conclusion that voicing assimilation is a matter of syllable structure. Obstruents assimilate in voice whenever they find themselves in a coda, and all consonants at the end of a word are necessarily in a coda. Within a word, however, consonants will syllabify with a following sonorant, ending up in an onset, which does not condition the assimilation. Kessler (1993) later demonstrated that for Classical Sanskrit this analysis is appropriate for virtually all sandhi processes that seem to be conditioned by word boundaries.[5]

So one clear advantage of the OM syllabification is that it allows rules of neutralization and assimilation to be stated more naturally. Under this theory, words syllabify according to a principle of maximum onset (Pulgram 1970:47–57, Kahn 1976:42–46, Clements 1990:299), except where morphology adds material to the left (prefixes or compounding stems). Whenever a consonant cannot go into an onset, either because the onset templates forbid a particular cluster or because syllabification does not occur across the pertinent domain (after left-appended material or at the end of words), then it is subject to phonological neutralizations that are in part lexically determined, then to feature-sharing with the next segment. This is virtually a textbook example of universal patterns of consonant assimilation (Cho 1990).

It might not be amiss to give one or two examples of this working word-internally. Table 3 gives the singular and plural of an adjective that ends in a consonant; nouns are essentially identical. It is difficult to find a word that takes all possible inflections with the text of the R̥gveda; the example used here is fleshed out with starred forms supplied by analogy (Macdonell 1916:53, Whitney 1889:146–147).

************* Insert Table 3 about here *****************

Here one sees that underlying *j* is preserved as such when it can be in an onset (before a vowel), but is neutralized to a velar and is subject to voicing assimilation when it is stranded in a coda. Verbs follow the same pattern. Grassmann (1875:1115–1118) lists (among others) the following forms for the root *yu(n(a))j-* 'yoke': *yunájmi*, *yuñjanti*, *yunájat*, *yujyāva, ayujmahi*, *yuyujre*, versus *yunakṣi*, *yunákta*, *yuṅkté*, *yuṅkṣvá*, *yuṅgdhvám*, *ayukthās*, *ayugdhvam*, *áyukṣātām*. These alternations are easily accounted for by saying that *j* can syllabify with a following vowel, *m*, *r*, or *y*; but before *s*, *t*, *th* and *dh* it must stay in the coda and undergo neutralization to *k* and subsequent assimilations. This agrees with the fact that stops are found word-initially before sonorants but not before obstruents.

A good deal of information about syllabification can be gleaned from sandhi behaviour itself, once the syllabic theory of sandhi is accepted. The most convincing evidence comes from those situations where adjacent segments give different results word-internally and between words, since this directly indicates that those segments can form an onset sequence, protecting the first element from coda-coditioned changes. I have already mentioned how word-final obstruents lose their own laryngeal features and link to the [voice] feature of the next segment, but do not assimilate to sonorants in the same word, e.g., *mahatā́* 'big' (RV 7.104.16), *satyáḥ* 'true' (RV 1.145.5), *śakráḥ* 'mighty' (RV 7.104.20), *asmi* 'am', RV 1.164.37. Similarly, stops fail to nasalize before word-internal nasals: *rátnam* 'gift' (RV 1.41.6). So it is clear that obstruents syllabify with following sonorants, exactly as seen in word-initial sequences.

Labial sonorants appear to have a special inclination to head onset clusters. Since *m* can occur in Sanskrit words before *n, y, r, l,* and *v*, before which it disarticulates and/or assimilates word-finally, it must form onset clusters with those sounds in e.g. *mrakṣakṛ̥́tvā* 'rubbing to pieces' (RV 8.61.10). Similarly, *v* must be syllabified with *ṇ* (*grā́vṇām* 'pressing stones', RV 10.85.4), because *āv* regularly become *au* before a consonant at the end of a word (not shown in Table 2). The segment *v* must also be syllabified with *y* (*suhavyám* 'fit for invocation', RV 1.74.5), since *av* becomes *o* word-finally before *y*. One might generalize that labial sonorants syllabify with (non-labial?) sonorants.

3.2.3 Word-internal clusters.

Thanks in part to its complex morphology, Sanskrit has many word-internal consonant clusters. If OM and the syllabic theory of sandhi are correct, then these should consist only of legal onsets, possibly preceded by legal codas which undergo sandhi changes. And this seems to be the case, if one assumes that a legal onset is one that conforms to the general patterns seen at the beginning of words or indicated by sandhi. By this criterion, I have identified the following word-internal onsets, using a machine-readable copy of the Ṛgveda (Ananthanarayana and Lehmann 1970). Obstruents or *h* followed by a sonorant: *kṇ*, *kn*, *kny*, *km*, *kmy*, *ky*, *kr*, *kry*, *kv*, *khy*, *khv*, *gṇ*, *gn*, *gny*, *gm*, *gmy*, *gy*, *gr*, *gry*, *gv*, *ghn*, *ghny*, *ghm*, *ghy*, *ghr*, *ghv*, *cm*, *cy*, *cr*, *chy*, *chr*, *chv*, *jñ*, *jm*, *jr*, *jry*, *jv*, *jy*, *ṭv*, *ḍy*, *ḍv*, *ḍhy*, *ḍhv*, *tn*, *tny*, *tnv*, *tm*, *ty*, *tr*, *try*, *tv*, *tvy*, *thn*, *thy*, *thr*, *thv*, *dn*, *dm*, *dy*, *dr*, *dry*, *drv*, *dv*, *dhn*, *dhny*, *dhm*, *dhy*, *dhr*, *dhry*, *dhv*, *pṇ*, *pn*, *pny*, *py*, *pr*, *pl*, *pv*, *by*, *br*, *bhṇ*, *bhn*, *bhm*, *bhy*, *bhr*, *bhv*, *śn*, *śny*, *śm*, *śy*, *śr*, *śrv*, *śl*, *śv*, *śvy*, *ṣṇ*, *ṣṇy*, *ṣṇv*, *ṣn*, *ṣm*, *ṣy*, *ṣr*, *ṣl*, *ṣv*, *sn*, *sny*, *sm*, *smy*, *sy*, *sr*, *sv*, *svy*, *hṇ*, *hn*, *hny*, *hm*, *hy*, *hr*, *hv*, *hvy*. Voiceless stops preceded by a sibilant (homorganic in the case of coronal stops): *śc*, *ścy*, *śch*, *śp*, *ṣk*, *ṣṭ*, *ṣṭy*, *ṣṭr*, *ṣṭry*, *ṣṭv*, *ṣṭh*, *ṣṭhy*, *ṣt*, *ṣth*, *ṣp*, *ṣpr*, *ṣph*, *sk*, *skr*, *st*, *str*, *stv*, *sty*, *sth*, *sp*, *sph*. Voiceless unaspirated stops followed by a sibilant: *kṣ*, *kṣṇ*, *kṣm*, *kṣy*, *kṣv*, *ks*, *ṭs*, *ts*, *tsn*, *tsm*, *tsy*, *tsv*, *pś*, *ps*, *psny*, *psv*. Sonorant plus sonorant: *ḻh*, *ṇy*, *ṇv*, *ṇvy*, *ny*, *nv*, *my*, *mr*, *ml*, *mv*, *vṇ*, *vn*, *vy*, *vr*, *vl*.[6] All of the extant intervocalic clusters consist of either the above onset clusters, or the same preceded by one of *k, ṅ, ṭ, t, n, p, m, r, l, s*, or the clusters *ṅk*, *rk*, *rt*, with the appropriate sandhi between the coda and the onset (recall that the normal sandhi of *r* at low prosodic levels is no change at all). These are exactly the word-final forms listed earlier as inputs to sandhi. The only exception here is *ṅk* in such words as *pṛ̥ṅkté* 'unites' (RV 10.95.9), whereas *k* is deleted after a nasal word-finally (*pratyáṅk* → *pratyáṅ*, RV 9.80.3). Perhaps the explanation is that *ṅk* is legal word-internally, since it falls within the same general pattern as *rk* and *rt* (sonorant plus stop). On the other hand, *ṅkt* is a legal sandhi of *ṅ* plus *t*; possibly the original *k* is basically deleted, but the text simply prefers to spell the optional sandhi *k* when it agrees with a stop that is etymologically justified (cf. imp. 3rd sing. *pṛ̥ṇáktu*, RV 1.84.1).

In all, there are over 400 different types of intervocalic consonant clusters (including clusters found at the juncture of compound words). The fact that all of them are explicable as onset clusters or the sandhi output of permitted codas before onset clusters is a strong indication that OM syllabification is operative in Sanskrit.

The same result is obtained when one looks at specific morphological operations. Some examples of the inflection of *j* stems were given above, where it was seen that *j* could be retained before vowels and sonorants, but not before obstruents. Verb inflections also give information about the syllabification of sequences of sibilant followed by stop. For example, in the root *diś-* 'point', the *ś* is converted to a stop when found in a coda, in such forms as *didiḍḍhi* (RV 2.32.6), where it converted to *ṭ*, caused the imperative suffix *-dhi* to retroflex, and assimilated. In the *s*-aorist *ádikṣi* (RV 5.43.9), it converted to *k* before *s*, causing retroflection in the latter. But the *ś* remains as a sibilant before *t* in the forms *dídeṣṭu* (RV 7.40.2) and *didiṣṭa* (RV 10.93.15; the change of *śt* to *ṣṭ* is a regular phonological change within words). This seems to be a counterexample to the claim of Clements (1990:298–299, 317) that core syllabification rules do not follow language-specific onset templates but instead obey the Sonority Sequencing Principle, whereby onsets must increase in sonority. In Clements's theory, sibilants and stops are of equal sonority, and so *ś* should stay in a coda before *t* or *th*, becoming a stop. It also calls into question Steriade's finding (1982:312–321, followed by Cho 1990:70–71 and Vaux 1992:297) that sibilants in Sanskrit do not syllabify with stops.

It might be objected that there are cases where a palatal converts to a velar, even when it should be syllabifiable in an onset. For example, *ric-* 'leave' has a perfect participle *ririkvā́ṃsaḥ* (RV 4.24.3); *arc-* 'shine, praise' has a nominalization *arkaḥ* 'ray, hymn'. That these are not coda-inspired depalatalizations can be seen from the fact that the velars do not subsequently assimilate voice. The confusion arises because the velars are actually the primary form. The palatals historically arose from velars by sound change before front vocoids, then spread analogically throughout the verb inflections, especially the finite forms. So *arka-* can be straightforwardly derived from *ark-* (or treated as an historical survival). Similarly, the few cases in the verbal paradigms where *c* or *j* are avoided before *v* or *r* are no doubt to be considered survivals; presumably palatalization was held off longer in this phonetically unfriendly environment.

In a similar vein, it has also been claimed that *dn* is not a possible onset because the passive participle ending *-ná-* causes preceding dental stops to turn into *n*, as would be expected in sandhi (Vaux 1992:296): *bhid-* 'split' → *bhinnáḥ* (RV 1.32.8). But the same does not obtain in words like *udnáḥ* 'waves' (RV 8.32.25). Perhaps the participial paradigm is a survival from an earlier period when *tn* and *dn* were not acceptable onsets. One can easily imagine a diachronic state of flux between a system that rejects these clusters for being homorganic and one that accepts them by analogy with the other stop plus nasal sequences.

3.2.4 Late metre.

The OM theory has its own metrical evidence, although here the evidence is very late. Hermann (1923:257, 261) and Vaux (1992:298) have called attention to the fact that later Sanskrit poetry, such as that of the Epics and of Buddhist Hybrid Sanskrit texts, have many clusters that do not make the preceding syllable heavy. Citing Varma (1929:95), Vaux lists for the Epics the clusters *pr*, *kṣ*, *śr*, *tr*, *kr*, *dr*, *br*, *vy*, *bhr* and *ddh*; for BHS, he lists *sr*, *ṣy*, *ṣṭh*, *nn*, *rh*, *sm*, *sph*, *sp*, *sn*, *sth*, *sk*, *sv*, *dhy*, *cy*, *tv*, *gr*, and *jñ*. The obvious conclusion is that these are onset clusters, although Epic *ddh* and BHS *nn* are a little hard to believe.

3.2.5 Summary of the evidence for OM.

Consonant sandhi as an assimilatory phenomenon makes the most sense if one interprets it not as a word boundary phenomenon with some unaccountable overlap into morphology, but as a property of syllable codas. Therefore if consonant C fails to undergo before word-internal X the changes that it would suffer if a word boundary (ergo syllable boundary) intervened, then the first explanation must be that C is not in a coda, i.e., that CX is a possible onset cluster. Since this often happens in a binary cluster which the CS theory would obligatorily split up, this is strong evidence for the OM theory. It is reinforced by the fact that complex onsets do occur word-initially, and that all medial clusters are analysable as codas (with appropriate sandhi) followed by onset sequences whose pattern is consistent with the patterns evidenced by studying sandhi effects and the inventory of word-initial sequences (as well as rather late evidence from the metre).

Specifically, the onset template accepts:

1. A segment followed by a segment of greater sonority. This pattern is found widely throughout the languages of the world, and is predicted by the Sonority Sequencing Principle (Sievers 1876:111–113), which says that sonority should increase monotonically the closer one gets to the sonority peak, the vowel. The universal sonority hierarchy used in Vogel (1977:135–136) and Clements (1990:292–298) fills the bill exactly: obstruents are less sonorous than nasals, which are less sonorous than liquids, which are less sonorous than glides.

2. Non-palatal voiceless unaspirated stops before *s*. The evidence comes mostly from word-initial sequences, Epic metre, and the authority of the Taittirīyaprātiśākhya. But sandhi tests cannot confirm this, since sounds like *k* would remain unchanged before *s* whether in a preceding syllable or not. And the sequence violates Clements's version of the Sonority Sequencing Principle, although it would be acceptable by the original and more common formulation that says that fricatives are more sonorous than oral stops (Sievers 1876:112, Kiparsky 1979:432; cf. Clements 1990:315–316).

3. Voiceless fricatives (the sibilants *ś*, *ṣ*, *s*) before voiceless stops, with the dental *s* assimilating in place to a following coronal. This also violates the Sonority Sequencing Principle, except for those versions that permit sequences that are equisonorous or nearly so (Hermann 1923:5), but seems well justified not only by word-initial sequences and BHS metre but also by the sandhi of *ś* and *ṣ*. The case of *s* is harder to prove, since it only disarticulates at the end of phonological phrases. Steriade (1982:312–333) argued that *s* does not form onset clusters, on the basis of reduplication and the

fact that *s* disappears between stops in the morphology. In her view, the fact that a word-initial *s* before an oral stop is not copied by reduplication processes means that it is not syllabified into the onset, and hence is invisible to the copying process. But in Gothic, the entire *s*-stop cluster reduplicated, even though otherwise Gothic too only reduplicated one consonant: *stai-staut* 'pushed' vs. *gai-grot* 'wept' (Kiparsky 1979:435), so syllabic invisibility may not be the proper analysis of sT peculiarities in reduplication.

4. Labial sonorants (*m*, *v*), and *h* before sonorants. From the standpoint of the Sonority Sequencing Principle, this is shocking. The glide *v* should be among the most sonorous of all sounds, and hence not followable by any other segment in the onset.[7] And certainly nasals should not be followed by nasals, if one subscribes to the version of the theory by which onset consonants must strictly increase in sonority.[8] There does seem good evidence for the clusters, however. Words begin with *ny*, *nv*, *my*, *mr*, *ml*, *vr*, *vy*, *hn*, *hr*, *hl*, *hv*, *hy*. While *m* disarticulates or assimilates before onsets when in a coda, it is found word-internally before *n*, *y*, *r*, *l*, and *v*. Similarly *av* is found before *ṇ*, *y* and perhaps *r* word-internally, instead of monophthongizing to *e*. And the morphology shows that *h* is always converted to a stop when in a coda, so *hṇ*, *hn*, *hm*, *hr*, *hl*, *hy*, *hv* must be onset sequences.

4 Synthesis.

4.1 Description.

Up to this point I have painted two radically different pictures of Vedic Sanskrit syllabification. On the one hand, there is the language where consonant clusters split in two, with one consonant being syllabified with the preceding syllable, regardless of the quality of the consonants in question. On the other hand there is the language where the bulk of the phonology and much of the morphology centres around the distinction between consonant clusters that do and those that do not syllabify together into onsets. There would seem to be a fundamental discrepancy here.

It has been suggested that the two systems be reconciled by assigning them to two different historical periods. Chatterji (1926:252) suggested that Vedic Sanskrit was a cluster splitter, but Classical Sanskrit was an onset maximizer; in fact, he seemed to believe that all intervocalic clusters in Classical Sanskrit formed onsets, which for him explained why the native scripts write virtually all consonant strings plus the following vowel (C^*V) as a single bundle.[9] Vaux (1992) preferred the more moderate approach of saying that Classical Sanskrit changed from a CS strategy to an OM strategy where onsets are structured pretty much as presented here (a major difference being that he agrees with Steriade 1982 and Cho 1990 that sibilants do not syllabify with a following stop). These approaches, however, do not take into consideration the unity of sandhi and morphophonemics that was outlined here, which requires OM syllabification even within the Vedic period. The hypothesis of an historical change would necessitate either developing a completely different explanation of why the phonology seems to behave as if there is OM, or a radical reconstruction of the text of the Ṛgveda, arguing that much of the sandhi visible in the text is an innovation of later editors. It also requires one to reject synchronic explanations for those features in later Sanskrit that still require CS syllabification, such as variations in the weight of reduplicate prefixes. All such forms would have to be seen as fossilized survivals, and correct formulations for new words would have to be seen as artificially constructed. All this is possible, of course, but is on its face a rather complicated argument.

Selkirk (1982:360–379) favours resyllabification as a general solution to such syllabification paradoxes. Under such a scheme, there would be an OM limited to suffixes, with clusters forming onset clusters when legal, and others splitting into coda and onset, with the codas subject to neutralization. In addition, there would be a later resyllabification not only across word boundaries but between all syllables. At this point, a word like *va-jra* would become *vaj-ra*, making the verse metre work out right. This could be refined by providing that OM takes place only in an intermediate lexical stratum; reduplicating prefixes are not subject to it because they are affixed earlier in the derivation, and compounds are not subject to it because they are later.

What is troubling about this analysis is that it posits that the postlexical phonology undoes much of what the earlier phonology worked to accomplish. Whereas onsets are preferred in suffixation, the postlexical phonology prefers to make heavy syllables, even if it has to take onsets apart (Borowsky *et al.* 1984:36). Whereas the lexical phonology is hostile to having segments like the palatal *j* in a coda, the postlexical phonology will take *j* out of a perfectly legal onset cluster in order to place it in a coda. It is curious that different parts of the phonology should have such violently conflicting notions of optimality. A more empirical problem is that some parts of the grammar appear to be both CS-oriented and OM-oriented at the same time, defying any neat stratification. Consider again the alternation of *i* versus $\emptyset$ in the perfect forms:

yuyuj- 'yoked'	*yuyujma* 1 pl. (RV 8.61.12)	*yuyujé* 1 sg. (RV 1.161.6)
jajñ- 'was born'	*jajñiṣe* 2 sg. (RV 10.133.2)	*jajñe* 1 sg. (RV 9.114.2)
ūc- 'said'	*ūciṣe* 2 sg. (RV 3.22.3)	*ūce* 1 sg. (RV 10.40.5)

The *i* appears between consonants after a heavy syllable, but it is omitted before vowels. Since the measure of heaviness is indifferent to whether the final cluster of the stem could form a legal onset (as does the *jñ* in *jajñiṣe*), this *i* must be deleted (or added) in the CS stratum of the grammar. But the *i* is also sensitive to the phonology of the suffix, in that it is always omitted before a vowel. This latter behaviour is idiosyncratic, since *i* normally becomes *y* before a vowel, so it cannot simply be referred to the postlexical phonology. Now suffixes would be added at the OM stratum of the grammar, as is clear from the case of *yuyujma*, where *j* is not neutralized because it forms part of an onset cluster with the suffix. So the treatment of *i* appears to be determined at two different strata with conflicting syllabification properties. An even clearer case of the inadequacy of resyllabification is presented by *jajñiṣe* itself. The *i* is present because the preceding syllable is heavy, but the second *j* exists as such because it is in an onset cluster *jñ*. Postlexical resyllabification will not account for the discrepancy that the *j* behaves both as a coda element and as an onset element at the moment when the perfect suffixes are added.

The cleanest solution involves recognizing ambisyllabicity. Instead of concluding that CS and OM occur at different times historically or at different parts of the grammar, it is better to say that both are present simultaneously. For *jajñiṣe*, the second *j* is both in the first syllable, making it heavy, and in the second syllable's onset, rescuing it from neutralization. The rule would be that in any domain allowing free syllabification (i.e., a root and its right-appended elements, viz., suffixes), an onset is maximized according to the template discussed in the section on evidence for OM. But the first consonant of a cluster is also ambisyllabically linked to the preceding syllable, regardless of the prosodic domain.[10] This clearly happens immediately after short vowels, since the preceding syllable is made heavy, but internal evidence is lacking in other cases. The phonology is neither tearing down ('minimizing') an onset, nor is it moving a segment into a hostile position: elements in codas can contain forbidden features as long as they are linked to an onset position, which licenses them.

4.2 Motivation.

This is a rather different case of ambisyllabicity than the familiar Germanic ones. In those languages ambisyllabicity probably arose because of a drive to align stress and moric prominence: i.e., stressed syllables had to be heavy. Therefore a single consonant after a short stressed vowel had to be geminate. The fact that gemination was largely redundant made it particularly susceptible to phonetic loss, leaving ambisyllabicity as its structural remnant (Kuryłowicz 1948:208). In Sanskrit, on the other hand, stress is not involved,[11] and there is no obvious breakdown of an earlier gemination. Rather, the system arguably dates back to Proto-Indo-European. One will note that early Greek poetry had a very similar metrical system, where for example *kr* made the preceding syllable heavy even though *kr* is good at the beginning of a word and *k* is not legal at the end of a word. Clearly some other principle is coming to an accommodation with the maximum onset requirement, but what is it?

Taking a cue from Germanic, one might hypothesize that the ambisyllabicity is primarily conditioned by the preceding syllable. Since there is positive evidence for ambisyllabicity only after a light syllable, that is, after a short, unchecked vowel, one might hypothesize that the short vowel is attracting a consonant. But this scenario is weak on several counts. First, there is no evidence that the ambisyllabicity does not occur after other types of syllables: the diagnostic of syllable weight simply yields no answer after a heavy syllable. Secondly, it doesn't really make sense for a language to unconditionally want to close syllables after a short vowel. In Germanic, doing so after a short stressed vowel is a way of aligning prominences of weight and stress. In Sanskrit, closing all short syllables would simply serve to make as many syllables heavy as possible. Most importantly, the ambisyllabicity only happens when there are consonant clusters. With a single consonant, Sanskrit has a phonologically unpredictable lexical contrast between simple and geminate, the latter always expressed as long: *vúkas* 'wolf' (RV 1.105.18) vs. *vr̥kkás* 'averter of disease?' (RV 1.187.10). A single consonant is never made to be ambisyllabic just because it follows a short vowel, although that is exactly what we would expect if the motivation is simply to check such vowels. After all, the Germanic languages make single consonants ambisyllabic, and Netherlandic even appears to be the opposite of Sanskrit, making single consonants ambisyllabic, but not clusters (Hulst 1985:65). A similar argument would apply to a hypothesis that Sanskrit attempts to balance a maximum coda requirement against the maximum onset requirement: singleton consonants should not be immune.

The most obvious explanation is that the main motivation for the ambisyllabicity is the consonant cluster itself. A cluster of two or more consonants wants to be bisyllabic at all levels of the grammar, even in contexts where there is a separate constraint for maximum onsets. In these cases both constraints can be solved by means of ambisyllabicity, since all relevant consonants can be found in codas,[12] especially when linked to a following onset.

I would further speculate that the specific property of clusters that makes them want to be bisyllabic is weight. Here I conceive of weight not as a property peculiar to syllables but as a feature inherent in segments that is inherited by the constructs they are found in. In general segments have one mora apiece, although geminates have two. In a construct, moras are summed up to the limit of two (reflecting a general tendency in phonology to make binary distinctions), so segments can make a rime light or heavy. A syllable is headed by the rime, so it inherits only the weight of the rime, not that of the onset. But the weight of the consonants in the onset could still be significant for certain other purposes.

That heavy segments (i.e., geminates) want to be bisyllabic is a universal (Clements 1990:321–322). Sanskrit generalizes this constraint to heavy interludes in general. Thus Sanskrit distinguishes between light and heavy interludes, where a light one is a single, simple consonant, and a heavy one is a geminate consonant or any cluster: light VCV, heavy VC:V or $VCC^{+}V$. The language does not permit any distinctive length among elements in branching interludes (consonant clusters).[13] This is exactly analogous to a distribution often found in rimes in languages like Pali (Hermann 1923:260), where there is a distinction between short and long vowels, but not in branching rimes (closed syllables).

Why should Sanskrit have a more general heavy-interlude constraint? Possibly there is a connection with its verse form, which is sensitive to weight. It may be the case that languages select as their fundamental verse structure their fundamental timing unit. In Germanic languages, this is the stress foot; in Spanish and French, this is the syllable; in Sanskrit and Ancient Greek, this is weight. Conceivably languages timed by weight are particularly sensitive to the weight of interludes. A general requirement that heavy interludes be ambisyllabic would tend to make the weight of clusters metrically relevant, since that would make preceding light syllables heavy.

It will be noted that all documentable cases of ambisyllabicity in the R̥gveda involve onset clusters, so one may object that I should be talking about heavy onsets instead of heavy interludes in general. The distinction is important. Onsets may not generally be considered to have weight (inasmuch as they do not influence the weight of their syllables), but at least there is a strong consensus that they exist. Interludes are not generally given independent status in current phonological theory. Nevertheless structuralists such as Kuryłowicz (1948) found it useful to speak of the structure of intervocalic consonant complexes. And evidence to be presented in the next section indicates that further developments in Indic cannot be based on onset weight alone. Interludes are clearly not part of the prosodic hierarchy, since they do not align consistently with the codas and onsets of syllabls. Rather, they should be thought of as a sort of presyllabic organization of segments on a separate tier.

4.3 Later developments.

All the evidence stated thus far has been from internal reconstruction of the R̥gveda text. There is also evidence from Indian linguists, early inscriptions, and the structure of daughter languages. All this, however, is very late, perhaps six or seven centuries later than the composition of the core R̥gveda ca. 1000 BCE. There is the very real possibility that the language is significantly different by this time: I have already mentioned that Vaux believes that Sanskrit completely changed its syllabification strategy by this time. Sanskrit was still alive as the language of the Brahmanic class, but had evolved signifcantly beyond the Vedic stage. The popular Indic speech went much farther still, such that it is considered to have developed into separate languages, and in particular the phonology was much altered. Therefore it is probably best to consider these data not as evidence for Vedic Sanskrit per se, but as the next step in the evolution of the language. They are evidence insofar as they constitute a plausible development from the state of affairs posited for Vedic.

An immediately encouraging fact is that where my account calls for ambisyllabicity, the grammarians state that the consonant may be doubled, i.e., pronounced long. One might only wonder whether I should have been talking about gemination all along rather than ambisyllabicity. But some of the grammarians said that the lengthening was optional (e.g., Pāṇini 8.4.47, under the scope of *vā* in rule 45), and others (e.g. Śākalya apud Pāṇini 8.4.51) denied the lengthening altogether, so it seems safest to assume that Sanskrit had structural bisyllabicity, which alternated between ambisyllabicity (no lengthening) and gemination (lengthening). Similarly, the earliest inscriptions often, though inconsistently, showed doubling there (Varma 1929:64-67).[14] Finally, the Middle Indic dialects, which evolved from Old Indic dialects similar to the language of the R̥gveda, show clear effects of such doubling. A word like *atra*, for example, evolves into *atta*. If one takes ambisyllabicity of the *t* as the base form, with optional doubling, the resultant form *atta* is easy to understand, given the independent fact that these dialects simplified all tautosyllabic clusters. The change is difficult to motivate otherwise. (The argument that Middle Indic simplifications have an origin in Sanskrit doublings goes back at least to Jacobi 1881:609, and Cho 1990:194–231 and Vaux 1992 have worked to put that in a syllable-based context.) All this evidence indicates that the bisyllabicity of onset clusters after a vowel was a real property of early Indic.

Unfortunately, these doublings are not completely restricted to the contexts where ambisyllabicity was posited. The situation is really very difficult to evaluate. Allen (1953:79) despaired of even discussing the topic. Cho and Vaux have recently made heroic attempts to harmonize those accounts and to explain them in terms of syllable theory, but it must be admitted that many of the pronouncements make broad claims that other sources seem to step on. For example, Pāṇini doubled *s* before stops; the Taittirīyaprātiśākhya (14.9, apud Vaux 1992:286) doubled stops after *s*. But it is possible to reconstruct core environments that the great majority of the grammarians would agree on, and that are compatible with the evidence from the inscriptions and the daughter language developments. One is that an oral stop doubles between a short vowel and an approximant. This is quite close to my rule for ambisyllabicity. The restriction that the vowel must be short (due to Pāṇini 8.4.52) may mean that the bisyllabicity occurs only when it would actually have a non-vacuous effect on syllable weight. On the other hand, the rule could just be a reflection of the fact that the colloquial dialects resist closing a syllable after a long vowel (Hermann 1923:260).

The other doubling that occurs almost universally is that of a stop after *r*. Since *r* plus stop can never make an onset cluster, they would always be heterosyllabic, and the prior syllable therefore always heavy. I therefore know of no internal evidence that would tell whether these sequences are ambisyllabic in Vedic Sanskrit. The fact that Greek does not show this doubling may mean that this particular rule is an Indic innovation (Vaux 1992:288), and so may plausibly be post-Vedic. But if the actual rule is that heavy interludes (not just heavy onsets) want to be ambisyllabic, as suggested above, then this case naturally falls under the scope of that rule. The sequence *r* plus stop is the only clearly documented coda sequence in the language. Therefore it is a clear situation where a non-onset sequence can be made bisyllabic. In contrast, a sequence of two stops rarely is doubled in the inscriptions, and there is no evidence for that in the daughter languages. Any bisyllabicity there would require admitting them as coda clusters or as onset clusters, neither of which is possible for two stops.

This argumentation is a little too pat, however. One cannot afford to turn away from the cases where there is conflicting evidence. Many sources, for example, provide that nasals geminate before stops (Taittirīyaprātiśākhya 14.24 apud Vaux 1992:290) or after sibilants (14.9 apud Vaux p. 286), which are incomprehensible from a syllable-theoretic point of view (*nt* cannot be an onset, and *sm* cannot be a coda) and have caused much consternation over the years (Hermann 1923:256, Cho 1990:209–210). Ultimately in some dialects of Sanskrit the rule must have been restructured into a requirement that length be associated with some member of a heavy interlude. If this is the case, then it is difficult to know whether the doubling of stops after *r* is the result of an ambisyllabicity requirement, or whether it simply reflects the fact that *r* cannot be doubled, so the stop is doubled in compensation, as it were. At this point, it seems safest to consider the doubling of stops before approximants as good evidence of bisyllabicity, but to consider conclusions about other doublings to be tentative, pending a complete analysis of the Sanskrit doublings and Middle Indic developments.

4.4 Representation.

All the evidence in Sanskrit points to a representation for ambisyllabicity that is no different from that for gemination. Like geminates (on which see Hayes 1986), ambisyllabic segments show ambiguity, behaving as onset segments from the point of view of phonotactics, like coda segments for the purpose of determining weight, and yet with respect to pronunciation, like a single segment. They also show a certain inalterability: even though an ambisyllabic segment closes a syllable and makes position, it is not subject to coda neutralizations. Secondly, they appear to slip in and out of a pronunciation where they are pronounced long, as true geminates. This variation, which is both synchronic and diachronic, is easiest to account for if it is a matter of two noncontrastive phonetic realizations of the same structure. Most importantly, this ambisyllabicity and gemination never contrast, even though the language has both phenomena. Sanskrit contrasts long and short consonants between vowels, as in the aforementioned near-minimal pair *vŕ̥kas* vs. *vr̥kkás*. But this contrast obtains only in the case of a single segment, whereas ambisyllabicity in Sanskrit obtains only in the case of consonant clusters, so the difference between the two cases does not require any further structural distinction. The fact that ambisyllabicity and gemination have distinctive environments even when both occur in the same language bolsters the claim that they have identical representation. As for what that is, exactly, my inclination is to conceive of this as a segment being linked into the rime of one syllable and also into the onset of the next, but I do not believe that the data contain any crucial evidence that argue against other versions of autosegmental phonological theory, such as the intermediation of a tier of timing slots, or taking moras to be the top-level constituents of the syllable. What is crucial is that a single segment be linked simultaneously into two adjacent syllables.

4.5 Conclusion.

The Sanskrit data constitute a strong case for the existence of ambisyllabicity. They cannot be explained away by some of the alternatives presented to Kahn's data for English. Kiparsky's invocation of stress feet (1979) is obviously inapplicable here, and Selkirk's resyllabification (1982) cannot explain situations such as the formation of perfects, which is sensitive to both syllabifications simultaneously. The fact that this ambisyllabicity cannot be connected with the Germanic data in terms either of areal contact or of historically antecedent structures (the breakdown of contrastive gemination in the Germanic cases) bolsters the case for ambisyllabicity's being a general linguistic phenomenon. It is also encouraging that this ambisyllabicity eventually developed a phonetic correlate (length), which proves that the bisyllabic structures were a living part of the language, as opposed to the possibility that the maximal onset patterns found in the morphology were frozen remnants of an earlier phonology.

In addition, the fact that ambisyllabicity is associated with gemination both in Germanic (gemination develops in ambisyllabicity) and in Sanskrit (ambisyllabicity develops into gemination) bolsters the case for their having very similar bisyllabic representations. Indeed, the fact that ambisyllabicity and gemination never contrast in Sanskrit even though both are present in the language strongly supports the conjecture of Borowsky *et al.* (1984) that the two phenomena have identical phonological representations, distinguished only at a phonetic level. Although Suzuki (1985) and Kang (1991) claim to have found evidence of the contrast in Old English and Korean, respectively, their evidence is based on the very weak argument that alternations conditioned neither by onsets nor by codas must be conditioned by ambisyllabicity. Since alternations can be caused by factors other than syllable structure alone, it seems best to provisionally reject their claim in favour of that of Borowsky *et al.*

Notes.

[1]Some of these terms have been used in conflicting ways. ‘Ambisyllabic’ has occasionally been applied, and quite reasonably so, to describe any consonant shared between two syllables (e.g., Abu-Salim 1982:23), but I here follow the more conventional practice of restricting its use to short segments. ‘Geminate’ shall refer only to the case where a segment shared between two syllables is long. I shall use the cover term ‘bisyllabic’ to refer indifferently to either phenomenon. Although the term ‘interlude’ is most commonly applied to ambisyllabic consonants, I shall here employ it to refer to all intervocalic sequences of one or more consonants, regardless of their syllabification.

[2]The approximants, traditionally called semivowels, comprise the liquids *r*, *l* and the glides *y*, *v*, as well as the derived *ḻ*, *l̃*, *ỹ* and *ṽ*. Among the sonorant consonants, they exclude only the nasals, *m*, *n*, *ṇ* and derived *ṅ*, *ñ* and *ṃ*. The stops comprise the velars *k*, *kh*, *g*, and *gh*; the palatals *c*, *ch*, *j*, and *jh*; the retroflexes *ṭ*, *ṭh*, *ḍ*, *ḍh*; the dentals *t*, *th*, *d*, and *dh*; and the labials *p*, *ph*, *b*, and *bh*. The fricatives include the sibilants *s*, *ś* and *ṣ*, as well as the unarticulated (glottal) *h* and the derived sounds *ḥ*, ϕ and χ.

[3]For ease of comparison with virtually everything else written about Sanskrit, I use standard Indological transcription. The acute accent on vowels shows high pitch. Vowels are short unless marked by a macron, except that *e* and *o* are always long. Vowels have values as in IPA, except that short *a* is [ə]. Inferior dots show that consonants are retroflex, except in the case of *ṃ* where it shows lack of oral articulation, and *ḥ*, where it shows that the [h] is voiceless (the normal *h* is breathy-voiced). *h* after a stop is a sign of aspiration, not a separate segment. Inferior circles show that a liquid is syllabic. Tildes show that the sound is nasal, except that *ñ* marks a palatal nasal [ɲ]. *ś* is a palatal fricative [ʃ], *ṅ* is a velar nasal [ŋ]. *v* is pronounced [ʋ] at the beginning of a syllable, else [w]; phonologically, it is best treated as a glide.

[4]In general I will use ‘word-final’ as a cover term embracing also the boundaries between stems in compounds, and between prefix and stem, since most external sandhi changes happen in this environment as well as between full words.

[5]For Vedic, these include nasalization of stops before nasals; reduction of *ay* to *a* before a vowel; change of *ns* to *nn* or *ṃr* before a vowel; assimilation of *m* to glides and *l*; and disarticulation of *m* before *r*. In addition, the processes of dental assimilation to coronal stops and the defricativization of *ś* or *h* after a stop only happen to occur at word boundaries and so are compatible with either theory. I know of no process that occurs only at the end of the word in the broad sense here employed (i.e., also including the end of prefixes and compounding stems), but it is the case that *r* desonorizes, becoming eligible for syllable-final obstruent-like assimilations, only at the end of a word proper. Crucially, the change that is conditioned by word juncture in this case is a neutralization, not an assimilation.

[6]I have taken the glides *y* and *v* of the Ṛ̥gveda text at face value, even though in many cases the metre depends on their being syllabic. The glide and the vowel often alternate for the same word ad libitem, and the complexity of the resultant onset cluster does not seem to be a determining

factor.

[7]Cho 1990:201–204 makes the appealing hypothesis that *v* should be treated as an obstruent, because of this distribution and the fact that it is labiodental [ʋ]. But in every other respect it patterns with the glide *y* as the nonsyllabic form of *u*. Besides, the labiodental pronunciation is an innovation over Proto-Indo-European; calling *v* an obstruent would not explain why *w* could precede segments like *r* in other Indo-European languages, leaving traces even in the English *writing* system.

[8]As for *h*, Clements (1990:322) notes that laryngeals fall outside of the sonority scale, so that cross-linguistically they quite often pattern with obstruents, particularly if that is their historical origin. In Sanskrit, *h* derives from voiced aspirated stops.

[9]The only exception is that the unarticulated (laryngeal) sounds *ṃ* and *ḥ* are represented as diacritics on the preceding bundle, probably reflecting a phonetic analysis of them as being nasal or voiceless continuations of the preceding vowel.

[10]In contrast, the postlexical rule of liaison, which gives a coda consonant to the immediately following vowel anywhere within the half-verse, is a movement rule. The consonant remains in the coda long enough to undergo neutralization and assimilation, and yet after liaison, it does not count toward making the syllable heavy: *tát a-stu* 'let that be' → *tád a-stu* → *tá-da-stu* (RV 1.30.12).

[11]There is no direct evidence that Vedic Sanskrit had stress at all, although grammarians describe a Latin-like stress system for the later Classical language. At any rate, there is no correlation between ambisyllabicity and either that stress system or the Vedic pitch accent.

[12]Aspirates are apparently an exception, since in texts they virtually never appear in a coda (in the Rigveda they do so twice, in geminates). In gemination, the first element is replaced by the corrsponding unaspirated sound. This may simply be an acknowledgement that aspiration is noticeable only when a stop is released, but the first element of a geminate is not released. Cho (1990:207) is surely correct in concluding that aspirates should be treated as true geminates phonologically, with any deaspiration to be referred to the phonetic implementation. If this approach is taken, the specific lack of the sound *h* in a coda can be referred to as an accidental gap.

[13]The restriction is of course not present in all languages. In Italian, for instance /br/ in *libro* 'book' contrasts with /bbr/ in *labbro* 'lip'.

[14]In Greek as well, inscriptions often double ambisyllabic consonants, although this is less common for stops before liquids than for less favoured onset types such as obstruents before nasals or other obstruents (Hermann 1923:110–123).

References

Abu-Salim, Issam M. (1982). Syllable structure and syllabification in Palestinian Arabic. *Studies in the Linguistic Sciences* **12:1**. 1-28.

Allen, W.S. (1953). *Phonetics in ancient India.* Reprinted 1961, London: Oxford University Press.

Allen, W.S. (1962). *Sandhi: the theoretical, phonetic, and historical bases of word-junction in Sanskrit.* 's-Gravenhage: Mouton.

Ananthanarayana, H.S. & W.P. Lehmann, eds. (1970). *Ṛgveda.* Computer file on deposit at the Oxford Text Archives.

Anderson, John & Charles Jones. (1974). Three theses concerning phonological representations. *JL* **10**. 1–26.

Booij, Geert E. (1983). Principles and parameters in prosodic phonology. *Linguistics* **21**. 249–280.

Borowsky, Toni, Junko Itô & Ralf-Amin Mester. (1984). The formal representation of ambisyllabicity: evidence from Danish. *NELS* **14.** 34–48.

Chatterji, Suniti Kumar. (1926). *The origin and development of the Bengali language.* Reprinted 1970, London: Allen & Unwin. Part I: Introduction, phonology.

Cho, Young-mee Yu. (1990). *Parameters of consonantal assimilation.* PhD dissertation, Stanford.

Clements, George N. (1990). The role of the sonority cycle in core syllabification. In John Kingston and Mary E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech.* Cambridge: Cambridge University Press. 283–333.

Collinge, N.E. (1985). *The laws of Indo-European.* Amsterdam: Benjamins.

Grassmann, Hermann. (1875). Wörterbuch zum Rig-Veda. 4. unveränderte Aufl., 1964. Wiesbaden: Otto Harrassowitz.

Gussenhoven, Carlos. (1986). English plosive allophones and ambisyllabicity. *Gramma* **10**. 119–141.

Hayes, Bruce. (1980). *A metrical theory of stress rules.* PhD dissertation, Yale. Distributed 1981, Indiana University Linguistics Club.

Hayes, Bruce. (1986). Inalterability in CV phonology. *Lg* **62**. 321–351.

Hermann, Eduard. (1923). *Silbenbildung im Griechischen und in den anderen indogermanischen Sprachen.* (Ergänzungsheft zur Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen, 2) Göttingen: Vandenhoeck & Ruprecht.

Hock, Hans Heinrich. (1980). Archaisms, morphophonemic metrics, or variable rules in the Rig-Veda? *Studies in the Linguistic Sciences* **10**. 59–69.

Hooper, Joan B. (1978). Constraints on schwa-deletion in American English. In J. Fisiak (ed.) *Recent developments in historical phonology.* The Hague: Mouton. 183–207.

Hulst, Harry van der. (1985). Ambisyllabicity in Dutch. In Hans Bennis and Frits Beukema (eds.) *Linguistics in the Netherlands 1985.* Dordrecht: Foris. 57–66.

Jacobi, H. (1881). Zur Genesis der Prâkṛitsprachen. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* **25**. 603–609.

Kahn, Daniel. (1976) *Syllable-based generalizations in English phonology.* PhD dissertation, MIT. Published 1980, New York: Garland.

Kaisse, Ellen. (1985). *Connected speech.* Orlando: Academic Press.

Kang, Seok Keun. (1991). Moraic representation of ambisyllabicity: evidence from Korean. *Studies in the Linguistic Sciences* **21:2**. 89–100.

Kessler, Brett. (1993). Syllables and sandhi in Classical Sanskrit. Paper delivered at the West Coast Conference on Formal Linguistics XII, Santa Cruz, Calif.

Kiparsky, Paul. (1972). Metrics and morphophonemics in the Rigveda. In Michael K. Brame (ed.) *Contributions to generative phonology.* Austin: University of Texas Press. 171–200.

Kiparsky, Paul. (1979). Metrical structure is cyclic. *LI* **10**. 421–441.

Kuryłowicz, Jerzy. (1948). Contribution à la théorie de la syllabe. Reprinted in his *Esquisses linguistiques*, 2nd ed., 1973. München: Wilhelm Fink. Vol. 1. 193–220.

Lass, Roger. (1985). Minkova noch einmal: MEOSL and the resolved foot. *Folia Linguistica Historica* **6**. 245–265.

Macdonell, Arthur Anthony. (1916). *A Vedic grammar for students.* Reprinted 1941, Oxford: Oxford University Press.

Pāṇini. *Aṣṭādhyāyī of Pāṇini*, in Roman transliteration by Sumitra M. Katre. Austin: University of Texas Press, 1987.

Picard, Marc. (1984). English aspiration and flapping revisited. *Canadian Journal of Linguistics* **29**. 42–57.

Pulgram, Ernst. (1970). *Syllable, word, nexus, cursus.* The Hague: Mouton.

Rice, Keren D. (1990). Predicting rule domains in the phrasal phonology. In Sharon Inkelas and Draga Zec (eds.) *The phonology-syntax connection.* Chicago: University of Chicago Press. 289–312.

Seebold, Elmar. (1972). *Das System der indogermanischen Halbvokale.* Heidelberg: Winter.

Selkirk, Elisabeth O. (1982). The syllable. In Harry Van der Hulst and Norval Smith (eds.) *The structure of phonological representations.* Dordrecht: Foris. Part II. 337–383.

Sievers, Eduard. (1876). *Grundzüge der Lautphysiologie.* Leipzig: Breitkopf und Hartel.

Steriade, Donca. (1982). *Greek prosodies and the nature of syllabification.* PhD dissertation, MIT.

Suzuki, Seiichi. (1985). The role of syllable structure in Old English poetry. *Lingua* **67**. 97–119.

Vaidika Saṁśodhana Maṇḍala. (1972). *Ṛgveda-saṁhitā with the commentary of Sāyaṇācārya.* Vol. 1–2, 2nd ed., 1972–1976; v. 3–4. 1st ed., 1941–6. Poona: Vaidika Saṁśodhana Maṇḍala.

Varma, Siddheshwar. (1929). *Critical studies in the phonetic observations of Indian grammarians.* Reprinted 1961, Delhi: Munshi Ram Manohar Lal.

Vaux, Bert. (1992). Gemination and syllabic integrity in Sanskrit. *The Journal of Indo-European Studies* **20**. 283–303.

Vennemann, Theo. (1982). Zur Silbenstruktur der deutschen Standardsprache. In his (ed.) *Silben, Segmente, Akzente.* Tübingen: Niemeyer. 261–305.

Vogel, Irene Barrie. (1977). *The syllable in phonological theory, with special reference to Italian.* PhD dissertation, Stanford.

Whitney, William Dwight. (1889). *Sanskrit grammar.* 2nd ed. Reprinted 1964, Cambridge, Mass.: Harvard University Press.

	AP	TP	RP
‘prosperity’ (1.1.3)	*pó-ṣam*	*pó-ṣam*	*pó-ṣam*
‘on all sides’ (1.1.4)	*viś-vá-taḥ*	*vi-śvá-taḥ*	*viś-vá-taḥ, vi-śvá-taḥ*
‘Agni’ (1.1.1)	*ag-ním*	*ag-ním*	*ag-ním, a-gním*
‘may he conduct’ (1.1.2)	*vak-ṣa-ti*	*va-kṣa-ti*	*vak-ṣa-ti, va-kṣa-ti*
‘please’ (1.1.6)	*aṅ-gá*	*aṅ-gá*	*aṅ-gá, a-ṅgá*

Table 1: Syllabification according to the rules of most Prātiśākhyas, such as the Atharvaprātiśākhya (AP), and the Taittirīyaprātiśākhya (TP) and Rkprātiśākhya (RP) in particular. References to the Rgveda (RV) throughout this paper are in the form of book (*maṇḍala*), hymn (*sūkta*) and verse.

	k	*ṅs*	*ṭ*	*t*	*ns*	*p*	*m*	*ay*	*r*[1]	*av*	*s*	*as*	*ās*
pausa	*k*	*ṅ*	*ṭ*	*t*	*n*	*p*	*m*	*e*	*ḥ*	*o*	*ḥ*	*aḥ*	*āḥ*
k(h)-	*k*	*ṅ*	*ṭ*	*t*	*n*	*p*	*ṅ*	*e*	*χ*[4]	*o*	*χ*[4]	*aχ*[4]	*āχ*[4]
g(h)-	*g*	*ṅ*	*ḍ*	*d*	*n*	*b*	*ṅ*	*e*	*r*	*o*	*r*	*o*	*ā*
c(h)-	*k*	*ṅ*	*ṭ*	*c*	*ñ*[2]	*p*	*ñ*	*e*	*ś*	*o*	*ś*	*aś*	*āś*
j-	*g*	*ṅ*	*ḍ*	*j*	*ñ*	*b*	*ñ*	*e*	*r*	*o*	*r*	*o*	*ā*
t-	*k*	*ṅ*	*ṭ*	*t*	*n*[2]	*p*	*n*	*e*	*s*[4]	*o*	*s*[4]	*as*[4]	*ās*[4]
d(h)-	*g*	*ṅ*	*ḍ*	*d*	*n*	*b*	*n*	*e*	*r*	*o*	*r*	*o*	*ā*
n-	*ṅ*	*ṅ*	*ṇ*	*n*	*n*	*m*	*n*	*e*	*r*	*o*	*r*	*o*	*ā*
p(h)-	*k*	*ṅ*	*ṭ*	*t*	*n*[3]	*p*	*m*	*e*	*ϕ*[4]	*o*	*ϕ*[4]	*aϕ*[4]	*āϕ*[4]
b(h)-	*g*	*ṅ*	*ḍ*	*d*	*n*	*b*	*m*	*e*	*r*	*o*	*r*	*o*	*ā*
m-	*ṅ*	*ṅ*	*ṇ*	*n*	*n*	*m*	*m*	*e*	*r*	*o*	*r*	*o*	*ā*
y-	*g*	*ṅ*	*ḍ*	*d*	*n*[8]	*b*	*ỹ*	*e*	*r*	*o*	*r*	*o*	*ā*
r-	*g*	*ṅ*	*ḍ*	*d*	*n*[8]	*b*	*ṃ*	*e*	∅[5]	*o*	∅[5]	*o*	*ā*
l-	*g*	*ṅ*	*ḍ*	*l*	*l̃*	*b*	*l̃*	*e*	*r*	*o*	*r*	*o*	*ā*
v-	*g*	*ṅ*	*ḍ*	*d*	*n*[8]	*b*	*ṽ*	*e*	*r*	*o*	*r*	*o*	*ā*
ś-	*k*	*ṅ*	*ṭ*	*c*	*ñ*	*p*	*ṃ*	*e*	*ś*	*o*	*ś*[6]	*aś*[6]	*āś*[6]
ṣ-	*k*	*ṅ*	*ṭ*	*t*	*n*	*p*	*ṃ*	*e*	*ṣ*	*o*	*ṣ*	*aṣ*	*āṣ*
s-	*k*	*ṅ*	*ṭ*	*t*	*n*	*p*	*ṃ*	*e*	*s*	*o*	*s*[6]	*as*[6]	*ās*[6]
h-	*g*	*ṅ*	*ḍ*	*d*	*n*[8]	*b*	*ṃ*	*e*	*r*	*o*	*r*	*o*	*ā*
vowel	*g*	*ṅ*[7]	*ḍ*	*d*	*n*[7,8]	*b*	*m*	*a*	*r*	*av*	*r*	*a*	*ā*

Table 2: Sandhi of word-final consonants. Column heads are the finals, row heads are initials. [1]*r* is regularly unchanged within a compound. [2]*ns* may become *ṃś* before *c*, *ṃs* before *t* within clitic group. [3]*ns* rarely becomes *ṃḥ* before *p*. [4]Becomes *ḥ* if the word-initial stop is followed by a sibilant. *s* is often unchanged within phonological phrase. [5]When *r* deletes, preceding short vowel is lengthened. [6]Final *s* is deleted before sibilant plus stop. [7]Nasal is usually doubled after a short vowel. [8]After a long vowel, *ns* behaves like *s*/*ās*, but with nasalization of the vowel. Regularly so before a vowel, sometimes before an approximant.

	Singular	Plural
Nominative	*uśík*	*uśíjas*
Accusative	*uśíjam*	*uśíjas*
Instrumental	**uśíjā*	*uśígbhis*
Dative	**uśíje*	*uśígbhyas*
Ablative	**uśíjas*	*uśígbhyas*
Genitive	*uśíjas*	*uśíjām*
Locative	**uśíji*	**uśíkṣu*

Table 3: The declension of a typical consonant-stem adjective, *uśíj-* 'striving'.