Hiatus Avoidance and Metrification in the Rigveda

DIETER GUNKEL AND KEVIN RYAN

Ludwig-Maximilians-Universität München

University of California, Los Angeles

1. Introduction

Using new corpus resources for Rigvedic poetics, we address various aspects of the poets' treatment of vowel hiatus and pre-vocalic shortening (correption), including their strategies for avoidance of these phenomena in certain contexts. Using observed vs. expected tests, we demonstrate, for one, that hiatus avoidance is correlated with degree of metrical-prosodic juncture. For example, hiatus is actively avoided both at the caesura in trimeter verse and between pādas, but its avoidance is weaker in the latter case. In conducting these tests, we control for a confound (interference) from pre-vocalic shortening, which requires us to address the problem of whether it was optional or obligatory within the hemistich. It turns out that /$\bar{V}_1\#V_2$/ patterns as metrically heavier than /$\breve{V}_1\#V_2$/ in the aggregate, though not as heavy as /$\bar{V}C\#V$/. We then discuss different means by which the poets avoid setting up underlying hiatus junctures, including lexical avoidance (i.e., word selection), morphological avoidance (i.e., allomorph selection), and syntactic avoidance (i.e., word reordering). This last technique, whose existence we support with two different tests, is particularly striking, in that syntax (word order) is shown to be sensitive to phonotactics (marked junctures).

2. The corpora

Our study is based on two electronic texts of the Rigveda. The first is a *padapāṭha*-like text created by Alexander Lubotsky in order to prepare his concordance (1997).[1] This text, which we refer to as the *sāśapāṭha*, serves to approximate the underlying phonological form of the Rigveda.[2] The second text, which approximates its surface form, is derived from van Nooten and Holland's metrically restored Rigveda (1994) via the slightly improved version provided by the Linguistics Research Center of the University of Texas at Austin.[3] We refer to

1 We thank Prof. Lubotsky for generously sharing it with us.

2 We are told that Jost Gippert coined the term *sāśapāṭha*.

3 http://www.utexas.edu/cola/centers/lrc/RV/ (26 March 2011).

Stephanie W. Jamison, H. Craig Melchert, and Brent Vine (eds.). 2011.
Proceedings of the 22nd Annual UCLA Indo-European Conference. Bremen: Hempen. 53–68.

it as the restored text. We aligned the corpora pāda for pāda,[4] converted *ch* to *cch* after short vowels,[5] and marked all tokens of loc. sg. *tvé*, the particle *ū/u*, and simile-marking *ná* as unalterable in sandhi, as well as ca. 75% of nominal duals in *-e*, *-ī*, and *-ū*.[6] In the restored text, we also marked underlyingly pre-vocalic *-ā*, *-ī*, *-ū*, *-e*, and *-o* (from underlying *-o* and *-aḥ*)[7] as short for purposes of scansion.[8] In trimeter verse, unambiguous caesurae were also marked.[9] While both texts can be further improved, these changes bring them significantly closer to phonological/phonetic form.

3. Hiatus: preliminaries

When considering the role that hiatus plays in Rigvedic composition, it is important to distinguish among a number of related phenomena. We refer to the underlying juncture of a word-final and a word-initial vowel, e.g., /bráhma ukthā́/ at 8.33.13d, as *underlying hiatus*. A Rigvedic poet could realize such underlying /V_1#V_2/ junctures in two basic ways. One was to allow the juncture to surface, e.g., /bráhma ukthā́/ → *bráhma ukthā́* at 8.33.13d, resulting in *true surface hiatus*.[10] The other was to resolve the hiatus by the fusion or gliding processes familiar from the classical language, e.g.,

> /ágne pū́rvāḥ án<u>u u</u>ṣásaḥ vibhāvaso/ → *ágne pū́rvā án<u>ū</u>ṣáso vibhāvaso* 1.44.10a,

4 In the few cases where there were discrepancies, the restored text was brought in line with the *sāśapāṭha*.

5 In the restored text we also made the change where a word boundary intervenes, e.g., *pṛthú cchardíḥ* 1.48.15c; on the phonology of *(c)ch* cf. Kobayashi 2004:59–60, 67–81, with references.

6 The rest of this time-consuming task remains to be done. The as yet unmarked *pragṛhya* duals are evenly distributed throughout the Rigveda. On the unalterable forms cf. Oldenberg 1888:455–6n2 and Malzahn 2001, with references. In the restored text, combinations such as /átha u/ → *átho* were also marked as *pragṛhya*. On the metrics of *ná* cf. Vine 1978 and Pinault 1985 and 1997.

7 We use the standard citation forms *-aḥ*, *-e*, and *-o* for the underlying sequences from Indo-Iranian **-as*, **-ai*, **-au*, none of which seems to pattern as a long monophthong in the Rigveda.

8 Cf. Oldenberg 1888:389–93; Malzahn 2001:108–59.

9 By "unambiguous" we mean verses where there is a word boundary after the fourth or fifth metrical position, not both, as well as verses where there are word boundaries after both, but the monosyllable in the fifth position was identifiable as a proclitic or enclitic.

10 This is opposed to *false surface hiatus*, i.e., the (graphic) surface juncture of a word-final and word-initial vowel that does not arise from underlying hiatus, e.g., /sómāḥ áraṃkṛtāḥ/ → *sómā áraṃkṛtāḥ* 1.2.1b.

/parāyatī́m mātáram ánu acaṣṭa/ → *parāyatī́m mātáram ánv acaṣṭa* 4.18.3a.

The poet's choice was guided by potentially conflicting phonotactic, prosodic, and metrical factors.

For example, when a poet is choosing whether to resolve underlying hiatus, phonotactic factors militate against permitting it to surface and drive the phonological processes that repair it. Crucially, these processes affect syllable count and weight as well as the distribution of word boundaries. For example, in /ca imā́ḥ/ → *cemā́ḥ* 1.190.4c, gliding results in the conversion of /L#LH/ → [HH] (where H = a heavy syllable, L = a light syllable, and # marks word boundary). The meter—the poets' target rhythmic patterns—makes strong claims on precisely the same three things, with the result that metrical preferences may (additionally) motivate or block processes like fusion and gliding (cf. /ca imā́n/ → *ca imā́n* 5.50.2b). Finally, metrical-prosodic phrasing must be taken into account. To continue with the same example, fusion and gliding do not apply across the caesura or larger phrasal boundaries. In this study, we set aside hiatus resolution and focus on the phonotactic, metrical, and prosodic factors involved where the poets allow underlying hiatus to surface.[11]

4. True surface hiatus and pre-vocalic shortening

It has long been known that the Rigvedic poets avoid true surface hiatus for phonotactic reasons (cf. Oldenberg 1888:434–40).[12] There are metrical grounds for avoidance that closely interact with the independent, phonotactically motivated operation of pre-vocalic shortening (henceforth PVS) in the Rigveda. Where the poets permit a /V_1#V_2/ juncture to surface, if V_1 is a short vowel (*a*, *i*, *u*), the sequence surfaces without any changes, e.g.,

/śumbhánti áśvarādhasaḥ/ → *śumbhánti áśvarādhasaḥ* 5.10.4b.

---

11 On gliding cf. especially Oldenberg 1915 = 1967:1255–69 and the overviews of the scholarship on the topic in Seebold 1972:60–4 and Malzahn 2001:165–75. On fusion cf. Oldenberg 1888:440–4 and Malzahn 2001:160–4.

12 It has also long been recognized that the poets avoided certain vowel combinations more stringently than others in underlying hiatus. For example, hiatus with a low V_1 is claimed to be more avoided than hiatus with a high V_1, a generalization that our own tests confirm (though we find that both types are actively avoided). A detailed treatment of how the specific phonological features (including vowel height, backness, length, and accent) of V_1 and V_2 affect the stringency of hiatus avoidance in various contexts is beyond the scope of the present paper.

If V_1 is a long vowel (*ā*, *ī*, *ū*), however, scholars agree that it usually (if not always) undergoes PVS, the evidence being metrical not graphic, e.g.,

/ácikradat vŕ̥ṣaṇam pátnī áccha/ → *ácikradad vŕ̥ṣaṇam pátnĭ ácchā* 4.24.8c,
/váreṇyaḥ hótā adhāyi vikṣú/ → *váreṇiyo hótă adhāyi vikṣú* 1.60.4b,

where PVS renders an underlyingly long V_1 more suitable for metrical positions (here 9 and 6 in trimeter verse) where a light syllable is preferred. PVS also has the potential to render an underlyingly long V_1 less suitable for preferentially heavy metrical positions, as in

/śatám ráthebhiḥ subhágā uṣā́ḥ iyám/ 1.48.7c,

where shortening, if it applied, would result in a relatively rare (see table 1 below) implementation of the eighth position of a trimeter verse with a light syllable,

→ *śatáṃ ráthebhiḥ subhágă uṣā́ iyáṃ.*

Scholars disagree on how the poets treated such cases. Some assume a more regular implementation of the meter and a less regular application of PVS (Arnold 1905:134–5, Malzahn 2001:160–4) and read

→ *śatáṃ ráthebhiḥ subhágā uṣā́ iyáṃ*

in these cases. Others believe that PVS was obligatory (except perhaps at the caesura), and that the implementation of the metrical positions involved was less regular (Oldenberg 1888:465–9), reading *subhágă uṣā́*.[13] To sum up the problem, we have two competing theories, a compositional grammar in which PVS is mandatory and another in which PVS can be overridden under metrical pressure.

We can begin to further articulate the role of PVS in the Rigvedic poetic grammar by taking a closer look at the avoidance of placing the first vowel of true surface hiatus junctures [V_1#V_2] in preferentially heavy metrical positions. In order to do this, we first quantify the *weight propensity* of each metrical position in eight-, eleven-, and twelve-syllable verse.[14] Weight propensity is measured in

13 One can get an impression of just how rare this is in Malzahn 2001:160–2, according to whose count obligatory PVS of *-ā* and *-ī* in trimeter verse would yield only seven less-than-regular cadences in the Rigveda.

14 The corpus for these tests consist of all eight-, eleven-, and twelve-syllable verses in the restored text, excluding 761 eleven- and twelve-syllable lines in which the location of the caesura was not immediately clear to us and was left unmarked.

terms of how often each metrical position is filled with a heavy syllable.[15] The shaded positions in table 1 are occupied by light syllables more than 50% of the time.

Table 1. Weight propensity of metrical positions in (a) eight-, (b) eleven-, and (c) twelve-syllable verse

	Metrical positions											
	1	2	3	4	5	6	7	8	9	10	11	12
(a)	53%	77%	67%	79%	8%	93%	5%	82%				
(b)	49%	86%	52%	96%	63%	12%	40%	97%	4%	98%	76%	
(c)	51%	87%	51%	95%	67%	10%	37%	97%	3%	98%	1%	83%

We then tally the percentage of word junctures (starting) in each position that involve true surface hiatus.[16] In graphs given as figures 1–2 below, the dotted line plots the weight propensity of each position and the solid line plots incidence of true surface hiatus (starting) there. Note the clean negative correlation: essentially, the heavier a metrical position is, the less readily the poets fill it with the first vowel of a true surface hiatus juncture, confirming that V_1 is (usually/relatively) short.

One way of approaching the undecided issue of whether PVS was obligatory or not is to compare the way the poets versify true surface hiatus junctures where V_1 is underlyingly long with the way they versify true surface hiatus junctures where V_1 is underlyingly short. If we assume, as a null hypothesis, that PVS is obligatory and results in a total quantitative merger, we then expect true surface hiatus that arises from $/\bar{V}_1\#V_2/$ to be metrically indistinguishable from that which arises from $/\breve{V}_1\#V_2/$, all else being equal (where $\bar{V} = \{\bar{a}, \bar{\imath}, \bar{u}\}$, $\breve{V} = \{a, i, u\}$, and V = any vowel or diphthong). This prediction can be tested as follows. First, we divide the metrical positions that we already weighed into two classes, (preferentially) heavy and (preferentially) light, according to the syllable weight grade occupying them most frequently (see the shading in table 1 above).

15 Note that this method of establishing weight propensity is only an approximation, though sufficient for the present purposes. It does not, for one, take into account context-dependent weight, i.e., cases where the weight of one position is contextually constrained by that of another, e.g., the constraint against implementing both positions 2 and 3 with light syllables. It also ignores word and metrical boundaries, whose placement can interact with weight preferences. For example, the fourth position of the *triṣṭubh* is more uniformly heavy when the caesura falls after the fifth position than when the caesura follows the fourth.

16 Note that having the total junctures as a denominator controls for caesura- and bridge-like tendencies in the verse.

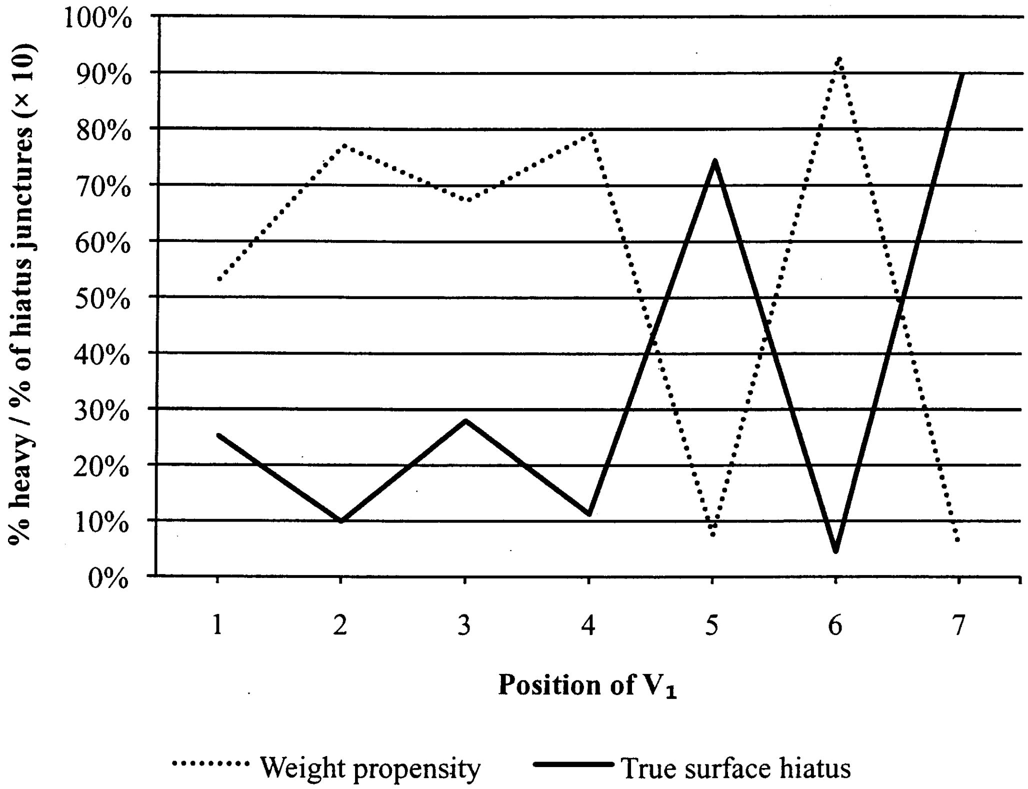

Fig. 1. Position of V_1 in eight-syllable verse

We then tally how often in true surface hiatus junctures

(1) /$\breve{V}_1$/ occupies a light position, e.g.,
/yát kārāve dáśa vṛtrā́ṇi_aprati/ → *yát kārāve dáśa vṛtrā́ṇi_aprati* 1.53.6c

(2) /$\bar{V}_1$/ occupies a light position, e.g.,
/jātó-jātaḥ jāyate vājī́_asya/ → *jātó-jāto jāyate vājī́_asya* 7.90.2d

(3) /$\breve{V}_1$/ occupies a heavy position, e.g.,
/gā́vaḥ dhenávaḥ barhíṣi_ádabdhāḥ/ → *gā́vo dhenávo barhíṣi_ádabdhā* 1.173.1c

(4) /$\bar{V}_1$/ occupies a heavy position, e.g.,
/mitráḥ rā́jānaḥ aryamā́_ápaḥ dhúḥ/ → *mitró rā́jāno aryamā́_ápo dhúḥ* 7.40.4b.

It is necessary to exclude junctures located at the caesurae in trimeter verse, due to the fact that caesura "mitigates" hiatus, as discussed in greater detail below.

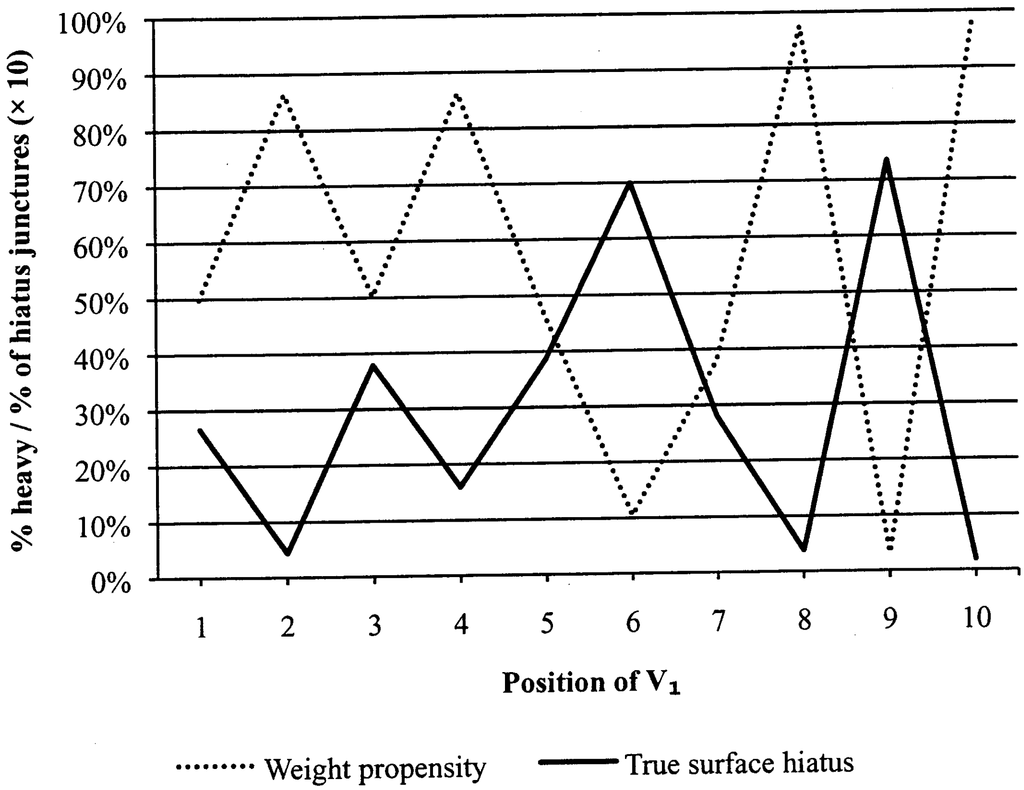

Fig. 2. Position of V_1 in eleven- and twelve-syllable verse

The resulting contingency table is given below. We find that $\bar{V}_1$ is significantly more frequent in preferentially heavy positions.[17]

Table 2. Vowels at hiatus junctures

	$/\bar{V}_1\#V_2/$	$/\breve{V}_1\#V_2/$	$\% /\bar{V}_1\#V_2/$
V_1 in light	1,551	94	5.7%
V_1 in heavy	695	138	16.6%

The greater weight of V_1 in the true surface hiatus junctures from $/\bar{V}_1\#V_2/$ vis-à-vis $/\breve{V}_1\#V_2/$ is further reinforced by the plot in figure 3, which depicts the correlation of the weight propensity of a metrical position (x-axis) and the incidence of true surface hiatus with underlyingly long V_1 (starting) there. The correlation, though far from perfect, is highly significant.[18]

17 Fisher's exact test two-tailed $F(1) = 3.3$, $p < .0001$.

18 Pearson's unweighted $\rho = .64$, weighted $\rho = .60$, both $p < .0001$.

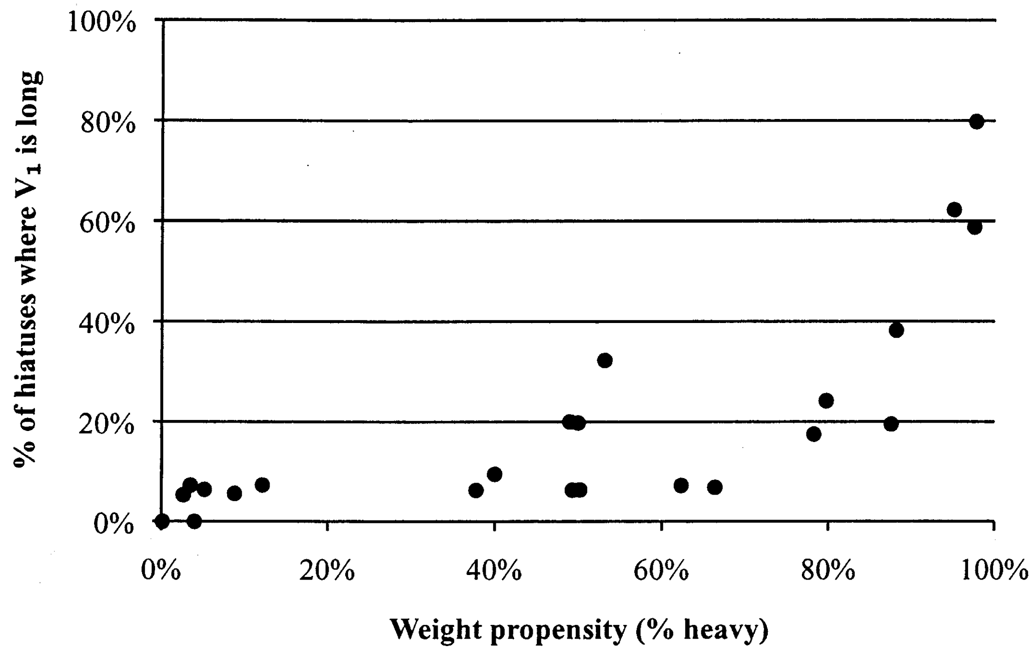

Fig. 3. Correlation of weight propensity and incidence of true surface hiatus with $/\bar{V}_1/$

In view of these data, a poetic grammar in which obligatory PVS results in the indistinguishable metrical treatment of all true surface hiatus junctures is impossible. We can also say for certain that the poets do not shorten an underlyingly long V_1 freely to fit the meter. The heavy:light skew of underlyingly long V_1 is 1.5, which is closer to that of a typical light syllable (e.g., 0.7 for $\underline{\breve{V}}$C#V) than it is to that of a typical heavy syllable (e.g., 9.1 for $\underline{\bar{V}}$C#V), though it is significantly different from both. In other words, in the aggregate, the poets treat underlyingly long V_1 in these junctures as having an intermediate weight between typical heavy and typical light.[19]

This leaves a number of possibilities open. If the poets operated with a grammar in which syllable weight was strictly binary, it is possible to retain the view that PVS was optional but preferred, i.e., underlyingly long V_1 could surface as short or long, but preference was given to the former treatment, resulting in an

19 If we exclude junctures at potential secondary caesurae, i.e., after the eighth position in trimeter verse (Arnold 1905:180) and after the fifth position in dimeter verse (where our own research suggests that word boundary is significantly more frequent than expected) on the assumption that these breaks, like primary caesura, mitigate hiatus, the heavy:light skew of underlyingly long V_1 drops to 1.9, but the distributional difference between underlyingly long and short V_1 remains significant.

aggregate weight closer to that of a typical light syllable. A further possibility, with suggestive parallels in Epic Sanskrit, Homeric Greek, and Kalevala Finnish versification (Ryan 2011), is that the poets were sensitive both to the binary categorical distinction between heavy and light syllables and to gradient weight distinctions within these categories. On this view, it remains possible that PVS was obligatory and resulted in the categorical change from long to short (hence heavy to light), but not in a total merger with underlyingly short V_1; it remained subcategorically longer/heavier, and the poets more readily located it in heavy metrical positions. We leave the problem at this point to further research.

5. Metrical-prosodic phrasing

Hiatus is sensitive to metrical phrasing: caesura ameliorates hiatus (Oldenberg 1888:359), and it is held that there is no avoidance across pādas (Arnold 1905: 71).[20] We assume here that metrical phrase structure is ontologically related to the prosodic phrasing of "ordinary," non-metrical language known as the Prosodic Hierarchy (Selkirk 1980, Nespor and Vogel 2007, among others), and that the poets preferentially align particular prosodic phrase boundaries with particular metrical phrase boundaries.[21] Controlling for the avoidance of true surface hiatus in heavy metrical positions discussed above, we show that the phonotactic constraints that militate against hiatus and motivate various hiatus avoidance tactics are in fact active across larger phrase boundaries. Hiatus avoidance thus provides (additional) positive evidence that the Rigvedic poets composed in units larger than the pāda.

We very briefly sketch the metrical-prosodic structure of eight-, eleven-, and twelve-syllable verse.[22] The basic unit of dimeter verse is an eight-syllable pāda. In the *gāyatrī* variety three such pādas (*a*, *b*, *c*) form a strophe, in the *anuṣṭubh*

---

20 Arnold (l.c.) explicitly states that "[e]ach verse in the Rigveda is an independent metrical unit as regards Sandhi: a final vowel may therefore stand at the end of any verse, whether the next verse begins with a vowel or not, without hiatus arising" and that "hiatus seems to be quite as common at the end of verses *a* and *c* as at the end of verses *b* and *d*." This view goes back to Kühnau 1886:20 contra Westphal 1860:442, and is implicit in Oldenberg 1888:430 and 1907 = 1967:264–5 and later work, e.g., Korn 1998:10. Note also that Lubotsky (1993) argues that *after* the period of composition, the early editors of the Rigveda seem to have avoided particular intra-distich junctures, e.g., that of anudātta -*ā̆* plus udātta *é*- or *ó*- via nasalization of the -*ā̆*.

21 For a discussion of the relationship between the Metrical Hierarchy and the Prosodic Hierarchy cf. Hayes 1989, with references. For preferential alignment of prosodic and metrical phrases cf. Devine and Stephens 1984:128–33.

22 For a convenient overview cf. Korn 1998:9–21.

four (*a*, *b*, *c*, *d*), and in the *paṅkti* five (*a*, *b*, *c*, *d*, *e*). According to Oldenberg (1909 = 1967:1189–90), the structure of the *gāyatrī* strophe above the level of the pāda is [[*ab*]*c*] and that of the anuṣṭubh is [[*ab*][*cd*]], and it may be inferred from the extensional nature of pāda *e* of *paṅkti* verse, which is often a refrain or a variation on *d* (Oldenberg 1888:32), that its strophic structure is [[*ab*][*cd*]*e*]. To generalize slightly, in eight-syllable verse, pādas *a* and *b* form a distich, as do *c* and *d* if *d* is present.

In trimeter verse, where there is a regular caesura after either the fourth or the fifth syllable, the pāda has some internal structure, consisting of two hemistichs that straddle the caesura (ℭ), e.g.,

[[*tá ukṣitā́so*]$_{\text{hemistich}}$ ℭ [*mahimā́nam ā́śata*]$_{\text{hemistich}}$]$_{\text{pāda}}$ 1.85.2a.

The major trimeter varieties, *triṣṭubh* and *jagatī*, exhibit the same strophic structure as dimeter *anuṣṭubh*, i.e., [[*ab*][*cd*]].

6. Hiatus avoidance at metrical-prosodic boundaries

What we are interested in testing is the relative stringency of hiatus avoidance at the metrical-prosodic boundaries within the strophe: word (#), hemistich/caesura (ℭ), intra-distich pāda, i.e., between *a* and *b*, *c* and *d* (Ᵽ), and distich, i.e., between *b* and *c* (Đ). In order to do this, we compare the observed frequency of true surface hiatus at these boundaries with the frequency that would be expected if the poets were not avoiding hiatus at all. The observed frequencies are there for the tallying in the corpora. In order to generate the expected frequencies, we randomly shuffle equivalent metrical-prosodic constituents.

7. Shuffling

One way of modeling a Rigvedic poetic grammar with no hiatus avoidance at the caesura is to randomly re-combine pre- and post-caesural hemistichs that are identical with respect to verse type (*triṣṭubh*, *jagatī*) and the position of the caesura (4ℭ or 5ℭ).[23] For example, the hemistichs of the following *triṣṭubh* verses with caesura 4

[[*sá śévṛdham*] ℭ [*ádhi dhā dyumnám asmé*]] 1.54.11a
[[*sá citréṇa*] ℭ [*cikite ráṃsu bhāsā́*]] 2.4.5c

could be recombined to

23 Hyper- and hyposyllabic verses and verses without a caesura 4ℭ or 5ℭ were excluded from the test.

[[*sá citréṇa*] ℂ [*ádhi dhā dyumnám asmé*]]
[[*sá śévṛdham*] ℂ [*cikite ráṃsu bhāsā́*]].

This yields a metrically licit corpus of nonsensical trimeter verse in which pre- and post-caesural hemistichs are combined without regard to hiatus—a Rigged Veda. We then tally the incidence of hiatus at the caesura in this text to get the values we would expect if the poets did not avoid hiatus there at all.

As discussed above, heavier positions exhibit less hiatus simply because underlyingly long V_1 in hiatus (almost) always scans as light. In order to factor out this motivation for hiatus avoidance and isolate the contribution of hiatus avoidance per se across various metrical-prosodic junctures, hiatus incidence is computed here by taking the number of underlying short vowels in the relevant context standing in true surface hiatus (i.e., $/\breve{V}_1\#V_2/$) and dividing by the incidence of $/\breve{V}_1\#CV/$ in the relevant context. By considering only ratios of short vowels in hiatus to those not in hiatus (nor "making position"), we eliminate the weight preference of the position as a confound: any underrepresentation of $/\breve{V}_1\#V_2/$ relative to $/\breve{V}_1\#CV/$ must be attributed to hiatus avoidance, not light-syllable avoidance.

The shuffling is performed over one hundred iterations and a mean expected value is drawn from this set of Rigged Vedas to insure that expected values are not drawn from an outlier.[24] The observed (O) incidence of hiatus at the caesura is .1149 versus a mean expected (E) value of .2994. The observed-over-expected (O/E) value is .389, which is to say that the poets produced hiatus at the caesura about 39% as often as we would expect if they didn't avoid it at all. Shuffling the remaining prosodic constituents (words and pādas) in precisely the same way, we find the following hierarchy of hiatus avoidance:

word ≫ hemistich ≫ pāda ≫ distich.

The avoidance at word boundary and caesura comes as no surprise. The weak but significant avoidance at (intra-distich) pāda boundary does, however.[25] At the distich boundary, we also observe hiatus less often than we would expect, but not significantly so, given this test.[26] The numbers are given in tables 3–6 below.

24 This is a Monte Carlo method for gauging probability (Metropolis and Ulam 1949, Robert and Casella 2004, Rubinstein and Kroese 2007).

25 Monte Carlo $p < .02$ vs. chance.

26 Monte Carlo $p < .11$ vs. chance.

Table 3. Word

	$/\breve{V}_1\#V_2/$	$/\breve{V}_1\#CV/$	Hiatus incidence
O	1,883	16,925	.111
E	e.g., 4,610	e.g., 14,224	mean .329
O/E			= .337

Table 4. Hemistich

	$/\breve{V}_1\#V_2/$	$/\breve{V}_1\#CV/$	Hiatus incidence
O	490	4,263	.1149
E	e.g., 1,102	e.g., 3,651	mean .2994
O/E			= .389

Table 5. Pāda

	$/\breve{V}_1\#V_2/$	$/\breve{V}_1\#CV/$	Hiatus incidence
O	965	17,165	.056
E	e.g., 1,178	e.g., 19,768	mean .05915
O/E			= .949

Table 6. Distich

	$/\breve{V}_1\#V_2/$	$/\breve{V}_1\#CV/$	Hiatus incidence
O	606	8,778	.06191
E	e.g., 651	e.g., 9,895	mean .06448
O/E			= .96

8. A note on *sá* and *sáḥ*

The distribution of the allomorphs *sá* and *sáḥ* in the Rigveda was treated by Oldenberg on two separate occasions.[27] We find his discussion of the verse-internal distribution completely convincing. Essentially, phonotactic preferences motivate the poets to use *sáḥ* __#V and *sá* __#C, except where metrical factors (syllable count, positional weight preferences) override this distribution in favor of either contraction of *sá* with a following vowel or the placement of *sáḥ* before a consonant.[28] We disagree, however, with his suggestion that we read *sá* verse-finally, against the received text. His logic, in more current linguistic terms, is that *sá* is the elsewhere allomorph, the use of which is phonotactically unmotivated at verse boundary. On the contrary, the choice of *sáḥ* in this position seems well motivated. In light of the evidence for hiatus avoidance across pāda boundary (at least), and assuming that the choice of *sáḥ* was no more costly than *sá*, phonotactic preferences would favor *sáḥ* before vowel-initial verses, at least within the distich, and the strong metrical preference for heavy syllables verse-finally (see the weight propensities above) would favor the same allomorph elsewhere.[29] Though it may be pushing things too far, one might point to junctures like /sáḥ Đ agním/ 3.13.3bc and /sáḥ Đ asmā́kam/ 1.79.11bc as evidence for hia-

27 1888:462–5, 1907 = 1967:264–5.

28 Cf. the occasional use of movable *nu* __#C to "make position" by ancient Greek poets.

29 Regarding the tallies of hiatus incidence at verse boundary above, note that Lubotsky did not follow Oldenberg in this respect in the *sāśapāṭha*.

tus avoidance across distich boundary, which, as we saw above, is greater than chance, but not significantly so in our test.

9. Morphological hiatus avoidance tactics

It is worth considering how exactly the poets avoided hiatus at the hemistich boundary and above. Given that purely phonological processes (gliding, fusion) were not an option, they must have avoided hiatus by morphological, lexical, and/or syntactic means.[30] This is a topic for further research, with potential implications for interactionist models of (poetic) grammar. Morphological avoidance is possible where affix allomorphy provides a form with a consonant and a form with a vowel at the word edge, such as consonant-stem nominals with a locative singular in *-i* or an endingless locative. The distribution of the locatives *yā́mani* and *yā́man* 'course, flight' may be indicative of a larger pattern. The poets locate *yā́mani*, the variant that could potentially enter into a /V_1#V_2/ juncture, before a consonant-initial word or verse-finally 13×, and never before a vowel-initial word, whereas they locate *yā́man* both before consonant- and vowel-initial words, 10× and 20× respectively. The distributions of the two allomorphs are significantly different, but it should be noted that metrical factors must be controlled for. While lexical avoidance is difficult to diagnose given possible nuanced semantic differences between items, some initial testing suggests that hiatus avoidance is achieved in part by ordering. This is of particular interest, since weight effects are well documented in syntax (e.g., Wasow 2002 for natural language, Behaghel 1909 and Kiparsky 1968 for poetic language), but phonotactic factors are less frequently considered.

10. Syntactic hiatus avoidance tactics

In the Rigveda, there are a large number of *swappable bigrams*, i.e., two-word sequences that would have the same metrical template, including boundary location, if the words were reversed (taking sandhi into account), e.g., *devā́n víśvān* HH#HH (2×) and *víśvān devā́n* HH#HH (7×). A subset of these, the *polar swappable bigrams*, would exhibit hiatus in one order but not the other, e.g., *indra daddhi* HL#HL (2×) and a putative unattested ?*daddhi indra*. The vast majority of polar swappable bigrams are in non-hiatus order (89.5% by tokens, 89.1% by types), very significantly different from the 50% we would expect if hiatus were ignored ($p < .001$). Of polar swappable bigrams involving two accented words,

30 On metrically motivated morphological variation cf. Korn 1998.

and thus less likely to involve clitics, there is not a single attestation exhibiting hiatus, excepting cases where caesura intervenes.

Further evidence for syntactic hiatus avoidance comes from pādas that can be scrambled such that the resulting metrical template, accentual pattern, and word boundary locations are the same as in the original (again, taking sandhi into account), e.g.

/agnéḥ índrasya sómasya/ → *agnér índrasya sómasya* 2.8.6a

can be scrambled to a putative ?*agnéḥ sómasya índrasya* with hiatus, while scrambling

/kṣáyam bṛhántam pári bhūṣati dyúbhiḥ/ → *kṣáyam bṛhántam pári bhūṣati dyúbhir* 3.3.2c

to a putative ?*dyúbhir bṛhántam pári bhūṣati kṣáyam* does not result in hiatus. Hiatus is almost twice as frequent in scrambled pādas (4.7% in scrambles vs. 2.7% observed, $p < .001$). These tests suggest that the poets bring about a significant amount of hiatus avoidance by manipulating the word order, which is to say that phonotactic factors partly condition word order in the Rigveda.

11. Conclusion

Vowel hiatus and pre-vocalic long vowels are independently marked phonological configurations. Each can either be repaired with phonological alternations (such as gliding, fusion, shortening, etc.), thus violating faithfulness to the underlying form, or else avoided through the manipulation of the underlying material (through lexical, morphological, or syntactic means). A complete compositional grammar of the Rigveda must gauge the relative extent to which these different strategies are employed in different contexts, a goal that benefits in many ways from computing reasonable expected values and comparing them to the poets' actual versification choices.

References

Arnold, E. Vernon. 1905. *Vedic Metre in Its Historical Development*. Cambridge: Cambridge University Press.

Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25.110–42.

Devine, Andrew M., and Laurence D. Stephens. 1984. *Language and Metre: Resolution, Porson's Bridge, and Their Prosodic Basis* (American Classical Studies 12). Chico, CA: Scholars Press.

Hayes, Bruce. 1989. The Prosodic Hierarchy in Meter. In Paul Kiparsky and Gilbert Youmans (eds.), *Rhythm and Meter*, 201–60. Orlando, FL: Academic Press.

Kiparsky, Paul. 1968. Metrics and Morphophonemics in the *Kalevala*. In Charles E. Gribble (ed.), *Studies Presented to Professor Roman Jakobson by his Students*, 137–48. Cambridge, MA: Slavica.

Kobayashi, Masato. 2004. *Historical Phonology of Old Indo-Aryan Consonants* (Study of Languages and Cultures of Asia and Africa Monograph Series 42). Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Korn, Agnes. 1998. *Metrik und metrische Techniken im Ṛgveda: Streckformen in Trimeter-Versen*. Graz: Leykam.

Kühnau, Richard. 1886. *Die Trishṭubh-Jagatî-Familie: Ihre rhythmische Beschaffenheit und Entwicklung*. Göttingen: Vandenhoeck und Ruprecht.

Lubotsky, Alexander. 1993. Nasalization of the Final *ā̆* in the Ṛgveda. *Indo-Iranian Journal* 36.197–210.

———. 1997. *A Ṛgvedic Word Concordance* (American Oriental Series 82–3). New Haven: Americal Oriental Society.

Malzahn, Melanie. 2001. Sandhiphänomene im Rigveda als Reflexe von Archaismen und Dialektismen. Ph.D. diss., Universität Wien.

Metropolis, Nicholas, and Stanislaw Ulam. 1949. The Monte Carlo Method. *Journal of the American Statistical Association* 44.335–41.

Nespor, Marina, and Irene Vogel. 2007. *Prosodic Phonology: With a New Foreword* (Studies in Generative Grammar 28). Berlin: de Gruyter.

van Nooten, Barend A., and Gary B. Holland (eds.). 1994. *Rig Veda: A Metrically Restored Text with an Introduction and Notes* (Harvard Oriental Series 50). Cambridge, MA: Harvard University Press.

Oldenberg, Hermann. 1888. *Die Hymnen des Ṛgveda* I: *Metrische und textgeschichtliche Prolegomena*. Berlin: Hertz.

———. 1907. *sá* und *sáḥ* im Ṛgveda. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 61.833–5.

———. 1909. Zur Geschichte des Śloka. *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen* 219–46.

———. 1915. Zur Geschichte der Triṣṭubh, mit einem Exkurs: Zur Behandlung des auslautenden *-i* und *-u* im Ṛgveda. *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen* 490–543.

———. 1967 [1987]. *Kleine Schriften*. 2 vols. Wiesbaden: F. Steiner.

Pinault, Georges-Jean. 1985. Négation et comparaison en védique. *Bulletin de la Société de Linguistique de Paris* 80.103–44.

———. 1997. Distribution de la particule négative *ná* dans la Ṛk-Saṁhitā. *Bulletin d'études indiennes* 15.213–46.

Robert, Christian. P., and George Casella. 2004. *Monte Carlo Statistical Methods*². New York: Springer.

Rubinstein, Reuven Y., and Dirk P. Kroese. 2007. *Simulation and the Monte Carlo Method*[2]. New York: Wiley.

Ryan, Kevin M. 2011. Gradient Weight in Phonology. Ph.D. diss., University of California, Los Angeles.

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

Selkirk, Elisabeth. 1980. Prosodic Domains in Phonology: Sanskrit Revisited. In Mark Aronoff and Mary-Louise Kean (eds.), *Juncture: A Collection of Original Papers* (Studia Linguistica et Philologica 7), 107–29. Saratoga, CA: Anma Libri.

Vine, Brent. 1978. On the Metrics and Origin of Rig-Vedic *ná* 'like, as'. *Indo-Iranian Journal* 20.171–93.

Wasow, Thomas. 2002. *Postverbal Behavior*. Stanford, CA: CSLI Publications.

Westphal, Rudolf. 1860. Zur vergleichenden Metrik der indogermanischen Völker. *Zeitschrift für vergleichende Sprachforschung* 9.437–58.