6-24-2015

Investigating Rigvedic word order in metrically neutral contexts

Dieter Gunkel
University of Richmond, dgunkel@richmond.edu

Kevin Ryan

Investigating Rigvedic word order in metrically neutral contexts

Dieter Gunkel (LMU München)
Kevin Ryan (Harvard University)

Institut für Sprachwissenschaft Universität Wien, 24 June, 2015

1. Swappable bigrams

1.1 We extract all swappable bigrams from the RV

- Both orders are metrically equivalent

e.g. *dháne hité* (7x) ~ *hité dháne* (2x) "when the stake is set"

- Regardless of whether both orders are attested

e.g. *śárma ya$_c$cha* (8x) ~ *ya$_c$cha śárma* (0x) "extend shelter"

1.2 Swappables comprise a laboratory in which non-metrical effects on word order (e.g. syntax) can be investigated in a verse corpus without any possible interference from meter.

1.3 Operationalizing metrical equivalency for the present purposes:

- Same weight template in both orders (in the context of the line)
    - *apsú antár* | *amṛ́tam apsú bheṣajám* [1.23.19a]

      Within the waters is deathlessness, in the waters healing
      (trans. Jamison & Brereton here and throughout)
- Same incidence of hiatus in both orders (see previous ex.)
- Swapping does not introduce a violation of caesura
    - cf. *kakṣ́ívantaṃ* | *stotā́raṃ yā́bhir ā́vataṃ*

      with which you helped Kakṣīvant the praiser
- We exclude bigrams with prevocalic -Vn# in either order
- We exclude identical repeated words
    - e.g. *cākánanta cākánanta*

1.4 We also put aside bigrams in which one or both words is a monosyllable

1.5 Corpus is the RV, excluding certain minor meters/subcorpora (vālakhilya, decasyllables, etc.)

1.6 This leaves 3,113 swappable tokens (2,507 types = unordered word-pairs, e.g. {*dháne*, *hité*})

1.7 Of these, we've so far annotated 637 for syntax (all types with >2 tokens and some with 2)

2. Non-phonological factors in word order

As you would expect, syntax plays a significant role in determining word order within our bigrams, which are small meter-free zones, like tiny pockets of very ancient hieratic prose, in a sense.

2.1 Complementizers come first 97% (N = 32 tokens)

2.1.1 [(...) Complementizer Verb]

1.127.4a dr̥ḷhā́ cid asmā ánu dur [**yáthā vidé**]

Even the firm things give way to him, **as is known**.

4.54.1d [śréṣṭhaṃ no átra dráviṇaṃ **yáthā dádhat**]

[**so that he will establish** the best material wealth for us here]

2.1.2 [(...) Complementizer Noun/Adjective (...)]

1.112.5c [**yā́bhiḥ káṇvam** prá síṣāsantam ā́vataṃ]
tā́bhir ū ṣú ūtíbhir aśvinā́ gatam

[**with which** you helped **Kaṇva** desiring gain] —
with those forms of help come here, o Aśvins.

7.104.21 abhī́d u śakráḥ [paraśúr **yáthā vánam**]

The able one, splitting them [**like** an ax **a tree**]

10.16.1cd [**yadā́ śr̥táṃ** kr̥ṇávo] jātavedo | áthem enam prá hiṇutāt pitŕ̥bhyaḥ

[**When** you will make him **cooked to readiness**], Jātavedas,
then impel him forth to the forefathers.

2.1.3 A complementizer follows once (3%).

2.1.4 [Adjective Complementizer ...]

10.16.2ab [**śr̥táṃ yadā́** kárasi] jātavedo | áthem enam pári dattāt pitŕ̥bhyaḥ

[**When** you will have made him **cooked to readiness**], Jātavedas,
then deliver him to the forefathers.

2.2 Swappable {Object, Verb}

2.2.1 Non-imperative finite Verbs display OV order 78% (N = 37)

5.15.1ab prá vedháse kaváye védiyāya | [**gíram bhare**] yaśáse pūrviyā́ya

I bring forth my song to the ritual master and poet worth finding, to the glorious ancient one.

9.111.2fg tridhā́tubhir áruṣībhir [**váyo dadhe**] | rócamāno [**váyo dadhe**]

He has acquired vitality through the tripartite ruddy (cows) — shining, he has acquired vitality.

6.48.15de āvír [gūḷhā́ **vásū karat**] | suvédā no [**vásū karat**]

[**He will make** the hidden **goods**] visible;
he will make goods easy for us to find.

1.167.9ab nahī́ nú vo maruto ánti asmé | ārā́ttāc cic [chávaso **ántam āpúḥ**]

For they have not [**reached the limit** of your swelling strength], o Maruts, nearby to us nor even in the distance.

2.2.2 VO order: 22%

1.130.7a [**bhinát púro** navatím] indra pūráve

You, Indra, **split the** ninety **strongholds** for Pūru

2.2.3 Imperatives display OV order 77% (N = 22)

1.102.3d tvāyádbhiyo maghavañ [**chárma yaccha**] naḥ

Extend shelter to us who seek you, bounteous one.

7.5.9c váiśvānara [máhi naḥ **śárma yaccha**]

O Vaiśvānara ... , [**offer** great **protection** to us]

8.48.8ab sóma rājan mṛḷáyā naḥ suastí | táva smasi vratíyās [**tásya viddhi**]

King Soma, be merciful to us with well-being.
We are under your commandment: **know this**.

2.2.4 VO order: 23%

3.39.1d índra yát te jā́yate [**viddhí tásya**]

Indra — what is born for you, **know that**.

2.2.5 Finding: OV is more frequent than VO among imperatives in the Rigveda (confirming surmisal in Hock 2012). Note that imperative-initial orders such as

1.184.5c **yātáṃ vartís** tánayāya tmáne ca

Drive your circuit for the sake of our descendants and ourselves

may be V-initial for the same reasons that thetic sentences are often V-initial (cf. Devine and Stephens 2006: 150).

2.2.6 Participles are less clearly final, but the sample is small, with only 4 types. They display OV order 40% (N = 10)

6.45.2 avipré cid [**váyo dádhad**] | anāśúnā cid árvatā | índro jétā hitáṃ dhánam

Placing vitality even in the uninspired, even with a steed lacking speed, Indra is the winner of the stake that is set.

2.2.7 VO order 60%

6.20.9b [**bíbhrad vájraṃ** vṛtraháṇaṃ] gábhastau

[**bearing** his Vṛtra-smashing **mace**] in his fist.

2.3 Partly motivating VO order

2.3.1 In 13/19 tokens of VO order, the object is a complex NP that would be syntactically discontinuous in OV order. Compare

bhinát [**púro** navatím]	vs.	*[**púro**...] **bhinan** [...navatím]
you.split [strongholds 90]		*[strongholds...] you.split [...90]
bíbhrad [**vájraṃ** vṛtraháṇaṃ]	vs.	*[**vájraṃ**...] **bíbhrad** [...vṛtraháṇaṃ]
carrying [mace Vṛtra-smashing]		*[mace...] carrying [...Vṛtra-smashing].

2.3.2 This suggests that the poets may be "swapping" OV order to VO in order to avoid syntactically discontinuous noun phrases. Consider the following scenario:

- Head-final syntax would normally give [[navatím púro] bhinat].
- *Jagatī* meter demands that words shaped like *navatím* are localized immediately after the caesura — 95% are found there — narrowing down the orders to

  *[**púro bhinan** navatím] indra pūráve

  [**bhinát púro** navatím] indra pūráve.

- The latter is chosen (at the cost of verb fronting/swapping) to avoid a discontinuous *púro ... navatím 'strongholds ... 90'.

2.3.3 Among swappable Object & Verb bigrams, continuity is at stake 28x.

- Continuous as attested, discontinuous if swapped (as above): 25x.
- Discontinuous as attested, continuous as swapped: 3x
- Not exceptionless, perhaps due in part to formulaic analogy, where the basic formulae would be

  śárma yaccha{tu, tam, ta, ti} (21x)

  and the OV order would perhaps be retained by analogy in

  1.114.10d ádhā ca naḥ [**śárma yaccha** dvibárhāḥ]

  and then [**extend** doublelofty **shelter**] to us.

2.3.4 Metrical factors plus avoidance of NP discontinuity may partly motivate VO order.

2.4 Swappable {Modifier, Noun}

2.4.1 Adjective and Nouns display AN order 80% (N = 216).

3.18.4b [**bṛhád váyaḥ**] śaśamānéṣu dhehi

Set **rising vigor** upon those toiling

7.36.3c [**mahó diváḥ**] sádane jā́yamāno

Being born in the seat of **great heaven**

2.35.11 tád asyā́nīkam utá [**cā́ru nā́ma** | apīcíyaṃ] vardhate náptur apā́m

This his face and his [**dear** secret **name**] grow strong

6.13.6c [**víśvābhir gīrbhír**] abhí pūrtím aśyām

With **all my hymns** may I attain to fulfillment

2.4.2 NA order

9.20.2ab sá hí ṣmā jaritṛ́bhya ā́ | [**vā́jaṃ gómantam**] ínvati

For he sends to the singers **a prize of cattle**

2.4.3 Genitives and Nouns display GN order 92% (N = 48)

8.5.11a vāvṛdhānā́ [**śubhas patī**]

Having grown strong, o **lords of beauty**

8.51.6b sá [**rāyás póṣam**] aśnute

that one achieves the **thriving of wealth**

9.33.5c marmṛjyánte [**diváḥ śíśum**]

They keep grooming the **child of heaven**.

3.30.18c [**rāyó vantā́ro** bṛhatáḥ] siyāma

Might we be [**winners of** lofty **wealth**].

2.4.4 Noun Genitive order 4x (in 1 type)

9.86.33a rā́jā síndhūnām pavate [**pátir divá**]

The king of rivers, **the lord of heaven** purifies himself.

2.5 Partly motivating Noun Modifier order

2.5.1 Asymmetric hiatus potential: bigrams with hiatus potential seem to prefer their hiatus-avoiding order even when hiatus isn't at stake in the local context (on which more below). Compare

- No hiatus potential

  /mandrám hótāram/ (5x) AN

- Symmetric hiatus potential

  /cā́ru nā́ma/ (6x) AN

- Asymmetric hiatus potential

  /hótā mandráḥ/ (10x) NA, hiatus-protected order (V#...C#)

  /mandráḥ hótā/ (3x) AN, hiatus-exposed order (C#...V#)

2.5.2 This suggests that the poets may have sometimes adopted NA order (as in /hótā mandráḥ/) to avoid hiatus-exposure. If so, we expect to find a higher incidence of hiatus-protection among NA than AN — which we do ($p < 0.002$).

	Protected	Exposed
AN	dáśa kṣípaḥ (6x) máhi śrávaḥ (4x) yúvā kavíḥ (3x) madacyútā hárī (sic 2x)	mandró hótā (3x) śiváḥ sákhā (4x) satyáḥ sátvā (2x)
	Total: 15	**Total: 9**
NA	hótā mandráḥ (10x) vṛ́ṣā háriḥ (6x) sákhā śiváḥ (1x) mádhu priyám (4x) vṛ́ṣā sutáḥ (2x)	—
	Total: 23	**Total: 0**

2.5.3 This would reflect a trade-off between syntactic and phonological markedness, where the poets opt for marked syntax (head-initial NA) instead of marked phonology (vowel hiatus).

2.5.4 Inflection of formulae/formulaic analogy may also result in otherwise unmotivated looking NA order, e.g.

- Base

  vā́jasya gómataḥ 7x verse-finally in dimeter, non-swappable NA

- Inflection/analogy

  vā́jaṃ gómantam 3x in dimeter, swappable NA

2.6 In sum, we observe a general tendency for head-finality, perhaps stronger in some syntactic environments than others. In our current data set

Genitive Noun > Adjective Noun > Object Verb > Object Participle.

Complementizers are expectedly exempt from head-finality and very strongly initial.

3. Phonological factors in word order

3.1 To what extent is word order in swappables predictable from pure phonology?

3.2 Independent of syntax, semantics, meter, etc., phonology is known to affect ordering decisions

- e.g. hiatus avoidance (Gunkel & Ryan 2011)

- or more generally:
    - word size effects ("end-weight")
    - phonotactics (marked configurations are avoided; e.g. Shih & Zuraw 2014)
    - rhythmic optimization (beyond meter, e.g. clash avoidance; Shih 2014)

3.3 Methodology

3.3.1 Multifactorial model. When multiple factors influence a decision, testing each independently can be misleading, since factors might interact or be correlated with each other.

3.3.2 Consider two widely attested principles of ordering:

- Shorter items tend to go first
- More frequent items tend to go first

3.3.3 Examining each in isolation, both appear to be significant for Vedic swappables.

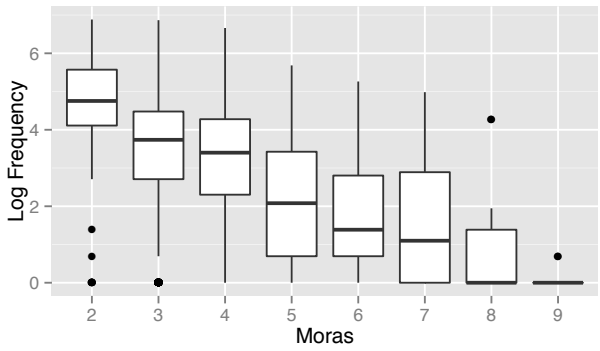
- Logistic model with only the former:

(Model A)	Estimate	Std. Error	z value	Pr(>|z|)
LongLast	0.15298	0.05419	2.823	0.00476 **

- Logistic model with only the latter:

(Model B)	Estimate	Std. Error	z value	Pr(>|z|)
FreqFirst	0.07969	0.01734	4.595	0.00000433 ***

- But, as we'll see below, only end-weight ("LongLast") turns out to be significant in the composite model with both frequency and end-weight.
- The issue here is that word frequency and word size are correlated. Thus, it would be premature to conclude that both principles are active merely because they both test positive individually.



3.3.4 Technical details

- (Binomial) logistic regression
- Mixed effects implementation, with random intercepts for word-pair identity
- Dummy response variable with no intercept (Lohmann & Takada 2014: 55, Levy in press §6.8.4)

3.4 <u>End-weight</u>

3.4.1 cf. Behaghel's (1909) fourth law (Gesetz der wachsenden Glieder)

3.4.2 cf. Pāṇini 2.2.34: In dvandvas, items with fewer vowels (i.e. syllables) go first.

- Kātyāyana adds: vowel length matters too; hence Wackernagel (AiGr §71b):

  "Bei gleicher Silbenzahl geht das Wort mit geringerer Morenzahl voran."

3.4.3 In bigrams, do we also see a heavy-final tendency? And if so, how is word weight best assessed?

- Syllable count
- Mora count (V = 1, VV = 2, coda = 1)
- Vowel-weighted mora count (V = 1, VV = 2, coda = .5)

3.4.4 Each metric is embedded in a model that controls for other significant factors, e.g.

	Estimate	Std. Error	z value	Pr(>|z|)	
LongLast	0.17268	0.06906	2.501	0.01240	*
HiatusPotential	0.84247	0.10086	8.353	< 2e-16	***
HeadFinal	1.70914	0.19946	8.569	< 2e-16	***
CompInitial	3.45785	1.08053	3.200	0.00137	**

3.4.5 Vowel-weighted moras perform best (ΔAIC +6 vs. syllables, +4 vs. moras)

3.4.6 End-weight is evidently not a side-effect of a "frequent-first" tendency.

- Model with both LongLast and FreqFirst (baseline)
- With only FreqFirst, ΔAIC -4.5 vs. baseline (significantly worse)
- With only LongLast, ΔAIC -1.1 vs. baseline (only marginally worse)

3.4.7 Thus, we'll ignore frequency from now on, and assess end-weight in terms of vowel-weighted moras.

	Estimate	Std. Error	z value	Pr(>|z|)	
LongLastHalf	0.1374	0.0401	3.426	0.000612	***
HiatusPotential	0.8568	0.1012	8.469	< 2e-16	***
HeadFinal	1.6778	0.1995	8.412	< 2e-16	***
CompInitial	3.4617	1.0810	3.202	0.001363	**

3.4.8 As one might expect, end-weight becomes a stronger tendency as the size discrepancy between items increases.

3.5 Hiatus potential

3.5.1 We determined swappability in the context of the pāda, e.g. *gúhā hitám* "placed in hiding" is considered swappable in

(a) *gúhā hitáṃ jánima némam údyatam* [9.68.5d]

the (one) birth deposited in secret and the other held aloft

in which both orders are symmetrical with respect to hiatus, but not in

(b) *gúhā hitám úpa niṇíg vadanti* [4.5.8b]

They confide privately what was deposited in secret [=the track/word]"

in which only one order introduces hiatus (or the possibility of fusion).

3.5.2 Thus, we'll say that bigrams like *gúhā hitám* have hiatus potential regardless of their local context.

3.5.3 Finding: bigrams with hiatus potential seem to prefer their hiatus-avoiding order even when hiatus isn't at stake in the local context.

	Estimate	Std. Error	z value	Pr(>|z|)	
LongLastHalf	0.13785	0.04007	3.440	0.000582	***
HiatusPotential	0.85617	0.10106	8.472	< 2e-16	***
HeadFinal	1.67639	0.19919	8.416	< 2e-16	***
CompInitial	3.46045	1.08062	3.202	0.001363	**

3.5.4 Thus, regardless of weight, V#...C# (hiatus-protected) is preferred over C#...V# (hiatus-exposed)

3.5.5 Possibly due to a kind of formulaic analogy (opt for word orders that permit widest use)

3.5.6 Evidently not due to avoidance of C#C (as in **hitáṃ gúhā*)

- "CCPotential" was tested as a factor and was not significant.

3.5.7 Puzzle: avoidance of *potential* hiatus seems only to work in the forward-looking direction (right edge).

- Consider e.g. *áheḥ svanā́d* "from the sound of the serpent." This has left-edge hiatus potential in one order (the attested order) but not in the other. Left-edge hiatus orders are evidently not avoided, though this could be an artifact of our data-set: Most of our bigrams are line-initial.

3.6 Pseudo-geminate avoidance?

- Avoid orders with identical C_1 and C_2 in -C_1#C_2- (correcting for sandhi, e.g. /t#d/ counts as geminate)
- Not a significant factor, but possibly only due to the paucity of relevant forms
- If visarga is treated as sibilant (and counted as geminate in conjunction with any following sibilant), the factor remains non-significant.

3.7 Thus, beyond syntax, we find evidence for two phonological effects on word order in our bigrams: end-weight and hiatus protection. Both are, however, rather weak compared to syntactic considerations.

Works cited

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

Devine, A. M. and Laurence D. Stephens. 2006. *Latin word order. Structured meaning and information*. Oxford: Oxford University Press.

Gunkel, Dieter and Kevin M. Ryan. 2011. Hiatus avoidance and metrification in the Rigveda. *Proceedings of the 23rd Annual UCLA Indo-European Conference*, ed. by Stephanie W. Jamison, H. Craig Melchert, and Brent Vine, 53–68. Bremen: Ute Hempen Verlag.

Hock, H. H. 2012. Come and get it: The Indo-European background of the Vedic éta … stávāma construction. *Proceedings of the 24th Annual UCLA Indo-European Conference*, ed. by Stephanie W. Jamison, H. Craig Melchert, and Brent Vine, 47–66. Bremen: Ute Hempen Verlag.

Jamison, Stephanie W. and Joel P. Brereton. 2014. *The Rigveda*. Oxford: Oxford University Press.

Shih, Stephanie S. and Kie Zuraw. 2014. Phonological factors in Tagalog adjective-noun word order variation. Paper presented at the 88th Annual Meeting of the Linguistic Society of America.

Shih, Stephanie S. 2014. Towards optimal rhythm. Ph.D. diss., Stanford University.

Wackernagel, Jakob. 1905. *Altindische Grammatik*, volume II, 1. Göttingen: Vandenhoek & Ruprecht.