Dvandvas, blocking, and the associative: the bumpy ride from phrase to word

Paul Kiparsky
Stanford University
Linguistics Department
Stanford, CA. 94306
kiparsky@stanford.edu

Dvandvas, blocking, and the associative: the bumpy ride from phrase to word

The oldest form of Sanskrit has a class of expressions which are in some respects like asyndetically co-ordinated syntactic phrases, in other respects like single compound words. I propose to resolve the conflicting evidence by drawing on Prosodic Phonology, Stratal Optimality Theory, and the lexicalist approach to morphological blocking. I then present an account of their semantic properties and of their historical development. The analysis points to a solution to the theoretical problem of non-monotonic trajectories in diachrony, a challenge for causal theories of change which claim that analogical processes are simplifying or regularizing. The idea is that optimization of such a highly structured object as a language does not proceed monotonically, but via a sequence of local optima.*

*This paper was presented at the 25th East Coast Indo-European Conference, June 2007. I gratefully acknowledge the comments and suggestions of Hans Heinrich Hock, Stanley Insler, Stephanie Jamison, Calvert Watkins, Luc Baronian, three reviewers, and two editors.

1 Vedic dvandvas

1.1 The synchronic puzzle

Sanskrit nominal compounds of all types are normally formed by combining bare nominal stems (sometimes with special stem-forming endings) into a compound stem, which bears exactly one lexical accent. The *dvandva* compounds of its earliest surviving literature diverge from this pattern. Each of their two constituents has a separate word accent and what looks like a dual case ending. These compounds (also known as copulative compounds, co-ordinating compounds, or co-compounds[1]) are invariably definite, have two members, and refer to conventionally associated pairs of divine or human beings, or to pairs of personified natural and ritual objects.[2]

(1) *parjány-ā-vā́t-ā* 'Rain and Wind', *mitrā́váruṇau, -ā* 'Mitra and Varuṇa', *mā́tār-āpitār-ā* 'father and mother', 'parents', *dyā́v-āpṛthivī́* 'Heaven and Earth', *indrā́viṣṇū́* 'Indra and Viṣṇu'

The type rapidly recedes after the oldest stage of the language recorded in the Rigveda, and is completely superseded in Classical Sanskrit by a highly productive regularized type of dvandva compound which, like other compounds, has a single accent,[3] no internal case ending, a singular, dual, or plural case ending on the final member depending on the cardinality of the compound's denotation, can have more than two members, and can be indefinite and inanimate.

(2) *vātavarṣáḥ* (Sg.) 'wind-and-rain' (collective), *mitrabandhu(-hīna-)* '(without) friends and relatives', *mātāpitarau* (Dual) 'father and mother', 'parents', *dyurātram* (Sg.) 'a day-and-night', *devamanuṣyāḥ* (Pl.) 'gods and humans'

Vedic dvandvas originated as combinations of two dual-inflected nouns.[4] Synchronically, however, they consist of nominal stems, like other Sanskrit compounds. The first member carries a stem-forming suffix (*-ā* in (1)). The final suffix is just the normal case/number inflection assigned to the whole compound stem.

(3) [[[mitrá-ā]$_{\text{NStem}}$ [váruṇa]$_{\text{NStem}}$]$_{\text{NStem}}$ -au]$_{\text{N}}$

This is how Pāṇini's grammar treats them (6.3.26 ff.), for two good reasons. First, the case ending of the second member cannot be part of the compound because it does not appear before derivational suffixes: 'accompanied by Mitra und Varuna' is *mitrā́váruṇa-vat-* (8.35.13), not

**mitrā́váruṇau-vat*. Secondly, the ending of the first member is not a case ending, for it is invariant, and independent of the case assigned to the compound itself: the Instrumental of *mitrā́váruṇau* 'Mitra and Varuṇa', is *mitrā́váruṇābhyām* (5.51.9), where the Instr.Du. ending *-bhyām* appears only once, on the whole compound (not **mitrā́bhyāṃváruṇābhyāṃ*, as would be expected if it were a case ending).

The Pāṇinian analysis (3) does raise the question what the *-ā* on the first member is. If it is not a dual case ending, what is it? I will argue that it is an morpheme, related to the dual case ending etymologically rather than synchronically, and that both are descended from the Indo-European instrumental ending *-h_1* in its comitative/sociative function. Anticipating that analysis, but without undue prejudice, in the examples cited below I gloss the associative dual ending of the first member of the compound as a dual without case. Thus '- ' glosses the compositional morpheme which I claim is an associative dual, while '- . ', '- . ' etc. gloss ordinary dual case/number endings.

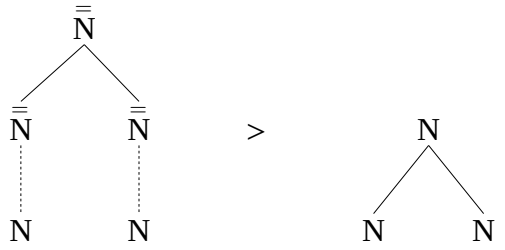
The Vedic dvandvas are synchronically problematic in other respects as well. They are a hybrid construction, patterning in some ways like syntactic phrases built from separate words by asyndetic co-ordination, in others like single compound words. I survey these seemingly contradictory properties in sections 2.1 and 2.2. In 2.3 I outline a theoretically consistent analysis that draws on Prosodic Phonology (Inkelas & Zec 1990), on Stratal Optimality Theory, and on the lexicalist approach to morphological blocking (Wunderlich 1996, Kiparsky 2005). The details of how these ideas jointly resolve the conflicting evidence are set out in 2.4–2.7. Section 2.8 concludes my synchronic analysis of dvandva compounds with a semantic account of how they come to denote pairs of objects.

1.2 The diachronic puzzle

In section 3 I show that the proposed synchronic analysis of Vedic dvandvas helps solve a theoretical problem raised by their diachronic development. Notwithstanding their archaic status in Vedic, they are probably an Indic innovation, with no exact counterpart even in Iranian.[5] With their mix of syntactic and lexical properties they constitute a systematic intermediate stage in the evolution from phrases to words. They originated in Indo-European conjoined phrases that had become formally opaque due to loss of the associative dual category. To a bird's eye, the change from asyndetic phrasal conjunction to compounding looks unproblematically like a straightforward reduction in unmotivated structural complexity. Since the first part denotes a

single individual, its ending no longer makes any contribution to the meaning of the conjoined phrase at the Vedic stage. Moreover, since dvandvas refer to specific pairs of individuals, they are semantically name-like, which makes their phrasal status superfluous. Their reduction to ordinary compound words eliminates this extra morphological and syntactic baggage.

(4)



This rather natural causal explanation raises some key questions about language change. Dvandvas of the intermediate Vedic type represent a wholly new kind of construction, indeed one that anomalous within Vedic morphology in certain respects. Because there is no pre-existing model for them, their rise cannot be explained by ordinary proportional analogical change. It is in fact a case of lexicalization, more specifically of the type referred to as , the merger of two or more words into one. Like grammaticalization, it instantiates a more general cross-linguistically well-documented tendency to tighter grammatical bonding. Such changes have been claimed to be a type of non-exemplar-driven analogical change, characterizable like ordinary analogical change as simplification, or optimization (van Gelderen 2004, Kiparsky in press). But zooming in on early Vedic reveals the construction in mid-change at a stage where it is an intricate but systematic blend of co-ordination and compounding. In fact, the historical record shows that (4) telescopes four steps, which Wackernagel (1905:149) summarized as follows:

(5) a. Co-ordinated nouns, usually dual, merge inseparably, revealing their coalescence by certain [phonological] phenomena, while still retaining the independent accentuation of both members.

b. The first member freezes into a fixed form.

c. The accent becomes restricted to the final stem syllable of the second member.

d. The first member receives the stem form.

The theoretical puzzle is why the simplificatory change from a somewhat anomalous type of phrasal co-ordination to wholly unremarkable compounding should traverse the path in (5), going

through several systematic intermediate stages even more complicated than the starting point, until settling down in its final state. Such non-monotonic trajectories pose a challenge for causal theories of change, in particular for the proposal that analogical changes in the broadest sense are simplificatory or regularizing processes. A point-by-point reconstruction of these changes on the basis of our morphological and semantic analysis accounts for the order in which they occurred, and shows how a simplificatory process may pass through a partial or incomplete stage where it actually complicates the grammar.

Such paradoxical non-monotonic optimization trajectories are actually quite common. They can reflect the competition between an old and a new grammar in one of two ways. Some involve partly arbitrary patterns of variation between an original and an innovative norm. Others — the dvandvas among them — involve hybrid structures which have some features of both. In both cases, the result is typically a 'bump' in the path, or even several, schematically like this:

(6)

In a scenario familiar from morphophonological analogy, a rule is simplified via one or more intermediate stages in which it coexists with the old one. The new form of the rule tends to apply in more productive formations (e.g. in inflection) while the old one continues to apply in less productive formations (e.g. in derivation). The rule's 'mitosis' (Kaye 1978) brings about a temporary increase in complexity (a bump in (6)). Such cases call into question the idea that analogical change and grammaticalization optimize the grammar globally by some overall measure at each step, and suggest that local optimization with respect to some structural features can be introduced even at the price of complicating the system as a whole.

A mundane parable may be useful here. Suppose the palaces of Versailles need a paint job. A painter is hired to do it. He estimates it will take him 30 years. When will Versailles be 'improved'? Certainly once the whole job is finished. Certainly not after the first few brushstrokes — they will look like graffiti. Probably after the first palace is completely painted.

Possibly after the facade of the first palace is completely painted. And certainly not after the first floors of some of the facades have been painted.

Analogical change is like that slow housepainter. Optimization of a complex highly structured object does not necessarily proceed monotonically. Language has many local optima because it is highly structured. Complexity may be eliminated step by stem via a series of local optimizations that affect particular structural features, even at the price of complicating the overall structure. Linguistic theory already furnishes tools for modeling such 'bumpy' trajectories. They include the distinction between morphological and phonological words and between lexical and postlexical strata in grammar, and the factorization of structure by ranked constrints provided by Optimality Theory.

To make meaningful claims about relative grammatical complexity, we need a way to quantify it. We don't have a generally accepted one for now, and it remains to be seen whether one can be found. Classical generative grammar's formal reduction of complexity to length is problematic in many ways, and it does not work at all in a canonical OT framework, where grammars differ only in the ranking of universal constraints. In any case, it is not clear that there is any empirical need for a global evaluation measure that evaluates the complexity of an entire grammar.[6] For now, what we can say with confidence is that there are at least two known aspects of linguistic complexity that do govern the progress of acquisition and change.

The first is the amount of *idiosyncratic information* that a given lexeme, rule, process, paradigm, or other element imposes upon the grammar of a language (understood here as including its lexicon). Other things being equal, the more idiosyncratic an element is, the later it will be learned and the earlier it will be lost or reanalyzed. How the idiosyncratic information is specified — whether in lexical entries through listing or feature specifications, by language-specific rules or constraints, by stipulated rule orderings or constraint rankings, or by a combination of these devices — is a theory-dependent question which we have to leave in its unresolved state here. For our purposes it suffices that virtually any framework of grammatical description provides *some* way of representing idiosyncrasy, from which its informational overhead can be computed.

There is also a fairly clear sense in which we can compare the relative structural complexity of *output representations*. Such a comparison is presupposed by Hawkins' (2004) M F principle. Formal asymmetries in the Minimalist framework, such as the preference of

M over M proposed by Roberts and Roussou (2003) and the H P and L M principles of van Gelderen (2004), may turn out to be reducible to considerations of relative complexity as well, as these authors themselves suggest (see further Roberts 2007, Ch. 3, for a review of the issues and related ideas).

The two kinds of relative complexity, of grammars and of outputs, are obviously related, and for the lexicon they may largely coincide; Anttila MS. argues specifically that the complexity of a lexical item correlates with the amount of language-specific ranking information it requires. On the other hand, neither is fully reducible to the other. Complex structures may be generated by simple grammars, and vice versa. For example, simple grammars (e.g. in an OT syntax, high-ranking F) can produce complex syntactic structures that mirror the semantics closely.

Both types of complexity are certainly needed to account for the direction of analogical change, and of the orderly growth of grammar in first language acquisition (Estigarribia 2007). In section 3 I sketch out how they may be unified in a constraint-based theory of grammar, and argue that paths like (6), including in particular the univerbation of phrasal conjunction, can be perspicuously modeled by constraint reranking in Stratal Optimality Theory.

2 Phrase or word?

2.1 Vedic dvandvas are phrases

Since inflectional endings and a single accent are hallmarks of wordhood in Sanskrit, the separate case ending and accent on each constituent diagnoses a Vedic dvandva as consisting of separate words.[7] Insler 1998 has supported the two-word analysis of dvandvas with a new argument from Rigvedic meter: unlike ordinary compounds, they can be placed across the caesura in a verse.[8]

(7) a. śáṃ na **índrā//pūṣáṇā** vā́jasātau (7.35.1)
success us Indra- //Pūṣan- . reward-winning-
'[may] Indra and Pūṣan [bring] us success in winning rewards'

b. ád íd **dyā́vā//pr̥thivī́** páry apaśyat (3.26.8)
Prt Prt heaven- //earth- . Preverb saw- .3
'and he surveyed Heaven and Earth'

Insler also points out that the constituents of a dvandva can be separated, not only by clitics, but even by full words.

(8) a. arcan **dyávā** námobhiḥ **pr̥thiví** (7.43.1)
venerate- .3. Heaven- . obeisances-Instr Earth- .
'they should venerate Heaven and Earth with obeisances'

b. **índrā** ha rátnaṃ **váruṇā** dhéṣṭhā (4.41.3)
Indra- . Prt wealth-acc Varuṇ- . most-giving
'Indra and Varuṇa give the most wealth'

Such discontinuous order ('tmesis') is the most compelling possible evidence for two-word status, because it is a robust cross-linguistic generalization that words cannot be syntactically divided (they are 'syntactic atoms').[9]

In sum: accent, case, caesura, and separability tell us that dvandvas consist of two words, presumably asyndetically conjoined into a syntactic phrase.

2.2 Vedic dvandvas are single words

However, other phonological and morphological data unequivocally show that Vedic dvandvas constitute single words (Wackernagel 1905: 149-50, MacDonell 1910:157).

The *ruki* rule (*s* → *ṣ* after nonlow vowels, velars, and *r*) applies regularly inside Vedic dvandvas. Only one dvandva has the requisite context, ***agníṣóm ā***, but it occurs 21 times, always with retroflexion (1.93.8, 10.66.7, voc. 1.93.1-12 11x, 10.19.1, AV. 1.8.2, 3.13.5, 6.54.2, 6.93.3, 8.9.14, 12.4.26, 18.2.53). In contrast, the *ruki* rule never applies in asyndetic phrasal conjunction:

(9) a. tváṃ tám brahmaṇas pat**e** **s**óma índraś ca (1.18.5)
you that Brahmaṇaspati- . Soma- . Indra- . and
'you, Brahmaṇaspati, Soma, and Indra'

b. índro dyávāpr̥thiv**í** **s**índhuḥ. . . (4.54.6)
' Indra, Heaven-and-Earth, Sindhu. . . '

c. ádit**iḥ** **s**índhuḥ pr̥thiví utá dyaúḥ (1.94.16 etc.)
'Aditi, Sindhu, Earth, and Heaven'

A second phonological argument comes from vocative accentuation. In Vedic, vocatives are regularly unaccented, except at the beginning of a sentence or *pāda* (line or half-line), where they

receive an accent on their first syllable, regardless of the location of the stem's inherent accent. Vocative dvandvas invariably obey this rule:

(10) a. sómam **indrābr̥haspatī** píbatam (4.49.6)
soma- Indra- -Br̥haspati- . drink- .2.
'Indra and Br̥haspati, drink the soma!'

b. **ágnīṣomāv** imáṃ sú me śr̥ṇutáṃ vr̥ṣaṇā hávam (1.93.1)
Agni- -soma- . this- well my hear- .2. bulls- . call-
'Agni and Soma, you bulls, hear well my call'

c. **índrāvaruṇā** mádhumattamasya vr̥ṣṇaḥ sómasya vr̥ṣaṇā́ vr̥ṣethām
Indra- -Varuṇa- . sweetest- bullish- soma- bulls- . quaff!'
(6.68.11)

'Indra and Varuṇa, you bulls, quaff this sweet, strong soma!'

d. samrā́jāv asyá bhúvanasya rājatho // **mítrāvaruṇā** vidáthe
rulers- this- world- rule-P .2. Mitra- -Varuṇa- . wisdom-
svardr̥śā (5.63.2)
sun-view-
' you rule over this world, Mitra and Varuṇa, in wisdom, in sight of the sun'

In contrast, each of a series of syntactically conjoined vocatives usually counts as a separate phrase:[10]

(11) a. **váruṇa mítrāryaman** várṣiṣṭhaṃ kṣatrám āśāthe (5.67.1)
'Varuṇa- . Mitra- -Aryaman- . supreme- . rule- . attain- . .2.
'you have attained supreme rule, Varuṇa, Mitra, Aryaman'

b. **ágna índra váruṇa mítra dévāḥ …** (5.46.2)
'Agni, Indra, Varuṇa, Mitra, gods'

c. máhi vo mahatā́m ávo // **váruṇa mítra** dāśúṣe (8.47.1)
great- . your- . great ones- . protection
textscsg.gen Varuṇa- . Mitra- . sacrificer- . '
'great is your protection, Varuṇa, Mitra, great ones, for the sacrificer'

These accentuation and retroflexion data show that dvandvas form a tighter prosodic unit than phrasal co-ordinations. If dvandvas are single words, this contrast follows. If they consist of two conjoined words, it is not clear how they can be distinguished from phrasal conjunction.

The two-word alternative can also be excluded on morphological grounds. Dvandvas serve as stems for the derivational suffixes *-vat (-vant), -a, -ya,* and *-ta-* (*-t ā*):

(12) *dyā́v āpr̥thivī́vat-* 'accompanied by sky and earth' (AV. 19.18.5), *mitrā́váruṇavat-* (8.35.13) 'accompanied by Mitra und Varuna', *maitr āvaruṇá-* (AB. etc.) 'relating to Mitra und Varuṇa', *maitr āvaruṇát ā* 'property of being M. and V.' (TS.), *maitr āb ārhaspatyá-* 'belonging to M. and B.' (ŚB.)

Since these suffixes otherwise take only single nominal stems as inputs, the data in (12) is inconsistent with a phrasal analysis of dvandvas.

The point is reinforced by the accentuation and vocalism of these derivatives. The suffixes *-a, -ya, -ta,* are accentually dominant, in that they eliminate the accents off the entire compound and attract it onto themselves (or in the case of *-ta,* onto the preceding syllable). No affix triggers this morphological accent deletion rule on anything larger than a word.[11] Also, the suffix *-a* causes the first vowel of its stem to be strengthened (ablauted) to so-called *vr̥ddhi* grade, in the case of *maitr āvaruṇá-* from *i* to *ai*. No affix triggers this strengthening on anything larger than a word.

It might be objected that most of the words in (12) are from the post-Rigvedic literature, and so might reflect a later stage of the language. But they are still morphologically based on the old-style dvandvas of the *mitrā́váruṇa-* type, double-accented and with the first member ending in the apparent dual *- ā, - ī* Therefore, they do provide evidence that these old-style dvandvas got treated as single words *before* they were morphologically regularized into ordinary stem-based single-accented dvandvas in later Vedic and Classical Sanskrit.

Vedic dvandvas obey two other restrictions that differentiate them from phrasally conjoined structures (as well as from later dvandvas). They are always binary, whereas phrases can have three or more co-ordinate members (as in (9) and (11a,b))[12]. And they are always definite, the majority functioning as proper names, with a smaller group denoting kinship relations and personified natural phenomena. The definiteness restriction is likewise unique to the Vedic dvandvas.[13]

It is worth mentioning here that dvandvas are largely irreversible,[14] though this does not get us much further because the order of classical dvandvas and of phrasal conjuncts is usually also fixed.[15]

Suppose we accept on the strength of the above data that dvandvas are words, consisting of a compound stem inflected with a case/number ending. What are the constituents of the compound stem? We have already seen that the second member, at least, is itself a bare stem. It appears as such in derivatives like (12), e.g. *mitrā́váruṇa-*. Therefore the dual case endings on the second member (*-ā, -bhyām*, etc.) are not part of the compound stem. They are added to the bare compound stem at the inflectional level, in the same way as they are added to any stem. When the compound takes a derivational ending, no case ending appears on it.

What about the first member? The presumption is that it is likewise a stem, its dual-like ending notwithstanding, for otherwise we would have to countenance a process of Word+Stem compounding, which is typologically at least rare. Its stem status is confirmed by several independent bits of morphological evidence. The ending is invariant *-ā* (*-ī*, if it is an *i*-stem) regardless of the syntactic role and inflection of the compound itself, as shown by its case inflection in (13):

(13) a. Genitive Dual *índrāváruṇayoḥ* (1.17.1, AV. 10.5.11 etc.)

b. Instrumental Dual *mitrā́váruṇābhyām* (5.51.9), *indrāgníbhyām* (8.40.5), *somāpūṣṇábhyām* (2.40.2), *agnīṣomābhyām* (AV. 12.4.26)

and by its number inflection in (14):

(14) a. Vocative Plural *indrāmarutaḥ* (2.29.3) ‘Indra and the Maruts’ (a set of more than two)

b. Nominative Plural *dyā́vāpṛthivī́ḥ* (AV. 8.9.16) ‘the heaven-and-earths’ (a set of six)

Only the second member has the syntactically appropriate inflection, with oblique case if the syntax so requires, and a plural rather than a dual ending if the compound denotes more than two things. If the construction consisted of two syntactically co-ordinated words, then both its members should receive the same case/number endings. Instead, the data indicate that case is assigned only to the whole compound *qua* inflected stem, and that the *-V̄* ending of the first member, is a stem-forming (compositional) suffix, which I call the for reasons that will become clear when we get to its semantics in 2.8 below,[16] and the account of its origin in 3.

In sum: phonology and morphology warrant the conclusion that dvandvas are built by putting together two stems (not two inflected words) into a compound stem, which then, like other stems,

can receive either derivational or inflectional endings. In this respect, Vedic dvandvas are built like all other Sanskrit compounds.

2.3 Resolving the contradiction

Applying the criteria for wordhood has given us contradictory results for Vedic dvandvas. By the tests in 2.1, they contain two co-ordinated words. By the tests in 2.2, they are single words consisting of a compound stem plus an inflectional ending, just like ordinary Sanskrit compounds.

In principle, there are several approaches to resolving such a contradiction.

Dvandvas could be a hybrid formation, with a mix of phrasal and lexical characteristics. Evidence that wordhood must be decomposed into partly independent bundles of properties has been accruing as linguistic descriptions have become more fine-grained, and linguistic theory now provides several (not necessarily incompatible) ways of bundling them. It has been noted for many languages that the 'words' relevant to prosodic constraints and operations are not always the same as the 'words' relevant to morphological constraints and operations. This discrepancy is modeled in Prosodic Phonology by referring morphology and phonology to distinct independent constituent structures, one representing , the other (Inkelas 1989, Inkelas and Zec 1990). Both representations constitute a complete top-to-bottom hierarchical organization for a sentence. In the default case, morphological words coincide with phonological words, but the grammar can stipulate mismatches between them under specific conditions. In addition, the two structures differ formally in certain ways (for example, morphological/syntactic categories are recursive).

The stratal hierarchy ('level ordering') complicates the concept of 'word' in another way. It provides a distinction between the and the , the latter typically a larger unit (one which includes clitics, for example). This distinction is defined in derivational terms and applies to both phonology and morphology, predicting certain interactions and correlations between them.[17] Proponents of Lexical Phonology/Morphology and of Stratal OT have argued that this division is required in addition to the one imposed by Prosodic Phonology.

Independent of enriched representations is the possibility of a constructional mitosis — that there are simply two distinct coexisting types of dvandvas in Vedic, one derived lexically, the other syntactically, with correspondingly different characteristics.

It is indicative of the Vedic dvandvas' morphological intricacy that we will need all three of

these ideas in order to make sense of their conflicting properties.

First, the recognition of two distinct coexisting structures is all but inevitable given the data of the preceding section. Recall that the conjunction 'Mitra+Varuṇa' *must* be a one-word compound stem when it is an input to morphology, as in (12) and (13):

(15) **mitrā́váruṇavantā** utá dhármavantā marútvantā . . . gachatho hávam
'Mitra- -Varuṇa-with- . and Dharma-with- . Marut-with- . come- .2 call-
(8.35.13)

'accompanied by Mitra and Varuṇa, by Dharma, by the Maruts, you two come to the call'

and it *must* be two separate words when it constitutes a discontinuous phrase in examples like (8), e.g.:

(16) cáksur máhi **mitráyor** ām éti priyáṃ **váruṇayoḥ** (6.51.1)
eye- . great- . Mitra- . there come- . .3 friendly- . Varuṇa- .
'there comes the great friendly eye of Mitra and Varuna'

The coexistence of the two structures is not in the least surprising because just about *every* type of lexical compound in Sanskrit has a potential syntactic paraphrase made up of the same lexical material.

(17) a. *viśpátiḥ* 'clan chief' (10.135.1) — a compound
b. *viśás pátiḥ* 'chief of the clan' (10.152.2) — a phrase
c. *(víśv ās āṃ) viśā́m pátiḥ* 'chief of (all) the clans' (6.15.1) — a phrase

Since the same grammatical relation can so often be expressed both morphologically as a compound and syntactically as a phrase, it would be strange if conjunction also could *not* be grammatically expressed in both these ways:

(18) a. *mitrā́váruṇayoḥ* 'Mitra and Varuṇa' (Gen.Du.) — a compound
b. *mitráyoḥ . . . váruṇayoḥ* 'Mitra and Varuṇa' (Gen.Du.) — a phrase

Dvandvas differ from other compounds only in that they so often look like their syntactic paraphrases.

So our first step towards a resolution of the contradiction is that the so-called Vedic dvandvas really are of two distinct coexisting types: syntactically generated asyndetic co-ordination structures (i.e. phrases rather than compound words), and lexically generated compound words formed by combining two stems into a compound stem, which then undergoes derivation or inflection. Split conjuncts and multiply inflected conjuncts are of the former type. On the other hand, all dvandvas which have derivational affixes, or which have the syntactically relevant inflectional ending only on their second member, or with any of the other single-word properties reviewed in section 1, are derived lexically.

This solves half the problem. The syntactic dvandvas are asyndetic co-ordination structures. Their peculiarity is that they have a dual case ending on both members. As we'll see when we get to the semantics of the dual in 2.8, this is a redundant and optional feature of the construction: either of the members can also be singular. The immediate problem that remains is that the lexical dvandvas have some unique quasi-phrasal features not found in any other type of compound. The next sections are devoted to them.

2.4 The structure of lexical dvandvas

Historically, the lexical dvandvas' phrasal features are residues of the syntactically conjoined phrases from which they originate. From the synchronic point of view, these features are not merely a random collection of anomalies. They bundle in a way which can be accounted for by the assumption that lexical dvandvas are single *morphological* words that consist of two *phonological* words (as proposed by Sproat 1985: 382–411, and in the framework of Prosodic Lexical Morphology by Han 1994).

Prosodic Phonology tells us that morphological rules and constraints refer to morphological units, and phonological rules and constraints refer to phonological units. The Vedic data bears out this predicted division of labor. Morphologically conditioned operations, such as derivational affixation and vocative accentuation, are defined on morphological words, whereas the phonological single accentuation constraint is imposed on phonological words, and caesuras are allowed at phonological word boundaries.[18]

When the phonological and morphological structures don't match, it can happen that the

caesura is placed in the middle of a morphological word — in fact, in the middle of a morphological stem, as in (19).

(19) a. virā́ṇ **mitrā́//váruṇayor** abhiśrī́ḥ (10.130.5)
Viraj- Mitra//Varuṇa- . privilege-
'Viraj was the privilege of Mitra and Varuṇa'

b. sajū́r **mitrā́//váruṇābhyāṃ** (5.51.9)
accompanying- Mitra// Varuṇa- .
'together with Mitra and Varuṇa'

They must be single morphological words because their first member has no case,[19] but two phonological words, each endowed with its own word accent, lying athwart the caesura.

But what about the *ruki*-rule? Since it is a phonological rule, Prosodic Phonology tells us that its domain must be phonological. So we would expect it to apply only within phonological words, and yet it applies in dvandvas across what we have concluded are phonological word boundaries (section 2.2). The key to the solution is that the *ruki*-rule in Vedic applies to *postlexical words.* That is, its domain includes also words created by postlexical restructuring processes, such as cliticization. It applies across clitic boundaries to initial *s*- in particles, pronouns, and certain verbs, mostly unaccented light verbs, provided of course the preceding word provides a *ruki*-triggering context (Wackernagel 1896:237). It also applies to full lexical words preceded by a particle or preposition that ends in a *ruki*-trigger. It hardly ever applies across ordinary full-fledged word boundaries beween content words ((20g) is a rare exception).

(20) a. nahí **ṣma** (8.7.21)
'not then'

b. rátho hí **ṣáḥ** (1.54.3)
charioteer for he
'for he was the charioteer'

c. yuváṃ kavī́ **ṣṭhaḥ** (10.40.6)
you- . wise-men- . are- .
'you two are wise men'

d. asmé tád indrāvaruṇā vásu **ṣyāt** (3.62.3)
us- that Indra-and-Varuṇa good- be- .3. .
'may that good be ours, Indra-and-Varuṇa'

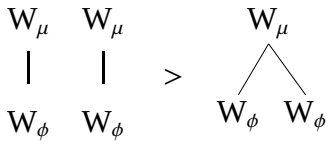
e. ghnánto vr̥trāṇi sūríbhiḥ **ṣyāma** (7.92.4)
killing- enemies- sacrificers- be- .1.
'with such sacrificers, may we kill the enemies'

f. sóma u **ṣuvāṇáḥ** sotr̥bhir ádhi **ṣṇúbhir**ávīnām (9.107.8)
'soma- Prt press-PP pressers- on backs- sheep- '
'soma is pressed by the pressers on the backs of the sheep'

g. gauryàṃ . . . padí **ṣitám** ámuñcatā (4.12.6)
Gauri- leg- tied- freed- .3.
he freed the Gauri cow, whose legs were tied'

This is clear evidence that the *ruki*-rule applies within *phonological* words, just as the theory predicts, including those phonological words which are formed postlexically through cliticization. If the *ruki*-rule in Vedic is a *postlexical word-bounded* rule, its domain includes the outputs of cliticization and univerbation. And exactly this is what I have elsewhere argued for on the basis of quite independent data having to do with the application of *ruki* in reduplication (Kiparsky 2007).

With this understanding, we can reconcile even this last diagnostic with the theory. In the lexical phonology, dvandvas are two phonological words, as we have supposed, but they are *merged into one phonological word in the postlexical phonology*. Letting W_μ and W_ϕ stand for morphological and phonological word, respectively, the following postlexical restructuring takes place.

(21) 

Crucially, the constraint that limits a phonological word to only one lexical accent *does not apply to postlexical words*. We know that already from other, much simpler facts. For example, postlexical words with two or more accents routinely arise when sandhi merges lexical words:

(22) a. sá devā́ṃ **éhá** vakṣati (1.1.2) Derivation: /ā́ ihá/ → [éhá]
he gods- Preverb here bring- .3.
'he will bring the gods here'

b. **távét** tát satyám (1.1.6) Derivation: /táva ít/ → [távét]
you- Part that- truth-
'that is your truth'

Therefore, when a dvandva is reparsed as a single phonological word postlexically, its accents remain.

Our analysis of lexical dvandvas now captures every aspect of their apparently inconsistent behavior.

(23) a. Morphology refers to morphological units:

1. Inflectional and derivational affixation: dvandvas are single stems.
2. Vocative deaccentuation: inflected dvandvas are single morphological words.

b. Phonology and meter refers to phonological units:

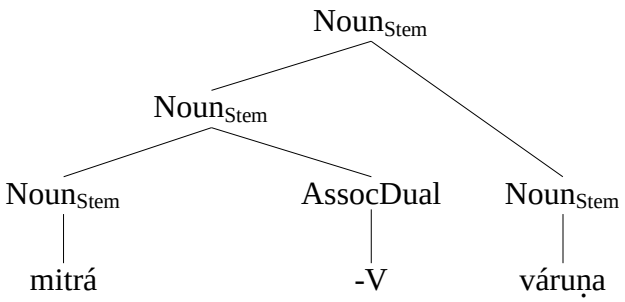
1. Single accentuation: dvandvas are two phonological words at the lexical level.
2. Caesura placement: dvandvas are two phonological words at the lexical level.
3. The *ruki*-rule: dvandvas are single phonological words at the postlexical level.

More precisely, the data show that caesura placement *can* treat dvandvas as two phonological words at the lexical level, not that it *always* does. Generally, metrical constraints can be enforced either on lexical representations or on postlexical prepresentations in Vedic versification. An example is contraction of vowels across word boundaries: $\bar{V}$ from /-V V-/ sometimes counts as two syllables, sometimes as one. This is comparable to the fact that words like *being* and *rhythm* can count either as two syllables or as one syllable in English poetry.

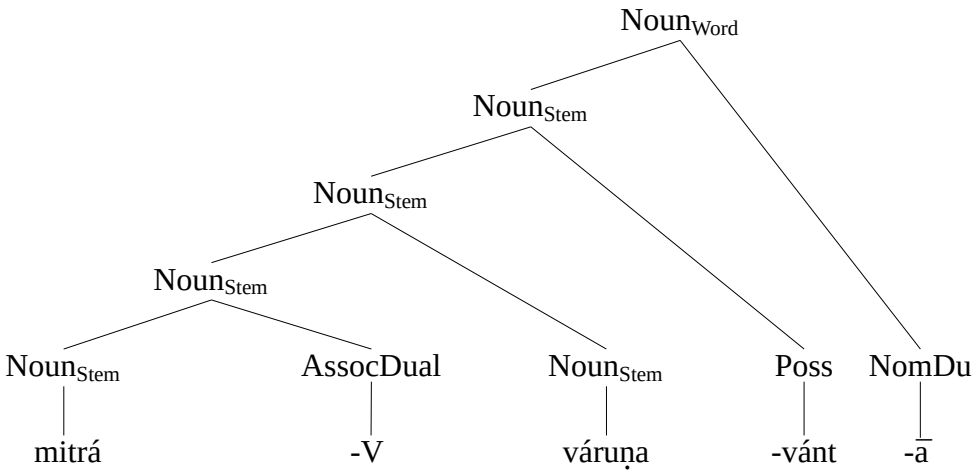
2.5 A derivation

I illustrate the morphological derivation of Vedic dvandvas with the example *mitrā́váruṇavant ā* (8.35.13) 'accompanied by Mitra and Varuṇa' (Instr.Sg.), a case-inflected adjective derived from a dvandva. Its morphological derivation proceeds as follows. (I have simplified and telescoped the steps in various harmless ways). Trees mark morphological constituency and brackets $(\ldots[\ldots]_{\omega}\ldots)$ mark prosodic constituency.

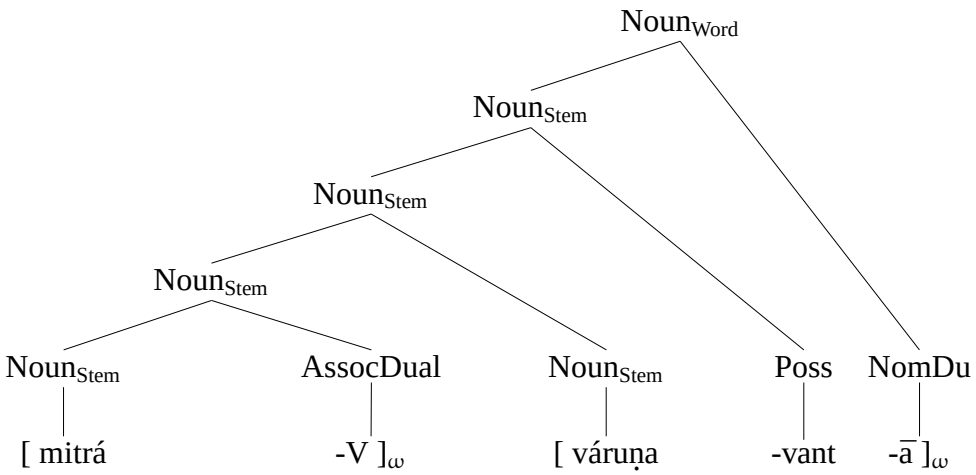
(24) a. Formation of dvandva stem *mitrā́váruṇa-*



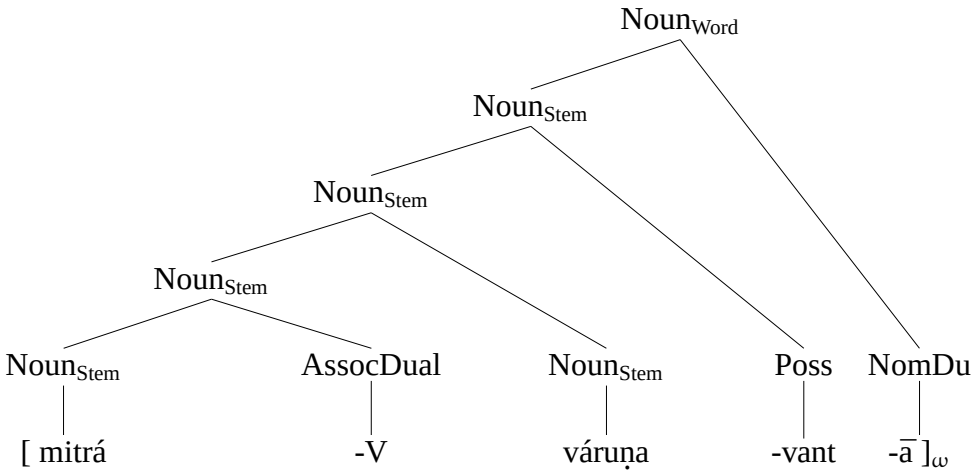
b. Affixation of derivational *-vant* and Instr.Sg. -$\bar{a}$



c. Prosodic parsing into phonological words (. . . [. . .]$_{\omega}$ . . .). The associative morpheme coerces the stem it attaches to into a phonological word; hence (by strict layering) the second member of the compound stem is a phonological word also. All but the first accent is erased within each phonological word. This is referred to in the literature as the Basic Accentuation Principle (BAP).



d. Postlexically, the compound is restructured into a single phonological word, in obedience to the constraint requiring alignment of morphological words with phonological words, neither affecting nor affected by the double accentuation (since the single-accent constraint on words is not active postlexically).



In terms of an OT grammar, the prosodic reparsing as a single phonological word involves a constraint reranking. A fuller presentation in diachronic perspective follows in section 3, but for now let us simply note that the prosody/morphology interface for Vedic dvandvas can be characterized by the two general constraints (25a,b) plus an idiosyncratic morpheme-driven constraint (25c) which corresponds to the associative suffix's idiosyncratic property of imposing phonological word status on its stem.

(25) a. $W_{\mu} \subseteq W_{\phi}$ (A morphological word must be contained in a phonological word.)

b. *W_{μ} (Minimize the number of phonological words.)

c. The associative ending is aligned with a phonological word.

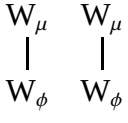
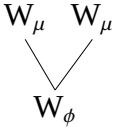
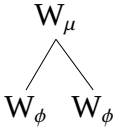
In the Vedic lexical phonology, the constraint (25a) which requires that a morphological word must be contained in a phonological word is dominated by the other two constraints, and hence violated when they must be satisfied, which is precisely in dvandva compounds.

(26) Vedic lexical phonology: *W_{μ}, -$\bar{a}]_{W_{\phi}}$ ≫ $W_{\mu} \subseteq W_{\phi}$

	*W_{μ}	-$\bar{a}]_{W_{\phi}}$	$W_{\mu} \subseteq W_{\phi}$
1. W_{μ} W_{μ} | | W_{ϕ} W_{ϕ}	**	✓	✓
2. W_{μ} W_{μ} \ / W_{ϕ}	**	*	✓
3. ☞ W_{μ} / \ W_{ϕ} W_{ϕ}	*	✓	*
4. W_{μ} | W_{ϕ}	*	*	✓

Lexical phonological rules apply to this representation. In the postlexical phonology, the constraint $W_{\mu} \subseteq W_{\phi}$ is promoted, triggering prosodic restructuring of dvandvas into single words. Postlexical phonological rules (such as the *ruki*-rule) apply at that point.

(27) Vedic postlexical phonology: $W_{\mu} \subseteq W_{\phi}$, *$W_{\mu}$ ≫ -$\bar{a}]_{W_{\phi}}$

	$W_\mu \subseteq W_\phi$	$*W_\mu$	$\text{-a}]_{W_\phi}$
1. 	✓	**	✓
2. 	✓	**	*
3. 	*	*	✓
4. ☞ 	✓	*	*

2.6 The ambiguous cases

Now for the hardest part of Insler's question '*mitrā́váruṇ ā* or *mitrā́ váruṇ ā*?'. What about ordinary adjacent pairs which could be either lexical or syntactic? For most dvandvas in the texts, the two parses are string-identical.[20]

(28) **mitrā́váruṇā** huve (1.23.5)
Mitra-and-Varuṇa- call- .1.
'I call on Mitra and Varuṇa'

Are they morphological words, or syntactic phrases? How can we tell? Does it matter?

Application of the abovementioned phonological diagnostics to the indeterminate cases reveals that they consistently count as single words (*mitrā́váruṇ ā*, not *mitrā́ váruṇ ā*). For adjacent duals that are morphologically ambiguous between compound and phrase, the phonological criteria, wherever applicable, point to the one-word structure, not to the phrase. For example, the phrasal conjunction option predicts that there should be syntactically conjoined sequences of dual vocatives with separate accents such as **mítr ā váruṇā* 'O Mitra-Dual, O Varuṇa-Dual', parallel to singular sequences like (11c). But, with noun pairs that allow compounding, such sequences

never occur. The vocatives are always unaccented, of the type *mitr ā́varuṇā*, and *mítr ā́varuṇā* in line- and *pā́da*-initial position.

Similarly, *agníṣóm ā́* invariably has *-ṣ* from the *ruki*-rule, in accord with the lexical analysis. If there were any phrasally conjoined instances of this surface string, they would show up as **agní sóm ā́* because the *ruki*-rule is inapplicable across the boundary between two lexical words. But no such cases are attested.

The phonology, then, indicates that morphologically ambiguous cases are normally lexical compounds. How is this to be accounted for formally? It is not possible to rule out the homonymous syntactic derivation directly in the grammar, for it must be available to generate the unambiguously phrasal constructions (such as the discontinuous instances). And no reasonable syntactic constraint on conjunction can rule out precisely those conjoined phrases that *look* like they *could be* compounds. But they can be ruled out *indirectly* by a principle which selects the simplest of a set of competing synonymous forms, in this case a compound over a phrase with the same form and meaning. It is a classic effect, motivated by economy of output structure.

2.7 Blocking

My treatment of blocking follows Wunderlich 1996 in positing two components of grammar, a and a . The generative component specifies the potential expressions of the language and their potential interpretations, and the blocking mechanism functions as a filter that resolves the competition between the potential expressions whose meaning is compatible with a given input meaning (the 'intended meaning') and sorts them into the optimal paradigms. The paradigms emerge from the interaction of , which requires that every cell is filled, and , which requires that it is filled only by one form, the simplest one that the generative component supplies.[21] The crucial and controversial assumption here is that blocking applies not only internally to the morphology or the lexicon, but holds between competing expressions of whatever kind, thereby generating the kinds of paradigms composed of a mix of analytic (morphological) and periphrastic (syntactic) forms that are ubiquitous in language (Poser 1992, Hankamer & Mikkelsen 2005; Embick 2007 and Embick & Marantz 2008 put forward a dissenting view).

A familiar English example will illustrate how morphology blocks syntax. The syntax of English generates not only the expressions *John didn't leave, Did John leave?, Didn't John*

leave?, but also *(*)John did leave.* Unless its auxiliary is stressed, *did leave* is superseded ('blocked') by the competing simpler expression *left*, which is generated by the morphology. The one-word expression *left* prevails over the two-word expression *did leave*, unless other factors, such as emphasis on the auxiliary *did*, prevent it.[22] This obviates the need for 'Affix-hopping' and similar operations that have been posited from *Syntactic Structures* down to Distributed Morphology.

For Sanskrit, the perfect furnishes a similar example. Morphological perfects are formed from most simple verbs, e.g. *cakā̆ra* 'has done', from *kr̥*. Derived verbs don't allow morphological perfects, nor do simple verbs of certain phonological shapes, e.g. V-initial superheavy roots. No verbs that allow morphological perfects allow periphrastic perfects *(*kar ā̄m cakre, *kr̥ṇav ā̄m cakre* 'has done'). Conversely, all verbs which don't allow morphological perfects do form periphrastic perfects, e.g. *mr̥g ā̄y ā̄m cakre* 'has hunted' (*mr̥g ā-ya-* 'hunt' is a derived verb), *īkṣā̄m cakre* 'has seen' (*īkṣ* 'see' is a superheavy root). As a result, each verb has one and only perfect, with simple and periphrastic forms in complementary distribution together constituting a

.

This is jointly ensured by F and M . When the simple forms are precluded, by phonological or other constraints, the periphrastic forms automatically fill the gap. The following simplified constraint system illustrates the idea.

(29) a. R : Superheavy roots beginning with a vowel don't reduplicate.

b. F : Express the meaning of the input.

c. M : Avoid complexity (for present purposes: minimize the number of words).

The blocking mechanism adjudicates between three candidates that the generative component offers for a given paradigmatic cell in the perfect paradigm: the analytic (one-word) form, the periphrastic (two-word) form, and the null form (a paradigmatic gap). The following tableau shows how the one-word form defeats the periphrastic form by M and the paradigmatic gap by F (set A), and how the more complex periphrastic form steps in when phonology blocks the one-word form (set B).

(30)

		R	F	M
A	1. ☞ *ca-kār-a*			*
	2. **kar-āṃcakre*			**
	3. ∅		*	
B.	1. **i-ikṣ-a*	*		*
	2. ☞ *ikṣ-āṃcakre*			**
	3. ∅		*	

Paradigms, on this view, are not listed, or directly generated by rules or constraints, and they are not primitives of the theory; rather, they emerge by OT constraints from the competition between expressions. Inflectional and productive derivational categories typically form paradigmatic subsystems constituted by blocking.

The availability of blocking simplifies the generative component by reducing many apparently arbitrary combinatoric restrictions to blocking effects.[23] It is also recommended by theoretical parsimony. Since blocking in morphology and the lexicon is required in all theories anyway, extending it to syntax obviates the use of distinct mechanisms of syntactic movement in word-formation, such as adjunction and metathesis. case, and dispenses with morphological operations such as. Furthermore, it captures an important generalization: the conditions that determine the distribution of simple and periphrastic forms are always more perspicuously stated on the simple form. For example, in Sanskrit there are obvious morphophonological reasons why superheavy V-initial roots and derived roots do not reduplicate,[24] but would not at all be clear why the complement of that class should resist Affix-Hopping. Similarly, in English it is easier and more natural to state the distribution of the comparative and superlative in terms of (largely phonological) restrictions on the distribution of *-er* and *-est* than in terms of a restrictions on the distribution of *more* and *most.* Accounts which derive words from phrases have no explanation for this generalization, as far as I can tell.

Blocking fails when requires the synthetic expression to express some extra information that is part of the intended meaning (the 'input' in the OT tableau). For instance, *John did leave* is acceptable when it expresses an extra component of focus meaning that *John left* does not express, in which case F selects it when that component of meaning is part of the input. In general, outside of paradigmatic subsystems there is usually no complete synonymy

between expressions, and consequently no categorical blocking. For example, *descend* does not block the hyponymous *go down* and *come down*, *glass shelf* does not block the hyponymous *shelf for glass, shelf for glasses, shelf of glass*, etc. Although Sanskrit usage at all stages favors compounds over their analytic paraphrases, so that a compound (if it is grammatically possible at all) is almost always the more frequent option, and often the only attested option,[25] phrasal equivalents sometimes do occur. For example, the compound *viśpáti* 'clan chief' (see (17)) occurs 31 times in the Rigveda, but the phrase *viśás páti* 'chief of the clan' is also found twice, and *viśám páti* 'chief of the clans' occurs three times. One reason to choose the full phrase may be to express the number feature on the complement, which is neutralized in the compound. Moreover, in three of the cases the phrasal construction is grammatically unavoidable because the genitive is modified by an adjective.

Since the dvandva compounds are dual, they are equivalent to the corresponding phrase in their number features. In this case, the syntactic construction is evidently not motivated by the need to express some extra semantic content as in (17). The phrasal construction might sometimes be motivated by the need to modify just one of the conjuncts. For example, the reason the conjuncts are separated in (31) may be so that the epithet *tuvij ā́ta* should refer specifically to Varuṇa, as in 2.27.1, 2.28.8 (it never applies to Mitra).[26]

(31) **mitráyor váruṇayoḥ** …tuvijā́táyoḥ (7.66.1)
Mitra- . Varuṇa- . …strong-natured- .
'on Mitra and on strong-natured Varuṇa'

But for the most part dvandva compounds should block the corresponding string-identical phrases, as indeed they do.

A similar instance of blocking arises in preverb+verb combinations, which are structurally ambiguous for similar reasons as dvandvas. Their separability ('tmesis', see (32b)) shows that they can be syntactically combined as separate words. But they can also be combined lexically, for a preverb+root combination also functions as an input to morphological affixation, including the formation of the absolutive, which results in single-word compounds like (32c), where tmesis as in (32d) is quite impossible.[27]

(32) a. íḷā́bhiḥ **sáṃ rabhemahi** (8.32.9) 'may we get hold of abundance'

b. **sám** iṣ́á **rabhemahi** (1.53.4) ‘may we get hold of food’

c. **saṃrábhyā** (10.94.4) ‘having got hold of’

d. ***sam iṣ́á rábhyā** ‘having got hold of food’

As in the case of dvandvas, we can ask about the status of ambiguous cases like (32a). Are they single words derived by inflecting a prefixed root? Or two separate adjacent words? Blocking gives priority to a morphological one-word expression over the corresponding syntactic two-word expression, unless the latter is required to express some additional intended meaning. No doubt tmesis, as in (32b), is motivated by metrical requirements, by focus/emphasis, or perhaps other rhetorical effects. In those cases where the two expressions are string-identical it can serve no such additional function. And once again, the phonology regularly treats preverb+verb combinations as one word rather than two; in fact, here it shows that they are single words even in the *lexical* phonology. They have a single accent (on the preverb if the verb is accented, as it usually is in subordinate clauses, on the verb otherwise), and the *ruki*-rule and *n*-retroflexion apply:

(33) a. *ví ṣyanti* (1.85.5) ‘they wet’, *pári ṣicyate* (4.49.2) ‘is poured’, *ní ṣīdata* (1.22.8) ‘sit down’

b. *pári ṇīyate* (3.2.7) ‘is led around’, *prá ṇonumaḥ* (1.78.1) ‘we shout’

So it seems that string-adjacent preverb+verb combinations also show near-categorical blocking of syntax by morphology, confirming my explanation for the distribution of dvandvas.[28]

The blocking principle is diachronically manifested in the reanalysis of phrases into words (univerbation), and indeed provides the theoretical explanation for the unidirectionality of this process. Over time, the preference for the simplest expression pushes ambiguous cases from their original phrasal analysis to their one-word analysis. In other words, it drives expressions along the grammaticalization trajectory in which ‘the parts of a constructional schema come to have stronger internal dependencies’ (Haspelmath 2004), turning clitics into suffixes, obliterating morpheme boundaries, etc. (for a more careful formulation see Kiparsky in press). In the case of ambiguous dvandvas like *mitrā́váruṇa-*, the single-word preference correctly predicts that the lexical (single-word) analysis will supplant the phrasal analysis. The evolution of dvandvas, then, can take its place among other well-documented grammaticalization processes.

In summary, the grammar conjoins nouns asyndetically in two ways: in the syntax by co-ordination, and in the morphology by compounding. There are unambiguous syntactic co-ordinations and unambiguous morphological compounds, but the bulk of attested dvandvas are superficially ambiguous between the two analyses. Three things indicate that the ambiguous cases are treated only as compounds: the general cross-linguistic preference for the simplest structure, the observed direction of language change in Sanskrit, and finally phonological data internal to the language. I attributed the preference for the compound analysis to an independently motivated blocking principle, whose operation is closely paralleled in Sanskrit preverb+verb combinations.

2.8 Semantics: the associative dual

We have concluded that the ending of the first member of Vedic dvandvas is not a case ending, contrary to appearances, but an associative derivational morpheme. Let us now try to pin down its meaning and function. This will complete the synchronic analysis, and prepare the ground for our exploration of its history in the next section.

The so-called of Vedic mark their stems as one of a pair of associated items, e.g. *dyáv ā* 'Heaven and Earth' (literally 'Heaven-Dual'). They are termed 'elliptic' because they were once thought to be derived from dvandvas by dropping their second member. But it has long been recognized that they are actually older than dvandva compounds and more widely distributed in Indo-European languages, and constitute the historical basis of the Vedic dual dvandvas (Wackernagel 1905). Although the older view got the historical relationship backwards, it was based on the real insight that the two formations are semantically closely related.

The -*ā* of elliptic duals, like that of first members of dvandvas, is an A D in the sense of Cysouw 2003, Moravcsik 2003, and Daniel 2005. It yields an inherent dual — the associative dual — which denotes, not a set of two Ns, but N and something else which forms a natural or conventional pair with N. For example: the regular duals of *father* and *mother* mean 'two fathers' and 'two mothers', but the associative duals of *father* and *mother* both mean just 'parents'[29]. And the regular duals of *night* and *dawn* means 'two nights' and 'two dawns', but the associative dual of *night* and *dawn* both mean just 'Night and Dawn' (personified as deities).

As described above, the associative stem is morphologically marked by a lengthening of the stem-final vowel; after a consonant stem it is -*ā*. The associative stem marks its stem as being a , with all attendant consequences.

The associative stem denotes the base as one of a pair. The other member of the pair need not be explicitly specified. If it is left implicit, it can be identified by contextual inference or convention. This is the elliptic dual. The other referent can be explicitly specified in one of several ways, either syntactically or morphologically. It can be expressed by a syntactically co-ordinated nominal. As is usual in Sanskrit, the co-ordination can either be marked by a conjunction (usually *ca*) or remain unmarked. In either case, the coordinated nominal can either be dual also, or it can be singular. Examples of co-ordinated duals are given in (34).

(34) a. Explicitly co-ordinated duals:

náktā ca cakrúr uṣásā (1.73.7)
Night- . and created- .3. Dawn- .

'they have created Night and Dawn'

b. Asyndetically co-ordinated duals:

mitráyor váruṇayoḥ (7.66.1)
Mitra- . Varuṇa- .

'on Mitra and Varuṇa'

Examples of the second conjunct in the singular are given in (35):

(35) a. Dual with explicitly co-ordinated singular:

mitrā́ ...váruṇo yáś ca sukrátuḥ (8.25.2)
Mitra- . ...Varuṇa- . who . and wise .

'Mitra and wise Varuṇa'

b. Dual with asyndetically co-ordinated singular:

vánaspátīn uṣā́sā náktam óṣadhīḥ (8.27.2)
trees- . , Dawn- . Night- . plants- .

'the trees, Dawn, Night, the plants'

These variants convey the same meaning in practice, for the associative dual by itself marks its stem as one member of a conventionally associated pair. The words *nákt ā* 'Night-Dual' and *uṣás ā* 'Dawn-Dual', *mitráyoḥ* and *mitrā́* 'Mitra-Dual', (34, 35b) by themselves already imply the pairs

‘Dawn and Night’ and ‘Mitra and Varuṇa’, respectively. Specifying the other conjunct does not add anything to the inferred meaning of the dual nouns but just makes it explicit.

The ‘missing’ conjunct is sometimes modified by an adjective, as detectable by a gender mismatch. In (36), the masculine duals *pitár ā* ‘fathers’ (‘parents’), and *dyáv ā* ‘heavens’ (‘Heaven-and-Earth’) are modified by *feminine* dual adjectives, evidently agreeing with the implicit dual second members *mātár ā* ‘mothers’ and *pŕ̥thivī* (Oliphant 1912: 35).

(36) a. pūrvajé pitárā
before-born- . . father- . .
‘the parents born before us’ (7.53.2)

b. ubhé dyávā
both- . . Heaven- . .
‘both heaven and earth’ (9.70.2)

It has been observed that 1/2P Dual and Plural pronouns have an associative meaning which is similar to that of the dual names just considered. E.g. *we* = ‘I and the other people in some (either implicit, or explicitly specified) group’ (Cysouw 2003, Daniel 2005). So, since the 1/2P Dual is in a sense an ‘elliptic dual’ (as noted already by Wackernagel), the ‘inclusory’ construction in (37) is just a special case of the construction (35a) (on which see Klein 1985:126 ff.).

(37) a. á yád ruhāva váruṇaś ca návam (7.88.3)
on when climb- .1 Varuṇa- . and boat- .
‘When we and Varuṇa (= Varuṇa and I) climb(Dual) on the boat’

b. úd yád bradhnásya viṣṭápaṃ gr̥hám índraś ca gánvahi (8.69.7)
up when sun- . height- . house- . Indra- . and climb- .1
‘When we and Indra (= Indra and I) climb(Dual) home to the height of the sun’

Importantly, the plural does *not* have an associative interpretation with names and kinship terms,[30] only with pronouns (‘we’ and ‘you’ mean ‘I/you and others’). This supports our analysis, according to which Sanskrit has an associative dual but not an associative plural in compounds.

Finally, the second conjunct can be specified morphologically, in a compound. The result is just a dvandva of the type under discussion here, e.g. *náktoṣā́s ā* (9.5.6) ‘Night and Dawn’. As described above, such a dvandva stem can then get either secondary denominal affixes, or case/number affixes. Because these affixes go on the whole compound, the first member remains

invariant in either case, and the oblique case appears only once, on the second member of the compound: *indr̄ap ū̱ṇóḥ* (1.162.2), *índr̄aváruṇayoḥ* (1.17.1) etc.

Gratifyingly, Vedic dual-dvandvas fit all established typological generalizations about associatives. Associative duals/plurals always attach to a semantically restricted subclass of nouns or pronouns. They are without exception , and tend to refer to humans (or divinities); most often they are proper names, kin terms, or titles (Moravcsik 2003). This is certainly true of Vedic dvandvas.

A well-known example is the Dyirbal associative dual *-gara* 'one of a pair' (Dixon 1972:230-1).

(38) a. burbula-gara banijɲu
burbula-A D come
'Burbula and another person are coming'

b. burbula-gara badibadi-gara banijɲu
burbula-A D badibadi-A D banijɲu
'Burbula, being one of a pair, and Badibadi, being the other of the pair, are coming'

Moreover, associatives are often *derivational* suffixes, as proposed here for Vedic. For example, in Hungarian the associative plural evidently does not belong to the inflectional paradigm. The relationship seen in Vedic between the associative interpretation and the ordinary number value also has parallels; for example, Japanese *-tati* is apparently ambiguous between an associative and ordinary plural.

(39) a. Hungarian *-ék*: *tanitó-ék* 'the teacher and his group', *János-ék* 'János and associates' (contrast *János-ok* 'the Jánoses = people called J.)'

b. Japanese *-tati*: *sensei-tati* (1) 'teacher and his group', (2) 'teachers'

3 History of Vedic dvandvas

3.1 Their rise

The Indo-European dual ending is reconstructed as *-h_1* (e.g. Greek *ósse* '(two) eyes').[31] It is formally identical with the comitative/instrumental singular suffix *-h_1* contained in such forms as Latin *bene*. Hans Heinrich Hock, Stephanie Jamison, and Calvert Watkins *(voce)* have suggested

to me that instrumental -h_1 and dual -h_1 are cognate. The link would be the associative dual, semantically just a short step away from the comitative, which is probably the original function of the instrumental case:

(40) a. *N-Comitative X* 'X with N'

b. *N-Associative X* 'N and X'

And of course comitatives have an associative function in some languages, such as Finnish:[32]

(41) me mentiin Mati-n kanssa
we went Matti- with
'Matti and I went'

So the historical connection between instrumental case and associative dual, from which the general dual number would in turn derive, is quite plausible.[33]

Another precondition for the rise of Vedic dvandvas is asyndetic conjunction. This syntactic construction is found in several other branches of Indo-European (Wackernagel 1905: 150), though it is hard say whether it goes back to the proto-language or not.

(42) a. OCS *Borisa Gl̀ěba* (more often . . . *i Gl̀ěba*, Žolobov 2002, Liukkonen 1973)

b. Av. *ahura⃗ěbya miθra⃗ěbya* 'for Ahura and Mithra', *pasuůā vīrai̯ā̊* 'cattle and men' (*Yt.* 13.10; see Watkins 1995:211 for the Indo-European formulaic background of this dvandva.)

c. Hitt. *naḫšaradduš weritemuš* 'Fear and Fright' (Puhvel 1977)

In spite of the fact that the combination of an associative dual and asyndetic conjunction was present in a number of Indo-European languages, only the Indic branch seems to have grammaticalized them into a productive morphological operation of dvandva compounding (although some other branches have independently developed dvandva compounds of the unremarkable Classical Sanskrit type, fn. 5). Perhaps this has to do with the productivity of its compounding, vigorous even in the earliest Vedic and reaching unprecedented heights in Classical Sanskrit.

3.2 Their loss

We can now solve the puzzle we began with. Univerbation reduces structure by condensing syntactically generated phrases into morphologically generated compound stems. Why then does the path from the pre-Vedic asyndetic conjunction structure to the regular classical dvandva compounds go via the complex Vedic dvandvas, with their stem-forming associative suffix, and their morphology/phonology mismatch? If reduction of structure is a kind of simplification or optimization, why does it increase the overall complexity of the system, if only at a transitional stage?

A mechanism which could be responsible for such increases of grammatical complexity is successive misanalysis and partial retrenchment in real-time acquisition. The idea is that learners can internalize their own and each other's wrong outputs, by using them productively and imitating one another; if they then incorporate them partially into their evolving grammar, hybrid structures and increased complexity can result. According to this scenario, misanalysis of conjoined phrases as regular morphological compounds at early stages of acquisition would have introduced pronunciations into the ambient language which, if retained after the full evidence prompts their 'correct' analysis as two phonological words, would have forced learners to assign them the hybrid structure of Vedic described in section 2.

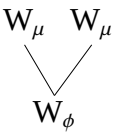
This account is incomplete because it provides no explanation for the orderly separation of of phonological and morphological properties that we have found. Obviously not any misanalysis is a potentially successful innovation. I propose that the missing part of the story is supplied by Optimality Theory, and its Stratal version in particular. It can model the observed bundling of properties, and the partial regularizations which give rise to them, on the basis of the idea that constraints are ranked and violable, and that constraint violations occur if and only if more highly ranked constraints compel them.

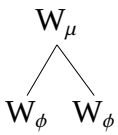
Let us return to the constraint-based analysis of dvandvas, this time presenting the constraint system in a more complete and principled way and bringing in the historical perspective. There are two pairs of core morphological constraints. The first pair, given in (43), impose a match between phonological and morphological words.

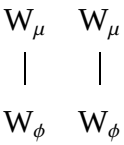
(43) a. $W_{\phi} \subseteq W_{\mu}$ (A phonological word must be contained in a morphological word.)

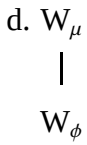
b. $W_{\mu} \subseteq W_{\phi}$ (A morphological word must be contained in a phonological word.)

These two constraints are respectively violated by mismatches like (44a) and (44b), and both are satisfied by structures where phonology and morphology are congruent, as in (44c) and (44d).

(44) a. 

b. 

c. 

d. 

The second pair of constraints minimizes structure by requiring a one-word analysis rather than a two-word analysis, one on the phonological tier, the other on the morphological tier. It is these constraints that diachronically drive univerbation, partly in conjunction with the constraints in (43).

(45) a. *W_ϕ (Minimize the number of phonological words.)

b. *W_μ (Minimize the number of morphological words.)

Constraint (45a) is violated twice in (44b) and (44c) and once in (44a) and (44d), and, constraint (45b) is violated twice in (44a) and (44c) and once in (44b) and (44d). The minimal structure (44d) is optimal on both counts. For (45) it is especially important to keep in mind the OT principle that constraints are violable and that they are violated only when higher-ranked constraints require it. Obviously not any pair of adjacent words are subject to analysis as one word, because of higher-ranked constraints that supersede (45). These include at least the constraints which govern syntactic structure (X-bar structure, government etc.) and morphological structure (contiguity of parts of a word, etc.). To simplify matters, let us assume here that such potential overgeneration of (45) is taken care of by the appropriate set of dominant constraints, and turn to the conjoined structures at issue.

The constraints in (43) and (45) are all we need to get the Classical Sanskrit one-word analysis of compounds (actually (43a) and (45b) don't do anything useful for this body of data, but I include them in the tableaux anyway for the sake of completeness). In the derivation of the pre-Vedic two-word structure and of the Vedic hybrid structure, some morphological/semantic constraint enforces the associative suffix. For synchronically arbitrary (but historically explicable) reasons this formative turns its stem into a phonological word. Let us represent this idiosyncratic property by the dominant language-specific constraint (46).

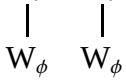
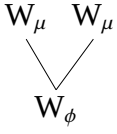
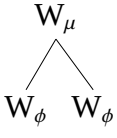
(46) -$\bar{a}]_{W_\phi}$ (-$\bar{a}$ is aligned with a phonological word)

This constraint defeats the one-word analysis that (43) and (45) would otherwise impose. Different rankings of these constraints generate the three stages in the evolution of dvandvas we have considered above: the initial Indo-Iranian stage of asyndetic co-ordination, the final Classical Sanskrit stage of one-word compounding, and the hybrid structure of the intermediate Vedic stage.

Change is typically constraint promotion. For example, regular sound change is modeled most simply as promotion of a phonological markedness constraint to undominated status in the postlexical phonology. The ranking may later spread to the lexical phonology. Descriptively, the effect is that the constraint becomes unviolated, and eventually becomes phonologized. If we assume a similar constraint promotion analysis for morphological change, then we can model the evolution of dvandvas with the constraints presented here as follows.

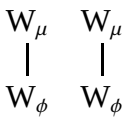
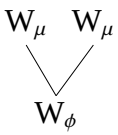
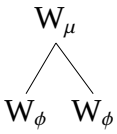
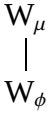
The starting point is the Indo-Iranian (and probably late Indo-European) system in (47). The associative morpheme enforces makes its stem into a word, which prevents compounding.

(47) Indo-Iranian ranking: -$\bar{a}]_{W_\phi}$, $W_\phi \subseteq W_\mu$, $W_\mu \subseteq W_\phi \gg *W_\mu, *W_\phi$

	-$\bar{a}]_{W_\phi}$	$W_\phi \subseteq W_\mu$	$W_\mu \subseteq W_\phi$	$*W_\mu$	$*W_\phi$
1. ☞ 	✓	✓	✓	**	**
2. 	*	*	✓	**	*
3. 	✓	✓	*	*	**
4. 	*	✓	✓	*	*

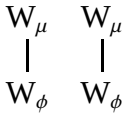
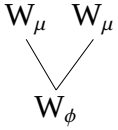
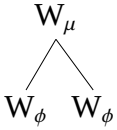
In the transitional system of Vedic, the lexical phonology differs from (47) in that the constraint (45b) *W_μ is promoted. (The arrows are just there to visualize the reranking; they have no formal status in the analysis.)

(48) Vedic (lexical phonology): *W_μ, -a͡]$_{W_\phi}$, $W_\phi \subseteq W_\mu \gg W_\mu \subseteq W_\phi$, *$W_\phi$

	*W_μ	-a͡]$_{W_\phi}$	$W_\phi \subseteq W_\mu$	$W_\mu \subseteq W_\phi$	*W_ϕ
1. 	**	✓	✓	✓	**
2. 	**	*	*	✓	*
3. ☞ 	*	✓	✓	*	**
4. 	*	*	✓	✓	*

The postlexical phonology of Vedic is already more advanced, with (43b) $W_\mu \subseteq W_\phi$ also promoted.

(49) Vedic (postlexical phonology), Classical: $W_\mu \subseteq W_\phi$, *$W_\mu \gg$ -a͡]$_{W_\phi}$, $W_\phi \subseteq W_\mu$, *W_ϕ

	$W_\mu \subseteq W_\phi$	$*W_\mu$	-ā]$_{W_\phi}$	$W_\phi \subseteq W_\mu$	$*W_\phi$
1. 	✓	**	✓	✓	**
2. 	✓	**	*	*	*
3. 	*	*	✓	✓	**
4. ☞ 	✓	*	*	✓	*

In the Classical Sanskrit system, this ranking is further extended to the lexical phonology. Once this innovative ranking is established, the associative suffix no longer has any visible effects, because the higher-ranked constraints defeat them. The suffix is still assigned to the first member of dvandvas by the appropriate morphological constraints (not spelled out here), but it can no longer induce phonological wordhood on its stem. At this stage, prevalent in late Vedic, but with early examples already in the Rigveda, dvandvas have a single accent on the final syllable, in spite of the associative, which is at this point always -ā.

(50) N.A.Du. *vā́tāparjanyā́* 'Wind and Rain', N.A.Du. *sū́ryācandramásā* Sun and Moon', Gen./Abl.Du. *indrāpū́ṣṇóḥ* 'Indra and Pūṣan', Instr/Dat./Abl.Du. *somāpū́ṣṇábhyām* 'Soma and Pūṣan'.

This reflects the ranking (49) throughout, by which dvandvas are single compound words both in the lexical phonology and in the postlexical phonology, hence with a single accent. Since they are single words at all levels of representation, the prediction is that these dvandvas are *not* split across a caesura. This is in fact correct (Insler 1998:288).

Like other oxytones, the C-stems are accentually mobile, with suffixal accent in the oblique cases *(cas faibles)*, e.g. Gen.Du. *indrāpū́ṣṇóḥ* 1.162.2. The fact that the loss of the accent results in oxytone stems is predicted by the account of Vedic accent given in Kiparsky (1984).

This late Vedic system finally passes into Classical Sanskrit with one additional change, the loss of the associative morpheme.[34] Once that happens, the dvandvas are just stem+stem compounds with no remaning phonological or morphological irregularities whatsoever. The nature and cause of this additional morphological change requires further investigation, but it is very likely connected to a semantic change in the category of dual. A clue which points to that conclusion is the fact that the elliptic dual of nouns is lost at the same time as the Vedic dvandvas as lost. From this point onwards the dual is just a number, independent of definiteness and humanness, parallel to the plural, with no associative interpretation (outside of first and second person pronouns, for which the associative interpretation is available in all languages). The fact that elliptic duals and Vedic dvandvas disappear hand in hand confirms the relationship between them that our analysis in section 2.8 posited.

Let us suppose that changes take place in minimal increments, and that they are initiated in the postlexical component (as has been argued in other work). On purely theoretical grounds, this implies the following micro-stages of univerbation.

(51) a. Postlexical promotion of (45b) *W_{μ},

b. lexical promotion of (45b) *W_{μ},

c. postlexical promotion of (43b) $W_{\mu} \subseteq W_{\phi}$ (Vedic),

d. lexical promotion of (43b) $W_{\mu} \subseteq W_{\phi}$ (Classical Sanskrit).

Change (51a) would have been essentially covert, with no visible effects. The various phonological and morphological innovations attendant upon univerbation would have come in with changes (51b) and (51c). These could have taken place in either order; attested Vedic is reached when both have taken place. The precise path is unfortunately not accessible in the historical record, and perhaps cannot be reached through comparative reconstruction either.

With the help of Prosodic Phonology, and a particular approach to blocking which crucially allows blocking interactions between words and phrases, our OT analysis has provided a synchronic rationale for the strange grammar of Vedic dvandva compounds, which in turn enabled us to trace their evolution from Indo-Iranian to Classical Sanskrit. The diachronic analysis supports the idea that grammaticalization/lexicalization, like ordinary exemplar-driven analogical change, reduces complexity by eliminating gratuitous structure and arbitrariness. In the

subsystem studied here, the change begins by eliminating the dvandvas' phrasal structure, synchronically unmotivated in view of their name-like semantics. The result is the Vedic stage that we examined at length, where the main sources of arbitrariness are the idiosyncratic morphological constraint (46), and the disparities between the lexical and postlexical systems, the grammatical locus of both being the associative dual suffix (originally the N.A.Dual ending), which imposes phonological wordhood on its stem. The subsequent regularization of dvandvas involves the elimination of this idiosyncrasy, first partially, then completely, by reranking in two steps. The relative complexity of the intermediate stages is a necessary consequence of the incremental progress of change in complex systems. Change in grammars is neither massively catastrophic, not gradient, but unfolds insmall discrete increments. On this understanding, the evolution of dvandvas is compatible with, and indeed supports, the view that grammaticalization is optimization.

Notes

[1] I believe the term ‘co-compound’ was first used by K.P. Mohanan in his analysis of Malayalam compounds (Mohanan 1982).

[2] Dvandvas of the Vedic type are also called dual dvandvas, doubly dualized dvandvas, or *devat ā*-dvandvas (‘deity-dvandvas’), in order to distinguish them from the dvandvas of the later language illustrated in (2), which have the structure of ordinary nominal compounds. Wackernagel 1905, Oliphant 1912, and Ryan MS. provide overviews of Vedic dvandvas.

[3] The location of this accent is not always attested, because accent was lost in post-Pāṇinian Sanskrit, but Vedic accented texts and Pāṇini’s rule 6.1.223 agree in putting it normally on the final syllable of the second member. In a few special cases the compound is accented on its first member’s inherently accented syllable (Pāṇini 6.2.34-37).

[4] ‘Two members, each preserving its own accent and dual form, coalesce into a compound.’ (Oliphant 1912: 46).

[5] Some other Indo-European languages have dvandva compounds, but they do not share the formal peculiarities of the Vedic ones, e.g. OE. *suhter(ge)fæderan* ‘uncle and nephew’, *aþumsweoran* ‘son-in-law and father-in-law’ (*Beowulf*), OIr. *gaisced* ’spear and shield’, ‘weapons’. I assume that they arose independently of the Sanskrit dvandvas.

[6] This is not to say that there might not be evidence for global complexity forthcoming, perhaps in the form of evidence for tradeoffs between phonology and syntax or other components of the grammar, or abstract arguments from philosophical theories and/or mathematical models of inductive inference based on Minimum Description Length (Kolmogorov Complexity, see Li and Vitányi 1997).

[7] Only one other, smaller group of putative compounds, also confined to the Vedic language, has a case ending and an accent on each member: determinative *(tatpuruṣa)* compounds with a genitive noun as first member, comprising names of gods such as *bṓ́haspáti* ‘Lord of Prayer’ (Bṛhaspati), *gnáspáti* ‘Husband of the Goddess’ (Agni). Such compounds can themselves be form dvandvas with another noun, in which case the result is multiply accented: *índr ābṓ́haspát ī* (4.49.5) ‘Indra and Bṛhaspati’. Although I do not deal with doubly accented determinative compounds here, I believe that some aspects of my morphological analysis carry over to them.

[8] Most examples cited in this paper are from the oldest Vedic text, the Rigveda, with the line number given in the usual style. Citations from later Vedic literature come from the Atharvaveda (AB.), the Taittirīya Saṃhitā (TS.) and the Śatapatha-Brāhmaṇa (ŚB.), and are explicitly identified as such. After these works, no new instances of Vedic-type dvandvas appear, and the old ones are almost all in quotes from the earlier literature. Therefore, the chronological layers of the language from this point on need not be distinguished for our purposes, and I will refer to them collectively as post-Vedic or classical Sanskrit.

[9] Thus 'tmesis' is a misnomer for such cases, for they cannot result literally from the splitting of dvandvas. The analysis of tmesis by clitics is not so clear-cut. It is also attested three times in determinative compounds with an inflected noun as first member mentioned in the preceding footnote: *śúṇaś cic chépam*, from the proper name *śúṇaḥśépa* 'Dog's Tail' (or maybe 'Dog's Penis') (5.2.7), and *nár ā ca śámsam* (9.86.42), *nár ā v ā śámsam* (10.64.3), from the proper name *nár āśámsa (< *nár ām-śámsa)* ('Praise of Men') (Agni). It constitutes less reliable evidence for wordhood than splitting by full words, for it could be explained as phonologically driven endocliticization (documented for Pashto by Tegey 1977, see Roberts 1997, Anderson 2005, Kopris & Davis MS.), especially if the constituents of dvandvas are prosodic words, as I argue they are. Insler 1998 dismisses the few instances of clitic intrusion into determinative compounds as a poetic artifice, while taking the robustly attested tmesis of dvandvas, which is often of the more radical type seen in (8), as a grammatical datum which proves their unique two-word status.

[10] The accentual contrast is somewhat blurred by occasional instances where even asyndetically coordinated singular vocative phrases are treated as a single unit, e.g. *bṛ́haspata indra* (4.50.11) 'Bṛhaspati, Indra!'. However, such cases are rare, and there are no exceptions in the other direction. Thus, the accentual criterion does draw a distinction between dvandvas and co-ordinated phrases.

[11] It is a process of the lexical phonology; as argued by Kiparsky 1984 for Vedic, and Kiparsky 2003 for Greek.

[12] As can regular dvandvas of the classical Sanskrit type: *kṣatra-viṭ-śūdrāḥ* (the three non-Brahmin castes), *eco 'y-av-āy-āv-aḥ* '*eC* is replaced by *ay, av, āy, āv̀* (Pāṇini 6.1.78).

[13] Contrast phrasal conjunction, e.g. *kṣápa usrā́* (6.52.15) 'nights and mornings', and classical dvandvas, e.g. *dadhi-ghṛ́tam* 'yoghurt and ghee'.

[14] There are just two cases where both orders are attested in Vedic: *dyáv āpṛthiví* (65x) ~ *pṛthivídyáv ā* (1x), and *náktoṣás ā* (5x) ~ *uṣás ānákt ā* (8x). Another instance of reversal, *v āt āparjanyà* (10.66.10, AV. 10.4.16, pro *parjány āvát ā*) is chronologically a later form than *parjány āvát ā*, and represents the innovative later type of dvandva with a single accent, discussed in section 3 below.

[15] For the order of members in classical Sanskrit dvandvas see Pāṇini's rules 2.2.32-34, confirmed by Wackernagel 1905:165 ff.; similar generalizations fix the order or members in modern Greek verbal dvandvas (Kiparsky 2009) and in the much studied 'irreversible binomial' phrases (for English see most recently Benor and Levy 2006).

[16] The reanalysis of a case ending to a stem-forming suffix in compounds is common. For example, the *-es* in German *Liebesbrief* 'love letter' is etymologically the Genitive Singular ending, but synchronically builds a compositional stem; the Genitive Singular of the noun is actually *Liebe*.

[17] Yet another type of hybrid category, probably not relevant for the dvandva problem, straddles traditional word classes such as verbs and nouns; Vedic examples might be the deverbal nouns that are verblike in assigning accusative or some other oblique case to their complements.

[18] For independent evidence that the single-accent constraint applies to phonological words, see Kiparsky and Halle 1984, where it is subsumed under the 'Basic Accentuation Principle' (BAP).

[19] That is, they bear the associative dual stem suffix, on my analysis. On the traditional analysis, they just have the 'wrong' case. Either way, there can be no question of syntactic co-ordination here.

[20] The spaces in romanized transcriptions have no counterparts in the Sanskrit recited or written text.

[21] The full story also covers semantic blocking; see Kiparsky 2005 for illustration with Sanskrit examples. Wunderlich makes a number of further assumptions, which together define the theory that he calls Minimalist Morphology. These additional assumptions are also not required here. Any theory of morphology which is lexicalist and which treats blocking as a relation between expressions will serve equally well.

[22] For instances of such systematic blocking in various parts of Vedic morphology, see Kiparsky 2005.

[23] Usually, it turns out that morphemes are combined freely subject only to general constraints on word structure. For example, affixes can be added whenever their feature content unifies with the feature content of the base and directionality requirements (represented by alignment constraints or perhaps in some other way) are satisfied.

[24] The former is excluded because the reduplicant would contract with the root and the result would be identical to the root itself. The latter is excluded because perfect reduplication is an operation on roots and not on stems, as independently attested by the phonological shape of the reduplicant.

[25] Pāṇini's grammar actually makes this preference explicit (by putting compounding under the major heading *vā* 'preferably'). Also, he lists certain compounds as being *nitya* 'obligatory' See Kiparsky 1979 for Pāṇini's treatment of variation and for evidence from Sanskrit usage that corroborates it. Note that there are cases where only the compound is grammatical, and cases where the compound is ungrammatical. The generalization concerns cases where both options are grammatical.

[26] Poetic artifice involving variation and parallelism may also play a role. It may be behind the pattern of variation in the hymn 7.66 from which this example is taken: the two gods are introduced as a pair with the syntactic conjunction *mitráyor váruṇayoḥ* in verse 1, then addressed individually in the singular (verses 3, 4, 7, 9, 11, 12, 17, 18) and finally together again in a dual co-compound (v. 19).

[27] A form like (32c) can only be derived by suffixation of *-yā* to the compounded root *sam+rabh*, not by prefixation of *sam* to the suffixed root **rabh-ya*. The most straightforward reason is that the bare root requires the allomorph *-tvā* (e.g. *labh-tvā* → *lab-dhvā* 'taking'), so the correct affix can only be selected after the prefix is already in place. See Kiparsky 2007 for the arguments in detail.

[28] Blocking is applicable also to combinations of so-called separable prefixes and verbs in German. It says that in a clause such as *wenn der Zug ankommt* 'when the train arrives', *an + kommt* 'arrives' is a single word (as the orthography writes it); the two-word parse that the syntax

provides for it is blocked. This agrees with the conventional one-word spelling, but I am not aware of any phonological evidence pro or con.

[29] E.g. *pitárau* means ‘parents’ in 1.121.5 and 10.131.5, and ‘two fathers’ in 10.85.14; in 10.115.1 *mātárau* means ‘two mothers’.

[30] Except for one (or possibly two) attested instances: *abhí samrā́jo* ***váruṇo*** *gṛ́ṇanty abhí* ***mitrā́so aryamā́*** *sajóṣāḥ* (7.38.4) ‘the united mighty rulers join in the song of praise, Mitra with Varuṇa and Aryaman’, literally ‘the Mitras (Plural), Varuṇa, and Aryaman’, and one other less clear case (Edgerton 1909).

[31] Following usual practice in Indo-European linguistics, I distinguish the reconstructed ‘laryngeal’ consonants with numerical subscripts to indicate their different vocalization and coloring properties; -h_1 is vocalized to *e*.

[32] See also the description of the Bunuba associative in Rumsey 2000:62-63 and Singer 2001:56. Other likely cases of comitative endings evolving into associative endings in co-compounds are cited in Wälchli 2005:249.

[33] Hans Heinrich Hock further draws attention to the associative instrumental construction RV. *yudhā́ yudham* ‘fight by fight’, (1.53.7), *purā́ puram* ‘fort by fort’, which Hoffmann 1960 connects with the *āmreḍita* (adverbial) forms *men āmenam, dhur ādhuram* found in the Jaiminīya-Brāhmaṇa, unfortunately of uncertain meaning.

[34] RV. *indrav āyú* (Dual) ‘Indra and Vāyu’ would be the earliest instance of the type, but it is disputed. Arnold (1905:123) argues on metrical grounds that ‘*indr āv āyú* must in all instances be restored’, because ‘it never stands at the end of a Triṣṭubh verse, or in any other position in which *a* is favored.’ On the other hand, Insler 1998 and van Nooten & Holland 1994 reject Arnold’s emendation and favor the usual reading *indrav āyú*.

References

A , S R. 2005. *Aspects of the theory of clitics.* Oxford: OUP.

A , A . MS. Gradient phonotactics in Optimality Theory. Stanford University.

A , E.V. 1905. *Vedic metre in its historical development.* Cambridge: CUP.

B , S B R L . 2006. The chicken or the egg? A probabilistic analysis of English binomials. *Language* 82.233-27.

B ´ -O , R . 1999. *Constraint interaction in language change: quantity in English and Germanic.* Ph.D. Dissertation, University of Manchester.

B ´ -O , R . 2007. Diachronic phonology. *The Cambridge handbook of phonology*, ed. by Paul de Lacy, Cambridge: CUP.

C , A . 1987. Allomorphy in inflexion. Beckenham: Croom Helm.

C , G. N. 1985. The problem of transfer in nonlinear morphology. Cornell Working Papers in Linguistics, Vol. 7.

F , J . 2004. Distributed reduplication. MS, MIT.

C , M . 2003. *The paradigmatic structure of person marking.* Oxford: OUP.

D , M . 2005. Understanding inclusives. *Clusivity*, ed. by E. Filimonova. Amsterdam: Benjamins.

D , R.M.W. 1972. *Dyirbal.* Cambridge: CUP.

E , F . 1909. Origin and development of the elliptic dual and of dvandva compounds. *Zeitschrift für Vergleichende Sprachforschung* 43.110-20.

E , D . 2007. Blocking effects and analytic / synthetic alternations. *Natural Language and Linguistic Theory* 25.1–37.

E , D , A M . 2008. Architecture and blocking. *Linguistic Inquiry* 39.1–53.

E , B . 2007. *Asking Questions: language variation and language acquisition.* Ph.D. Dissertation, Stanford iversity.

F , M . 2000. Der urindogermanische Dual — eine Klasse für sich. *125 Jahre Indogermanistik in Graz*, ed. by Michaela Ofitsch and Christian Zinko, 133–137. Graz: Leykam.

G , E . 2004. *Grammaticalization as Economy.* Amsterdam: John Benjamins.

H , E . 1994. *Prosodic structure in compounds.* Ph.D. Diss., Stanford.

H , J L M . 2005. When movement must be blocked: A reply to Embick and Noyer. *Linguistic Inquiry* 36.85-125.

H , M . 2004. On directionality in language change with particular reference to grammaticalization. *Up and down the cline: The nature of grammaticalization,* ed. by Olga Fischer, Muriel Norde and Harry Perridon, (Typological Studies in Language, 59.17-44). Amsterdam: Benjamins.

H , J A. 2004. *Efficiency and complexity in grammars.* Oxford: OUP.

H , K . 1960. Der vedische Typus *men āmenam*. *Zeitschrift für vergleichende Sprachforschung* 76.242-248. (Repr. 1975, in Karl Hoffmann, *Aufsätze zur Indoiranistik*, ed. by J. Narten, Wiesbaden: Reichert.)

I , S . 1989. *Prosodic constituency in the lexicon.* Ph.D. Diss., Stanford.

I , S , D Z . 1990. *The Phonology-Syntax Connection.* Chicago U.P. and Center for the Study of Language and Information.

I , S . 1998. *mitráváruṇā* or *mitrá váruṇā*? *Mír Curad: Studies in Honor of Calvert Watkins,* ed. by Jay Jasanoff, H. Craig Melchert, and Lisi Olivier. Innsbruck.

K , J . 1978. Rule mitosis: The historical development of Algonquian palatalization. *Linguistic studies of Native Canada,* ed. by Eung-Do Cook and Jonathan Kaye, 143-56. Vancouver: University of British Columbia Press.

K , P . 1979. *Pāṇini as a variationist*. Cambridge, Mass, and Poona: MIT Press and Poona University Press.

K , P . 1984. The Lexical Phonology of Sanskrit word accent. *Amṛtādhāra: R.N. Dandekar Felicitation Volume,* ed. by S.D. Joshi. Delhi: Ajanta Publications.

K , P . 2000. Opacity and Cyclicity. *The Linguistic Review* 17.351-367.

K , P . 2003. Accent, syllable structure, and morphology in Ancient Greek. *Selected Papers from the 15th International Symposium on Theoretical and Applied Linguistics,* by Elizabeth Mela Athanasopoulou. Thessaloniki.

K , P . 2005. The Vedic injunctive: Historical and synchronic implications. *The Yearbook of South Asian Languages and Linguistics.*

K , P . 2007. Reduplication in Stratal OT. Linda Uyechi and L.H. Wee (eds.), *Reality Exploration and Discovery: Pattern Interaction in Language & Life*. Stanford: CSLI Press. http://www.stanford.edu/~kiparsky/Papers/reduplication.pdf

K , P . 2009. Verbal co-compounds and subcompounds in Greek. *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT. MIT Working Papers in Linguistics #57*, ed. by C. Halpert, J. Hartman and D. Hill. http://www.stanford.edu/~kiparsky/Papers/greekcocompounds.pdf

K , P . In press. Grammaticalization as optimization. *DIGS VIII Proceedings*, ed. by Andrew Garrett and Diane Jonas. Oxford: OUP. http://www.stanford.edu/~kiparsky/Papers/yalegrammaticalization.pdf

K , P M H . 1977. Towards a reconstruction of the Indo-European accent, in *Studies in Stress and Accent. Southern California Occasional Papers in Linguistics* Vol. 4, ed. by L. Hyman, 209–238. Department of Linguistics, University of Southern California, Los Angeles.

K , J S. 1985. *Toward a discourse grammar of the Rigveda. Volume 1: Co-ordinate conjunction.* Heidelberg: Winter.

K , C A. & A R. D . MS. Endoclitics in Pashto: Implications for lexical integrity. Presented at the Fifth Mediterranean Morphology Meeting (MMM5). Fréjus, 15-18.9 2005.

L , G.L. l967. *Turkish grammar.* Oxford: Clarendon.

L , M P V ´ . 1997. *An introduction to Kolmogorov Complexity and its applications.* Second Edition. Heldelberg: Springer.

L , K . 1973. Das Götterdvandva im Slavischen. *Scando-Slavica* 19.249-53.

M D , A.A. 1910. *Vedic grammar.* Strassburg: Trübner.

M , M . 2000. Die Genese des idg. Numerus Dual. Ofitsch, Michaela and Christian zinko (edd.) *125 Jahre Indogermanistik in Graz*, 291-315. Graz: Leykam.

M -B ¨ , M . 2002. *Indogermanische Sprachwissenschaft*[8]. Berlin: de Gruyter.

M , E . 2003. A semantic analysis of associative plurals. *Studies in Language* 27. 469-503.

N , B. & G. H . 1994. *Rig Veda, a metrically restored text.* Cambridge, MA: Harvard University Press.

O , S G . 1912. The Vedic Dual: Part VI, The Elliptic Dual; Part VII, The Dual Dvandva. *Journal of the American Oriental Society* 32.33-57.

P , W J. 1992. Blocking of phrasal constructions by lexical items. *Lexical Matters*, ed. by Ivan Sag and Anna Szabolsci, 111–130. CSLI, Stanford, CA.

P , J . 1977. Devatā-dvandva in Hittite, Greek, and Latin. *American Journal of Philology* 98.396-405.

R , L . 1952. *Grammaire de la langua védique.* Paris: Champion.

R , T . 1997. The optimal second position in phonology. *Phonology in progress — progress in phonology,* ed. by Geert Booij and Jeroen van de Weijer. The Hague: Holland Academic Graphics.

R , I . 2007. *Diachronic Syntax.* Oxford: OUP.

R , I , A R . 2003. *Syntactic change: A Minimalist approach to grammaticalization*. Cambridge: CUP.

R , A . 2000. Bunuba. *Handbook of Australian languages*, Vol. 5, ed. by R.M.W. Dixon and B.J. Blake. Oxford: OUP.

R , K . 2006. The Ṛg-Vedic fate of the Devatā Dvandva (Double Dual). MS, UCLA.

S , R . 2001. *The inclusory construction in Australian languages.* Honours thesis, Linguistics and Applied Linguistics, University of Melbourne. http://eprints.infodiv.unimelb.edu.au/archive/00003035/

S , R W. 1985. *On deriving the lexicon.* Ph.D. dissertation, MIT.

T , H . l977. *The grammar of clitics: Evidence from Pashto and other languages.* Kabul: International Center for Pashto Studies.

W , J . 1905. *Altindische Grammatik,* II,1. Göttingen: Vandenhoeck & Ruprecht.

W¨ , B . 2005. *Co-compounds and natural coordination.* Oxford: OUP.

W , C . 1995. *How to Kill a Dragon. Aspects of Indo-European Poetics*. Oxford: OUP.

W , D . 1996. Minimalist morphology: the role of paradigms. *Yearbook of Morphology* 1995, 93-114, ed. by Geert Booij & Jaap van Marle. Dordrecht: Kluwer.

Ž , O.F. 2002. Raba dvojine v zvezah z dvema imenoma. Dvojna dvojina. *Slavistična revija* 50.269-275.