Uta Reinöhl*

Continuous and discontinuous nominal expressions in flexible (or "free") word order languages: Patterns and correlates

Abstract: This study explores continuous and discontinuous word order patterns of multi-word nominal expressions in flexible word order languages (traditionally referred to as "free word order" or "non-configurational" languages). Besides describing syntagmatic patterns, this paper seeks to identify any functional or other correlates that can be associated with different word orders. The languages under investigation are a number of Australian languages as well as Vedic Sanskrit, all of which have long been known for their syntagmatic flexibility. With respect to continuous order, evidence from several of these languages suggests that default ordering is primarily governed by functional templates. Deviations from default order, while maintaining continuity, can be attributed to different types of "focus" interpretations or heaviness effects. With respect to discontinuous order, I identify three sub-types. The most widespread one, "Left-Edge Discontinuity", involves one element placed in or near utterance-initial position. It shows a clear, if not an absolute, correlation with different kinds of focus interpretations, similarly to deviations from the default order in continuity. The other two types of discontinuity are linked to the behaviour of specific function words. Besides teasing out cross-linguistic similarities, this paper also sheds light on language-specific characteristics that affect the forms and functions of complex (i.e. multi-word) nominal expressions in flexible word order languages, such as the nature of 2nd position ("Wackernagel") elements.

Keywords: noun phrase, free word order, flexible word order, discontinuity, information structure, functional template, Wackernagel position, focus, non-configurationality, Australian languages, Vedic Sanskrit

1 Introduction

This paper discusses word order patterns of complex (i.e. multi-word) nominal expressions in flexible word order languages (also referred to as "free word

*Corresponding author: Uta Reinöhl [ʔuːtʰa ʁaɪnʔøːl], Universität zu Köln, Köln, Germany; Johannes Gutenberg-Universität Mainz, Mainz, Germany, E-mail: uta.reinoehl@uni-koeln.de

order" or "non-configurational" languages), as well as any functional or other correlates that can be associated with the different ordering patterns. While it seems that most languages around the world allow for the permutation of main constituents under specific conditions, flexible word order languages, well-known especially from the Australian continent, also show considerable variability with regard to word order within nominal expressions (e.g. Hale 1983; Heath 1986; Austin & Bresnan 1996; Louagie & Verstraete 2016). This variability involves both flexible ordering when the elements in question are placed in adjacency (i.e. when we are dealing with "continuous nominal expressions") and discontinuous placement (i.e. the phenomenon of "discontinuous nominal expressions" or, traditionally, "discontinuous NPs"). Notably, in these languages, such alternative ordering patterns do not seem to necessarily or obviously affect either the unit-hood of the nominal expression in question or internal "headedness" relations. While this ordering flexibility is relatively well-known, a study that comprehensively surveys both continuous and discontinuous patterns as well as any functional or other correlates is lacking so far.

This paper thus examines syntagmatic variation in languages that push the ordinary boundaries of word order flexibility. Discontinuous placement is perhaps the most radical instantiation of this flexibility and is probably the typologically most unusual of the features associated with the label of "non-configurationality". I will focus not only on discontinuity in this paper, however, but also spend some time on continuity, as adjacent ordering also shows greater flexibility in the languages discussed here in comparison to, e.g. English or French, with their much more tightly structured phrasal syntagms. Throughout the paper, the focus is on the ordering of non-phrasal elements, as phrasal structures inside complex nominal expressions (e.g. adnominal possessor phrases or relative clauses) generally seem to enjoy considerable word order freedom, whether or not a language otherwise displays particularly flexible word order. The non-phrasal elements in question are, in particular, "nominals", which may be entity- and/or property-denoting, given the absence of a clear-cut noun-adjective distinction in the languages at issue here. This last point is a crucial part in accounting for the flexibility found, as will be outlined in Section 2. Besides nominal elements, I also take into account the usual other types of non-phrasal elements figuring in nominal expressions, such as demonstratives, quantifiers and numerals.

Languages for which it has been claimed that word order variation in complex nominal expressions does not necessarily affect their unit-hood or "headedness" relations include at least the 29 Australian languages listed as showing syntactically "unrestricted" word order flexibility in Louagie and Verstraete (2016: 63–64, Table B-1 (iii)). Further languages that are relevant in the context of this paper are those categorised by Louagie and Verstraete as showing "flexibility of

adjective-like modifiers" while "determining elements [are] fixed at one edge" (Louagie & Verstraete 2016, Table B-1 (iic)) and at least some of the 18 languages for which the primary source indicates that flexibility is "restricted in frequency" (Louagie & Verstraete 2016, Table B-1 (iia)). With regard to the last-mentioned category, note that deviations from the default order seem to be comparatively infrequent in most, if not all, languages. Thus, this category reflects perhaps not so much a cross-linguistic difference, but rather whether or not an author comments on frequency. Furthermore, several other Australian languages mentioned in Louagie and Verstraete (2016) qualify, even though they are not categorised there as showing "flexible" word order, including Gooniyandi[1] (McGregor 1990, McGregor 1997), Nunggubuyu (Heath 1986), and possibly also Tiwi (Lee 1987). The reason is that, whenever the primary source mentions any "functional" correlate as influencing ordering, the language is excluded from the category of "flexibility" in Louagie and Verstraete's (2016) categorisation. A "functional" correlate may involve, e.g. an information-structural factor such as some sort of a focus interpretation. I will argue in this paper that factors of this sort play a role in most cases where ordering deviates from the default. Therefore, the fact that information-structural or other correlates can be identified should not be taken as a reason for excluding languages from categorisations such as "flexible word order", if this label is understood to be a syntactic one.

Outside of the Australian continent, a language that also shows a significant degree of syntagmatic flexibility within nominal expressions as well as other properties that have been associated with the label "non-configurationality" is Vedic Sanskrit (Reinöhl 2016: 33–42), the oldest attested variety of Indo-Aryan. Accordingly, this language is also included in the present study. Note that the selection of languages is not so much due to this kind of flexibility only being attested here, but rather to the fact that we are dealing with a topic still very much in need of exploration. I hence focus on languages for which this phenomenon has already been studied in some detail, which applies in particular to certain Australian languages, given their central role in the history of the non-configurationality debate.

While the number of potentially relevant languages is significant, this paper is in part exploratory and no attempt has been made to explicitly discuss all of

---

1 List of ISO-Codes & Glottocodes (Language; ISO-Code 639–3; Glottocode) of all languages mentioned, in the order of mention: English; eng; stan1293, French; fra; stan1290, Gooniyandi; gni; goon1238, Nunggubuyu; nuy; nung1290, Tiwi; tiw; tiwi1244, (Vedic) Sanskrit; san; sans1269, Jaminjung-Ngaliwurru; djd; djam1255, Ancient Greek; grc; anci1242, Latin; lat; lati1261, Polish; pol; poli1260, Warlpiri; wbp; warl1254, Jaminjung; djd; jami1236, Tagalog; tgl; taga1270, Warrongo; wrg; waru1264, Bardi; bcj; bard1255, Kayardild; gyd; kaya1319, Wardaman; wrr; ward1246, Thargari; dhr; dhar1247, Walmajarri; wmt; walm1241, Wangkajunga; mpj; wang1288, Ngiyambaa; wyb; wang1291.

the above-mentioned Australian languages. Determinants of main constituent order have received some attention for Australian languages (see Simpson & Mushin 2008: 26–27 for references) and for Vedic (e.g. Delbrück 1878; contributions in Hock 1991; Schäufele 1991b; Lühr 2015), but ordering within nominal expressions has been the primary focus of few publications. Thus, this study presents a first qualitative survey of patterns and correlates, drawing particularly on McGregor (1990, 1997) on Gooniyandi and Schultze-Berndt & Simard (2012) on Jaminjung. Since Vedic Sanskrit is the only language outside of Australia included in this study, and since its nominal syntax has not attracted much attention, this paper also includes a fair number of Vedic examples. These stem predominantly from a study of ca. 1000 clauses from the early prose text *Maitrāyaṇī Saṁhitā* (around 900 BC, Amano 2009: 1) analysed for this paper, a commentary of early Hindu ritual practices (see Acknowledgements).

Note that the claims in the present paper do not, or do not necessarily, extend to other Australian languages outside of the ones mentioned, nor to other early varieties of Indo-European besides Vedic Sanskrit. The latter has to do with the fact that most early representatives of Indo-European (e.g. Ancient Greek or Latin) show more syntactically rigid or "phrasal" nominal expressions, which can be linked to the existence of grammaticalised adnominal function words such as articles or adpositions (Himmelmann 1997; Reinöhl 2016).[2]

The evidence discussed in this paper, first with regard to continuity (Section 3) and then to discontinuity (Section 4), suggests that there are several cross-linguistic similarities between word order patterns as well as their correlates. In Section 3, I argue that in most, if not all, languages, a particular continuous order is the default, which in numerous languages appears to be primarily governed by a functional template. Such functional templates contain slots for, e.g. "qualifying", "classifying", "entity-denoting" and other roles (McGregor 1990), which can be flexibly occupied by elements, provided they are semantically compatible. Deviations from the default while maintaining continuity can be associated with different sorts of focus interpretations or with heaviness effects. In Section 4, I identify three types of discontinuity. In "Left-Edge Discontinuity", the most widespread one across the languages studied here, an element belonging to a complex nominal expression is placed at or near the left edge of the clause, while the other element(s) follow(s) in third position or later. In "Wackernagel Pronoun Discontinuity", a pronominal element that is part of a complex nominal expression is placed in 2nd

---

2 Still, some of the phenomena discussed here may also be attested in those languages (e.g. Devine & Stephens 2000 on discontinuity in Ancient Greek). The same applies to modern languages that display discontinuity, but otherwise have a relatively rigid structure of nominal expressions, such as Polish (Siewierska 1984).

("Wackernagel") position and thus away from the other element(s) it forms an expression with. "Conjunction-related Discontinuity", specific to Vedic, involves conjunctions separating elements forming a complex expression. As with deviations from the default order while maintaining continuity, an important cross-linguistic correlate of discontinuous patterns is focus interpretations, in particular in the case of Left-Edge Discontinuity. Section 5 at the end of this paper discusses in more detail the different types of focus interpretations associated with the various non-default word order patterns. Throughout the paper, I highlight how language-specific characteristics, such as the behaviour of 2nd position elements, act as constraints on forms and functions. The aim is to tease out which aspects are particular to language, language family or area, and which ones may be cross-linguistically common. In a few cases, genre effects will be mentioned, such as regarding specifics of the historical Vedic corpus in comparison to the different types of fieldwork data collected for the Australian languages.

2 Identifying complex nominal expressions in flexible word order languages

Before examining the different patterns and their correlates, it is necessary to discuss whether and how it is possible to identify complex nominal expressions in flexible word order languages, a point less than straightforward given the formal characteristics of such languages.

I assume in this paper, in line with numerous studies since Hale (1983), that two or more elements may form a complex nominal expression despite a lack of the obligatory syntactic integration that is normally associated with the label "noun phrase". Note that, since the languages discussed here lack such integration, the term "noun phrase" is avoided in this paper following Himmelmann (1997), and preference is given to "nominal expression", which is neutral with regard to whether or not an expression is phrasally structured. The justification for assuming that unit-hood as a (complex) nominal expression is indeed a relevant category in the languages at issue here is that there are clear differences between cases where elements work together in the same utterance to identify a participant or establish a concept, and cases where they do not, even if they share the exact same morphological marking.[3] In cases where several elements

---

3 In the absence of word order as a reliable criterion, shared morphological (in particular, case) marking figures prominently as an indicator for elements forming complex nominal expressions in analyses of Australian languages in the framework of Lexical-Functional Grammar (e.g. Simpson

work together in that way, we find recurring ordering patterns and recurring correlates, as will be discussed in Sections 3–5.

Even when there is a lack of formal clues, Himmelmann (1997: 117–119) argues, based on work by John Du Bois and Sandra Thompson, that a complex nominal expression can be identified as such on the grounds of discourse-referential uses. These include "tracking" uses of discourse participants (roughly, "referential" uses) as well as "non-tracking" uses (often called "non-referential" uses) as found, e.g. in predicative function. Two or more elements that together establish a concept, whether appearing in a "tracking" or "non-tracking" use, and show some formally nominal properties, are considered to form one complex nominal expression. This discourse-based understanding is adopted here.

However, it is important to note that the lack of formal clues often makes it difficult to decide in practice whether we are dealing with a real nominal expression in a discourse-functional sense. The challenge of deciding whether nominals form a complex expression or not has been addressed in Hale (1983) for Warlpiri (Pama-Nyungan). Hale, focusing on discontinuous cases, discusses the difference between what he calls "merged" and "unmerged" interpretations:

(1)[4]   I take it to be a fact of Warlpiri that a sentence like ... below is open to at least two interpretations:

Maliki-rli ø-ji   yarlku-rnu wiri-ngki
dog-ERG PERF-1OBJ bite-PST   big-ERG
'The/a big dog bit me.'
'The/a dog bit me and it was big.'

> ... On one interpretation, the expression *wiri-ngki* (big-ERG) is taken as a modifier of *maliki-rli* (dog-ERG), forming with it a unit expression ... This is the 'merged' interpretation. ... On another interpretation ..., *wiri-ngki* is simply predicated of *maliki-rli*. (Hale 1983: 38)

1991; Nordlinger 1998; Austin & Bresnan 1996, among others). However, it is important to be aware of the fact that shared case marking is not a sufficient criterion, as will become clear in this section.
4 Glossing abbreviations: A = transitive subject, ABL = ablative case, ABS = absolutive case, ACC = accusative case, COM = comitative case, CONJ = conjunction, DAT = dative case, DEM = demonstrative, DU = dual, ERG = ergative case, EVID = evidential marker, F = feminine, FOC = "focus" marker, FUT = future, GEN = genitive case, IMP = imperative mood, INC = inclusive, INS = instrumental case, LOC = locative case, LOT = large group (number), M = masculine, MAN = manner case, MID = middle, MODAL_ROOT = modal root, N = neuter, NEGIMP = negative imperative mood, NOM = nominative case, NONPAST = non-past, OBJ = object, OBL = oblique pronominal, OPT = optative mood, PASS = passive voice, PERF = perfective aspect, PL = plural, PRS = present, PRT = discourse particle, PST = past tense, REDUP = reduplicated, REMPST = remote past, REFL = reflexive, REL = relative pronoun, REP = repetitive aspect, S = subject, SEQ = sequential clitic, SG = singular, SUBJV = subjunctive, TOP = topic, TR = transitive, VBLZR = verbalizer.

The reasons why it can be difficult to decide whether elements form one nominal expression, i.e. allow for a merged interpretation, include word order flexibility, the absence of a clear-cut noun-adjective distinction, as well as the ubiquitous use of nominal forms in a wide range of grammatical roles.[5] The remainder of this section discusses the delimitation of real nominal expressions from apparent ones in more detail, given the discourse-functional understanding adopted above (see Reinöhl To appear, for a more detailed discussion).

First of all, cases where the elements in question are not found within the same "domain" are excluded from consideration (see Schultze-Berndt & Simard 2012). For instance, free topics in the left periphery preceding a clause are not considered to form complex expressions with nominals inside the clause, and the same applies to nominals in the right periphery, e.g. afterthoughts or right dislocations.[6] A problem is, of course, that it can be difficult to establish domain boundaries in flexible word order languages if one relies solely on syntactic evidence, in particular when dealing with discontinuous cases.[7]

In order to identify domain boundaries, Schultze-Berndt & Simard (2012) argue for Jaminjung that the prosodic domain of the "intonation phrase", intermediate between prosodic word and intonation unit, delimits boundaries within which elements form complex expressions, whether continuous or discontinuous. The importance of prosodic phrasing for delimiting information units has long been recognised, going back to work for instance by Chafe (e.g. 1994). Ex. (2) from Jaminjung illustrates a discontinuous nominal expression found inside an intonation phrase. The authors point out that the dog was already introduced into the conversation and is now referred to more specifically as 'a *dangerous* dog'. An interpretation of 'fierce' or 'dog' as, e.g. a free topic or an afterthought, seems implausible in this context. By contrast, ex. (3) does not involve one complex expression, but two separate ones distributed

---

5 In an early influential line of research, the very possibility that two or more elements could form a complex nominal expression in "non-configurational" languages like Warlpiri did not exist, namely in Jelinek's (1984) "Pronominal Argument" analysis and similar approaches such as in Laughren (1989) or Speas (1990). In Jelinek's (1984) approach, focusing on Warlpiri, only clitic pronominal elements in 2nd position fill argument roles, while any other nominal or free pronominal elements are individual adjuncts to those clitic pronouns. This line of argument was however soon debunked based on a number of deficits in accounting for facts of Warlpiri and flexible word order more generally (e.g. Simpson 1991; Austin & Bresnan 1996; Austin 2001).

6 I do not discuss constructions where there are morphological indications that the nominals in question do not form a complex expression, as in the case of the "part-whole" constructions mentioned in Schultze-Berndt and Simard (2012: 1030–1032). In other cases, nominals may be analysed as forming a complex expression despite a difference in case marking, a point I will also not further discuss here.

7 Thus, there is typically no evidence as in English, where function words may have to be repeated when occurring in separate domains (e.g. English *I saw **the** car, **the** red one.*).

across two intonation phrases, as shown by the backslash \. The second expression serves as an afterthought, adding information on the already introduced referent.

(2) Jaminjung (non-Pama-Nyungan)
     ^***mulanggirr*** *ngantha-ma-ya*   ***wirib*** \
     fierce   2SG>3SG-have-PRS dog
     'You have a ***dangerous* dog**!' (Schultze-Berndt & Simard 2012: 1035)

(3) Jaminjung (non-Pama-Nyungan)
     buyud*=biyang jabul=ni*   *burr-angu=rrgu=rndi* \
     sand=SEQ   shovel=ERG/INS 3PL>3SG-get/handle.PST=1SG.OBL=EVID
     buj-mawu   ***buyud*** \
     bush-dweller sand
     'They got **sand** for me with a shovel, **the bush kind of sand**.'
     (Schultze-Berndt & Simard 2012: 1027)

McGregor (1997) on Gooniyandi also takes into account prosodic phrasing (marked by slashes). The next example illustrates a right dislocation where a referent is at first expressed only pronominally, while lexical information is provided later in the following prosodic unit (cf. also Baker & Mushin 2008: 10–11):[8]

(4) Gooniyandi (non-Pama-Nyungan)
     goobardiya / *ngila* / *joonyoo* / ***niyaji-ya*** *yard ngaragbinmi* /
     Goobardiya   east   Joonyoo this-LOC   yard they:made:it
     mayari-ya /
     house-LOC
     'At Goobardiya and east at Joonyoo, **there** they made yards, **at the home-steads**.' (McGregor 1997: 101)

Examples of nominals in the left periphery include, as already mentioned, free topics as illustrated in ex. (5) from Jaminjung; in this case, the authors make reference to the pitch contour to underscore their analysis.

(5) Jaminjung (non-Pama-Nyungan)
     *[jarlig=biya]*_{TOP} *[bardawurru]*_{FOC} *gani-ma-ya* \
     child=SEQ   many   3SG>3SG-have-PRS
     '(As for) children, she has many ....' (Schultze-Berndt & Simard 2012: 1048)

---

8 McGregor (1997), however, does consider examples such as ex. (4) to be a type of "discontinuity" in contrast to the analysis adopted here. For a critique of McGregor's approach, see Schultze-Berndt and Simard (2012: 1028).

Another type that has been discussed in the literature (and which is even considered synonymous with discontinuity in Fanselow & Féry 2006: 2, fn.1) is split topicalisation. Here, the elements involved bear different information-structural roles and do not together identify one concept, and they are realised in separate prosodic domains, e.g. German **Zeitungen** *liest er nur eine, die* **TAZ** ('As for newspapers, he only reads one, the TAZ' (quoted from Fanselow & Ćavar in Schultze-Berndt & Simard 2012: 1047). Therefore, this type is also excluded here as only apparently involving complex nominal expressions following Schultze-Berndt & Simard (2012: 1047–1048).

Prosodic analysis is still somewhat rare for Australian languages, and there is of course no recorded evidence available for Vedic Sanskrit. However, certain prosodic clues make it possible to delimit a core domain in Vedic, as well, at least in certain clause types (see Reinöhl To appear). The left domain boundary is often clearly delimited by unaccented and/or accented Wackernagel elements. These require a single preceding, accented word as their host, which in turn marks the left clausal edge. At the right edge, we normally find the finite verb, which is unaccented in main clauses, carrying the lowest pitch sentence-finally.[9] The following example illustrates a case where two nominals follow the finite verb. While sharing the same inflectional marks, *ádābhyāḥ* likely functions as an afterthought, rather than forming a discontinuous nominal expression with *bṛhadúkṣaḥ marútaḥ viśvávedasaḥ* (see remarks in Delbrück 1878; Staal 1967; Schäufele 1991b).[10]

(6)   Vedic Sanskrit

bṛhadúkṣaḥ	*marútaḥ*	*viśvávedasaḥ*
richly_raining.NOM.PL.M	Marut.NOM.PL.M	owning_all_treasure.NOM.PL.M
právepayanti	*párvatān*	*ádābhyāḥ*
shake.PRS.3PL	mountain.ACC.PL.M	invulnerable.NOM.PL.M

'The **richly-raining Maruts**, **owning all treasure**, shake the mountains, **the invulnerable ones**.' (RV 3.26.4, adapted from Schäufele 1991a: 58)

Given the certain degree of ordering flexibility, it is of course hard to say whether all such cases in Vedic should indeed be discarded as examples of a discontinuous nominal expression or not. Ultimately, it has to be discourse criteria that give the

---

9 Information on the pitch accent system of Vedic derives, among other things, from the historical tradition of marking diacritics in manuscripts, metrical evidence, and descriptions by classical Indian grammarians.

10 The unaccented finite verb *vepayanti* 'they shake (causative)' is enclitically attached to an accent-bearing spatial adverb *prá* 'to, towards', with which the verb forms a semantic unit. By convention, the local particle and enclitic verb form are written as an orthographic unit.

answer, rather than formal indications, in order to avoid circularity. However, discourse criteria are not always applicable, as in the case of ritual formulas, and only certain text portions involve the kind of narrative structure that can guide one's assessment of information structure and textual coherence. Therefore, it seems advisable to err on the conservative side and exclude such cases – i.e. where nominals seem distributed over separate domains judging by word order – on principle.

Besides excluding cases where elements are not found in the same domain, it is important to note that nominals with the same morphological marking (if any) do not have to form one expression even when found inside one domain. In fact, this is possibly an even more important point to make, as these cases are harder to spot, with elements often appearing side by side and showing the same morphology. For instance, secondary predicates have been identified in previous studies as being easily mistaken for attributes in flexible word order languages, for instance McGregor (1997: 102–103), Schultze-Berndt & Simard (2012: 1028–1030); for a study focusing on secondary predicates in an Australian language see Simpson (2006b) on Warlpiri. Secondary predicates are not part of a nominal expression whose referent they add information about, but constitute an additional predicate beside the main predicate, temporally overlapping with the latter (cf. Himmelmann & Schultze-Berndt 2006). Consider the following example from Jaminjung (see Schultze-Berndt & Simard 2012: 1028–1030 for discussion):

(7)   Jaminjung (non-Pama-Nyungan)
       ***jarlag**=biyang jag   ga-rdba-ny   **mangarra** \
       good=SEQ   go.down 3SG-fall-PST plant.food
       'The **fruit** fell down **ripe**.' (cf. Schultze-Berndt & Simard 2012: 1030)

Vedic Sanskrit also frequently employs nominal elements as secondary predicates, in particular participles (see Casaretto & Reinöhl In prep.). Additionally, predicates in main and subordinate (in particular, relative) clauses, as well as in complex converb expressions, are also frequently morphologically nominal in Vedic, all of which may easily be mistaken as forming complex expressions with a nominal in argument role, for instance (see Reinöhl To appear, for details).

The relevance of being strict about what it takes to form a real complex nominal expression is underscored by a count of examples in the Vedic prose corpus. The application of the discourse-functional understanding adopted here necessitates the exclusion of about two thirds of the sentences as not actually involving real complex nominal expressions. All of these contain two or more nominal elements with the same case, gender and number marking.

3 Continuity

Over the next two sections, it will become clear that motivations for deviations from the default order in continuity, on the one hand, and motivations for placing elements discontinuously, on the other hand, are partially similar. Nonetheless, relative order in continuity and the various patterns of discontinuity are not simply, or not obviously, part of the same larger phenomenon. Firstly, the basic principle that what belongs together functionally is placed together is not violated with non-default continuity, while it is violated in discontinuity. Secondly, it is not obvious whether relative ordering in continuity shares certain motivations with the same relative ordering in discontinuity. Thirdly, it has not yet been sufficiently explored how phrase-internal flexibility (as, for instance, in Tagalog, see Himmelmann 1997: 162; Shih & Zuraw 2017) compares to flexibility that occurs across a larger domain, i.e. discontinuity. For these reasons, discussions of continuous and discontinuous cases are kept separate for the purposes of this paper.

3.1 Default order: Functional templates

Generally, and this point has been made by others before, it does not ever, or not normally (but cf. remarks by Tsunoda on Warrongo discussed below), seem to be the case that ordering is in some sense arbitrary in any language, in spite of what a label like "free word order" might suggest. Thus, it has been observed for several of the languages discussed here that nominal expressions have a default order. While order is not as constrained as, e.g., in English, most nominal expressions attest to a particular, recurring pattern (or, more precisely, to a small range of patterns).

Importantly, the default continuous order is described for several languages here as being primarily based on functional templates rather than on word class, a line of research going back to work by McGregor (1990) in particular. By contrast, English NPs, for instance, can be described as primarily based on word classes, both with regard to ordering and to obligatoriness restrictions (e.g. attributive adjectives typically occupy a slot before a noun and cannot on their own function as heads of NPs).[11] In the languages at issue here, relative ordering involves default positions for modifying elements as well as for modified ones ("heads"), instead of for adjectives and nouns. Examples of ordering governed by functional templates include Bardi, where modifiers typically precede their heads (Bowern 2012: 158), and the same is said of Kayardild (Evans 1995: 235–236), which also matches the general pattern in

---

[11] This is not to say that functional considerations have no role to play in English. For instance, they have been invoked in analyses of the relative order of modifiers.

Vedic Sanskrit. In Jaminjung, by contrast, modifiers typically follow the head (Schultze-Berndt & Simard 2012: 1020; see Louagie & Verstraete 2016: 41–42 et passim for references to modifier-modified order in further Australian languages).

The absence of an entrenched noun-adjective distinction plays a key role in understanding functional templates (cf. remarks in Baker 2001: 1438 or Nordlinger 2014: 237–238). Many nominal elements may be used both as property-denoting and as entity-denoting elements and, accordingly, appear either in slots for modifying or for modified elements. This point is illustrated by the following examples from Vedic Sanskrit, both of which involve the element *svargá-* (stem form). In ex. (8), *svargá* (a sandhi form of the locative form *svargé*) is used on its own in an entity-denoting sense, meaning 'in heaven'. In (9), by contrast, context tells us that *svargé* describes a property of *loké*, accordingly translated as 'heavenly'. In accordance with the default position for modifiers in Vedic Sanskrit, *svargé* precedes in this case.

(8)   Vedic Sanskrit
      svargá   *u*   *tvám ápi mādayāse*
      heaven.LOC.SG.M PRT 2SG   at   enjoy.SUBJV.MID.2SG
      'May you enjoy yourself in **heaven**' (RV 10.95.18)

(9)   Vedic Sanskrit
      sahá   ***svargé***   ***loké***   *bhavataḥ*
      together heaven.LOC.SG.M world.LOC.SG.M become.PRS.3DU
      '(the two) come together in the **heavenly world**' (MS I 4,7(2))

If the default functional template is taken into account, any superficial impression that one might have of somehow unconstrained word order flexibility quickly dissolves. Also, note that functional flexibility of open-class nominals does not extend to all members of the class equally. Thus, a particular lexeme's meaning may lend itself to a use as either entity- or property-denoting, but not to both, as observed by McGregor for Gooniyandi:

> The Entity and Qualifier functions are frequently realised by words corresponding to notional nouns and adjectives respectively, and there are some nominals which (in the present corpus) are restricted to one or the other of these two roles: for instance, *mayaroo* 'house' seems to always function as an Entity, and has not been encountered in the Qualifier role; on the other *hand, joodoo* 'straight' seems to be always a Qualifier, never an Entity. (1990: 256)

An element that is functionally restricted given its semantics will, in turn, show less syntagmatic flexibility. In other cases, differences in functional flexibility between elements may seem arbitrary at first glance. For instance, *loká-* (stem form) in the example above is normally only used in an entity-denoting sense,

i.e. meaning 'world', not 'worldly', in contrast to the flexibility of the semantically similar *svargá-* 'heaven, heavenly'. In Vedic, such irregularities can often be accounted for by processes of word derivation (productive ones or no longer productive ones) or by associations of the entity-denoting use with a particular gender (see Reinöhl To appear, for details).

Analysing syntagmatic patterns in more detail, McGregor (1990) develops an analysis for Gooniyandi which involves several slots for different modifying functions besides the slot for an entity-denoting nominal. The template looks as follows:

(10) (DEICTIC)^(QUANTIFIER)^(CLASSIFIER)^ENTITY^(QUALIFIER) (McGregor 1990: 253)

McGregor outlines how, for instance, a particular open-class nominal may be used, e.g., as a "classifier" or as a "qualifier", and is accordingly found in different positions in relation to another nominal in entity-denoting role. Thus, *jiginya* 'little' can precede or follow another nominal as illustrated in ex. (11) and ex. (12) according to function. As a "classifier", it precedes, but as a "qualifier", it follows.

(11)   Gooniyandi (non-Pama-Nyungan)[12]
       jiginya marla
       little   hand
       'a finger' (McGregor 1990: 272)

(12)   Gooniyandi (non-Pama-Nyungan)
       marla jiginya
       hand   little
       'little hand' (McGregor 1990: 272)

McGregor argues that the functional template accounts for the placement not only of "open-class nominals", but of all elements that are used in nominal expressions in Gooniyandi, i.e. also determiners, number words, proper nouns, kin-terms, subsection terms, nominative pronominals, and oblique pronominals (McGregor 1990: 256). For instance, the next two examples illustrate how numerals may be used either in "quantifier" or "qualifier" function. In the first ex. (13), *garndiwiddi* 'two' serves to establish the number of referents in question, a quantifier usage. For

---

12 McGregor (1990: 273) points out that *jiginya marla* 'finger' does not appear to be a case of "lexical compounding", with *marla* being ambiguous between 'hand' and 'finger'. A lexicalisation analysis may be suggested by the translation, but not by the Gooniyandi evidence.

ex. (14), McGregor explains that the two sons are already given referents, i.e. the sons do not have to be determined in terms of their number; instead, "number can be taken as a property or quality of the referent" (1990: 271).

(13) Gooniyandi (non-Pama-Nyungan)
ngaddagi garndiwiddi ngaloowinyi
my   two   son
'my two sons' (McGregor 1990: 272)

(14) Gooniyandi (non-Pama-Nyungan)
ngaddagi ngaloowinyi garndiwiddi
my   son   two
'the two of my sons, both of my sons' (McGregor 1990: 272)

The fact that syntagmatic slots are tied to functions rather than to parts of speech does not mean that there are no links between functions and certain lexical sub-groups. McGregor outlines how not all types of elements appear in all functions, and hence in all slots. However, all types of elements may appear in two or more different functions, as shown in the following Table 1:

Table 1: "Associations between lexical class and function" (McGregor 1990: 256).

Lexical class	NP roles that may be realized
Determiner	Deictic, Qualifier, Entity
Number words	Quantifier, Qualifier, Deictic
Proper nouns	Entity
Kin-terms	Entity
Subsection terms	Entity, Classifier, Qualifier
Open class nominals	Entity, Qualifier, Classifier, Quantifier
Nominative pronominal	Entity
Oblique pronominal	Deictic, Qualifier, Entity

While an analysis in terms of functional templates has a certain tradition in Australian linguistics, analyses that reach a similarly fine-grained level as McGregor's are still wanting for most of the Australian languages in question as well as for Vedic Sanskrit (but see e.g. Evans 1995: 233–251 for an analysis of nominal expressions in Kayardild inspired by McGregor's approach).

3.2 Heaviness

There are two major types of factors that can change default ordering in continuous structures, information structure and heaviness. I first discuss heaviness effects. An example from Vedic concerns structures with multiple modifiers. As pointed out, when juxtaposed, a modifier usually precedes the modified element in Vedic. However, in the case of several modifiers, we also find them preceding and following the head (cf. Delbrück 2009[1888]: 20) as seen above in example (6), repeated here:

(15) Vedic Sanskrit

    bṛhadúkṣaḥ   *marútaḥ*   *viśvávedasaḥ*

    richly_raining.NOM.PL.M Marut.NOM.PL.M owning_all_treasure.NOM.PL.M

    právepayanti   *párvatān*   *ádābhyāḥ*

    shake.PRS.3PL   mountain.ACC.PL.M invulnerable.NOM.PL.M

    'The **richly-raining Maruts, owning all treasure**, shake the mountains, the invulnerable ones.' (RV 3.26.4, adapted from Schäufele 1991a: 58)

A heaviness effect may be also involved in Kayardild: Evans (1995) explains that, while modifiers normally precede the entity-denoting element (1995: 237), a single modifier may also sometimes follow the latter (Evans 1995: 235). A possibly related phenomenon is found in Wardaman. Merlan (1994: 232, 234) notes that, while nominals in adnominal function typically follow the modified nominal, shorter modifiers enjoy greater word order freedom and also appear to the left of modified nominals.[13]

    In what way and to what extent heaviness is at play in cases of reversed default order is still very much in need of research. Theoretically, the question arises of whether a pattern triggered by a heaviness effect should be treated as derivative of the default functional template or as a template of its own. If deviations can be accounted for, or even predicted, given the number of modifiers, and/or possibly phonological or morphological weight, an analysis in terms of a derivation from the basic functional template seems appropriate. Thus, in Vedic Sanskrit, a modifier follows an entity-denoting nominal particularly often when two or more modifiers are involved (Delbrück 2009[1888]) which could be taken as an argument for such an analysis.[14]

---

13 In some languages, there seems to be a constraint on the number of adnominal modifiers (see Louagie & Verstraete 2016: 50), which might preclude heaviness effects on ordering.
14 While the majority of Vedic examples used for this study comes from prose texts, metrical constraints have been suggested to influence word order in Vedic poetic texts (see e.g. Hock 1997). However, Oertel (1940) has shown that permutations of ordering in otherwise identical

Studying heaviness once again requires a clear delimitation of what complex nominal expressions are. Loose appositions, for instance, may also involve a seemingly reversed order in comparison to the basic functional template. For example, while close appositions tend to precede proper names in Vedic Sanskrit (as in English *President Obama*), loose appositions follow (also as in English, e.g. *Obama, the President*; see Reinöhl To appear, for examples and discussion of the patterns in Vedic). Now, participles and particularly compounds are frequently used as loose appositions in Vedic Sanskrit and, by their morphological nature, tend to be comparatively long elements. Thus, their placement following the noun could, at first sight, be interpreted as a heaviness effect. Importantly, however, we are not dealing with integrated nominal expressions in the first place, but with two independent expressions of equal syntactic rank according to standard analyses. While this distinction can be made relatively easily on a theoretical level, identifying individual loose appositions in practice can be a difficult task (but see Reinöhl To appear, on Vedic, and Sadler & Nordlinger 2010 on appositional combinations of nominals in Australian languages).

3.3 Focus marking

I now come to the first of several instances where a deviation from default ordering is associated with focus marking of one kind or another. In the same way as in other literature, the meaning of the label "focus" varies somewhat across the relevant studies and, indeed, we seem to be dealing with several different information-structural parameters. This includes, e.g. emphasis, contrast, new information or the invocation of alternatives (see Krifka 2008; Matić & Wedgwood 2013). The range of uses will be taken up in more detail below in connection with discontinuity, as well as in Section 5.

Discussing non-default continuous order in Vedic Sanskrit, Delbrück (1878: 39) argues that the following example involves a contrastive focus reading (my terminology) of the 'Brahmins', as it is particularly *they* who should eat the food and not someone else.

---

verses are not explained by metre in numerous cases. A closer examination from the perspective of formulaic language promises deeper insights into this phenomenon (Ong 1982). Thus, Lindqvist (2015: 29–35) suggests that Parry's (1971[1928], 1971) notion of an essentially rigid formula as developed for the Homeric poems is too narrow when applied to the Rigveda, as we find a lot of variation of formulas including word order permutations.

(16) Vedic Sanskrit
      tám   ***brāhmaṇā́ś***   ***catvā́raḥ***   *prā́śnīyuḥ*
      DEM.ACC.SG.M Brahmin.NOM.PL.M four.NOM.PL.M eat.OPT.3PL
      'that (food), **four Brahmins** should eat' (TS 5.7.3.4, cf. Delbrück 1878: 39)

If informationally prominent in this way, the modified element may come first in Vedic. It remains subject to investigation how common this pattern is and whether other factors, such as main constituent order, may play a role.

   A focus interpretation of the semantic head is mentioned in Louagie & Verstraete (2016) also for Tiwi, quoting from Lee (1987). Numerals usually precede the nominal head in Tiwi, but the latter may be fronted in exceptional cases for "focus or for stylistic effect" (Louagie & Verstraete 2016: 38) as shown in the following example.

(17) Tiwi (non-Pama-Nyungan)
      pilayiki yirrara
      flag(M)   two(M)
      'two *flags*' (Louagie & Verstraete 2016: 39 quoting Lee 1987)

Note that, while elements may in principle fluctuate between being used as modifiers or modified elements as outlined in Section 3.1, this point does not seem to be at issue in the Vedic and Tiwi examples just quoted. In these cases, the numerals seem to retain their normal function and the 'Brahmins' and 'flags' remain the entity-denoting elements, no matter in what word order.

   Warlpiri is another language for which it has been claimed that the reversal of default ordering for modifying and modified elements signals a focus reading, in this case on the modifier, which is shifted to the left. Compare the following two examples:[15]

(18) Warlpiri (Pama-Nyungan)
      watiya   ngurrju
      tree/wood   good
      'a/the good tree/wood' (Laughren 1984, quoted in Legate 2000: 10)

---

15 Bowern (2012: 335) considers a focus reading when modifiers follow rather than precede their head in Bardi, evidence for which comes from the behaviour of a quantifier *niimana* 'much' (examples p. 273–274). However, she also mentions that there may be a difference in restrictiveness of the modifier, which suggests an account in terms of different slots being associated with different sub-types of functions along the lines of McGregor (1990). A possibly even better fit might be the modified proposal of McGregor's approach as sketched in Harvey (1992), which explicitly takes restrictiveness into account.

(19)   Warlpiri (Pama-Nyungan)
   ngurrju watiya wiri yama-ku-ju.   *Manu ngurrju watiya kawurnpa-ku.*
   good   tree   big shade-DAT-TOP and   good   tree   ash-DAT
   'It is a *good* big tree for shade, and it's *good* wood for ash (to mix with
   tobacco).' (Laughren 1984, quoted in Legate 2000: 10; italicization added
   based on Legate's analysis)

As with heaviness effects, the question arises whether such reversals should be
analysed as derivations from the basic functional template or as functional
templates in their own right. The last option, however, does not seem convinc-
ing. We are not simply dealing with an alternative functional template in
McGregor's sense where different slots are associated with roles such as "qual-
ifying", "classifying" etc., since such an approach could not capture the special
information-structural correlate. Instead, it seems that prominent information
structure changes the normal default ordering. Accordingly, a derivation
account seems plausible in this case, as well.

     Note that a number of descriptions of Australian languages mention order-
ing flexibility, but do not elaborate on information-structural or other correlates.
In a few cases, there are explicit claims that there are no identifiable correlates,
e.g. by Tsunoda on Warrongo (2011: 347–351). It must be noted, however, that
Tsunoda's grammar is based on the two last fluent speakers of Warrongo, the
main speaker not having spoken the language for fifty years (Tsunoda 2011: 45),
so that the difficulty of identifying any correlates may very well be due to the
limits of a corpus created under such conditions (cf. Tsunoda's insightful
remarks on the language and its last fluent speakers in the introductory chapter
of his grammar). In any case, it is a question for future research to what extent
reversal of default order is associated with focus of one kind or another across
flexible word order languages. It will be particularly relevant to compare lan-
guages based on whether the basic functional template is left-headed as in
Warlpiri or right-headed as in Vedic Sanskrit and Tiwi.

     Before we move on to discontinuous expressions, a note is in order on the
recent treatment of non-default continuous ordering in Louagie and Verstraete
(2016). As already indicated in the introduction, it emerges from their discus-
sion that flexible relative order is only categorised as such if there is no
mention in the literature of a "functional" (i.e. information-structural) corre-
late. For instance, in a section on "fixed word order", the authors remark with
regard to Tiwi that

> Some of these languages allow a change in word order for emphasis or focus, as in Tiwi,
> where the head nominal ... can be fronted for focus or for stylistic effect ... . Since such

changes have a clear functional motivation and are not the default, we do not regard this as counterevidence for NP constituency. (Louagie & Verstraete 2016: 38–39)

Louagie and Verstraete (2016) interpret the term "constituency" as involving fixed word order (besides other criteria). If this is how constituency is defined, if only in part, then a change of word order that does not affect the unit-hood or headedness of the expression presents a piece of evidence against it, for whatever information-structural or other motivation. In other words, if only reversal for no apparent reason whatsoever is accepted as instantiating real flexibility, it seems probable that we will never be able to diagnose flexible word order anywhere. Once a detailed, discourse-based study is conducted, it is expected that information-structural or other correlates will be found. Therefore, it seems best to retain the concept of word order flexibility also in such cases where information-structural or other correlates such as heaviness effects can be identified.

4 Discontinuity

Louagie & Verstraete (2016: 51) count 49 languages in their sample of 100 Australian languages for which some kind of "discontinuity" is an option, and this number may even be higher since discontinuity is not explicitly excluded for a number of the other languages.[16] In any case, there is a wealth of grammatical descriptions in which at least some examples are found. Several descriptions also contemplate conditioning factors, some of which are cited below. Importantly, however, not all descriptions of "discontinuity" in the literature fall under the definition adopted here.

"Discontinuity" is understood here as encompassing cases where two (or more) elements appear to form a complex expression while not being placed in juxtaposition, but where the latter would in principle be an alternative possibility, given changed information-structural conditions, for instance. If discontinuity was the only option, we would not be dealing with actual syntactic flexibility. I also exclude cases where it seems plausible to analyse the example as a case of another element "splitting" the complex nominal expression, rather than the members of the expression standing apart. An example from Vedic Sanskrit are cases where a "focus"-marking particle, which always attaches to the right of the host that it has scope over, is placed after an adnominal modifier

---

16 However, it is not always clear whether an attempt has been made to separate apparent from real cases of complex nominal expression in language descriptions that mention discontinuity as a possibility.

and thus "splits" that modifier from its following head. Cases that look like they may not involve "real" complex nominal expressions are also strictly excluded, along the lines discussed in Section 2. Thus, for instance McGregor's (1997) "Type B Discontinuity" is not discussed here, as it does not conform to the requirement of prosodic domain integrity. Also, any example type for which someone proposes some kind of "unmerged interpretation" in the sense of Hale's (1983) understanding will not be considered either.

The following discussion is organised according to three distinct syntagmatic types of discontinuity.[17] These syntagmatic types show specific functional correlates, but no one-to-one correspondences (possibly with the exception of the minor Conjunction-related Discontinuity). All three types have been identified as discontinuity for at least one of the languages under investigation in previous studies, while not labelled in the same way. While examples for all types are attested for Vedic Sanskrit, the literature on Australian languages displays a strong focus on what is here labelled "Left-Edge Discontinuity", which involves the placement of one element in a position at or near the left edge of the domain in question. An important difference between the more widely attested Left-Edge Discontinuity and the other two types is that whereas the former involves all kinds of elements, the latter are linked in different ways to the behaviour of specific function words. Wackernagel Pronoun Discontinuity involves the insertion of a pronoun in 2nd ("Wackernagel") position, which forms a complex expression with elements elsewhere in the sentence. Conjunction-related Discontinuity involves the optional placement of conjunctions in between elements that form a complex expression. Whether further types should be added or whether one or more of the three types should be sub-divided is a question for further research; at this point, the different formal characteristics paired with their specific functional correlates motivate the tripartite categorisation adopted here.

As with continuous patterns, the comparison reveals that we find formal as well as functional similarities across the languages examined. In part, the findings corroborate ideas that deviations from the juxtaposed default often come with somehow informationally prominent interpretations, such as different types of focus marking. At the same time, however, there is also considerable variety in interpretations including the attestation of ones that are not informationally prominent in any obvious way.

---

17 This classification is loosely based on Reinöhl (2016: 122–131). In that study, the categories are ordered into four types from "weakest" (Type 1) to "strongest" (Type 4) in terms of the "degree" of disruption. Left-Edge Discontinuity roughly corresponds to what is Type 4 in that study.

4.1 Left-Edge Discontinuity

In Left-Edge Discontinuity, a nominal element placed at or near the left edge[18] of a domain (typically corresponding roughly to a clause-size unit) forms a complex expression with another element (or elements) following later in the same domain. One or more unrelated element(s), typically pronouns, discourse markers and/or verbal elements, is placed in 2nd position (sometimes characterised as a "Wackernagel" position or simply "P2", see Wackernagel 1892; Halpern & Zwicky 1996; Anderson 1993, Anderson 2005). Optionally, further unrelated material, typically not more than one word, may stand in third position.[19] The schema can be formalised as follows (where E_i stands for an element that forms part of a complex nominal expression).

(20) E_i P2 (...) E_i (...)

Left-Edge Discontinuity is in a way the most radical type of discontinuity, but seemingly also the most common one. On the one hand, it cannot be accounted for by syntactic (if optional) word order constraints on any of the elements belonging to the nominal expression, in contrast to Wackernagel Pronoun Discontinuity. On the other hand, the integrity of the nominal expression is often seriously disrupted by several words standing in-between the elements in question, which sets it off in particular from Conjunction-related Discontinuity.

    With regard to the question of the domain, some authors including McGregor (1997), Bowern (2012) and Schultze-Berndt & Simard (2012) explicitly state that discontinuous expressions that conform to what is called here Left-Edge Discontinuity are uttered in the same type of intonation units found

---

18 In some cases, a clause-introducing conjunction, an adverb, or other kind of adjunct precedes the left-edge element in question (e.g., see ex. (43) below). Note that most examples in the literature do seem to involve an element that is the left-most element in the domain in question, but there are exceptions.

19 Cases where the nominal expression immediately brackets the 2nd position without additional intervening material have been analysed in some studies as 2nd position elements being inserted into or "splitting" the nominal expression (e.g. Pullum 1981; also Taylor 1990 for Ancient Greek), instead of the nominal expression being placed discontinuously. However, it is important to note that 2nd position elements at least in some languages obey word order rules of their own, instead of being inserted "into" a nominal expression (unlike, e.g. the Vedic focus particle mentioned in the introduction to Section 4). Furthermore, the information-structural correlates appear to be similar, whether or not it is only 2nd position elements or also additional material in third position that separates the elements in question. This motivates the unified analysis adopted here.

with continuous expressions. At the same time, however, prosodic informa-
tion has not yet become available for several languages that seem to display
this type, a more reliable classification of which will have to await further
research.

    Examples for Left-Edge Discontinuity show variation in several respects.
Firstly, the 2nd position shows variation with regard to which elements are
placed into it. For instance, the Vedic "Wackernagel" position can be filled by
accented and unaccented discourse clitics and/or pronominal forms, sometimes
up to three or so at the same time (see Reinöhl In prep, for examples). By
contrast, the 2nd position is filled in at least some of the Australian languages
by pronominals and/or a TAM base or a verbal element.[20] The following exam-
ples from Vedic Sanskrit (ex. (21)), Kayardild (ex. (22)) and Gooniyandi (ex. (23))
illustrate some of the variation in 2nd position.

(21) Vedic Sanskrit
      sárva   enam   **r̥távo**   jinvanti
      all.NOM.PL.M DEM.ACC.SG.M season.NOM.PL.M quicken.PRS.3PL
      '**all seasons** invigorate him' (MS I 6,9(2))

(22) Kayardild (non-Pama-Nyungan)
      dan-da   kunya-a   walbu-wa nga-ku-l-da kurrka-n!
      this-NOM small-NOM   raft-NOM 1-INC-PL-NOM take-NEGIMP
      jungarra kurrka-tha **walbu**,   dan-da   mutha-a   dangka-walad
      big(NOM) take-IMP   raft(NOM) here-NOM many-NOM person-LOT(NOM)
      'Don't let's take this small raft! (Let's) take the **big raft**, there are lots of
      people here.' (Evans 1995: 249–250; glossing as in original)

(23) Gooniyandi (non-Pama-Nyungan)
      ye   **yaanya** lamajjidi   **yoowooloo** niyi-nhingi/
      [false start] other   we:picked:him:up man   that-ABL
      'Oth... We got **another Aboriginal man** from there.' (adapted from
      McGregor 1997: 95)

The following example from Thargari involves two syntactic words in between
the relevant elements. Arguably, both the pronoun and the verbal complex
could be analysed as being "P2" material, rather than an element in 2nd position

---

20 It does not affect categorisation here whether or not the 2nd position is referred to in the
description in question as a "Wackernagel"-type position.

and another one in 3rd position, given that both pronominal and verbal (or auxiliary) elements are typical 2nd position material.

(24) Thargari (Mantharta) (Pama-Nyungan)
yiṇa ŋaḍa muḍuru-ṇi-nʸa   **waya**
that I   straight-VBLZR-PST   wire
'I straightened **this wire**.' (Klokeid 1969: 37, quoted in Louagie & Verstraete 2016: 52)

A second area in which we find variation is that Vedic Sanskrit strictly only allows a single word in initial position (see ex. (21)) with very few exceptions, while a number of Australian languages have been described as optionally allowing a word or an entire constituent in initial position preceding the 2nd position, most famously Warlpiri (see Louagie & Verstraete 2016: 71–72, Table B-3).[21] The following examples illustrate this optionality with examples from Walmajarri:

(25) Walmajarri (Pama-Nyungan)
wajpal-warnti-ø   pa-lu   **paja-jinyangu**
white_man-PL-NOM   MODAL_ROOT3-3PLS   many-very
pirri-pirri-yan-i
toward-REDUP-go-PST
'**A lot of whites** were coming around' (Hudson 1978: 89, cf. Louagie & Verstraete 2016: 71)

(26) Walmajarri (Pama-Nyungan)
Parri-jarra-ø kurriny-ø ø-pila   warta-ngurni yan-an-i
boy-DU-NOM   two-NOM   MODAL_ROOT1-3DUS   behind-ABL   go-REP-PAST
wangki-wangki-karra
word-REDUP-MAN
'**Two of the boys** came up from behind us talking away.' (Hudson 1978: 90)

---

21 Variation in Warlpiri between either only a single word (belonging to a complex expression) or the entire complex expression being placed initially before the Wackernagel position is discussed by Hale (1983) and Simpson (1991, chapter 5), among others. However, Warlpiri is categorised in Louagie and Verstraete (2016: 71–72) as a language where the clitic complex always follows the first constituent. Wangkajunga is listed in Louagie and Verstraete (2016: 71–72) as allowing for either a word or constituent in initial position preceding 2nd position elements, but Jones in her grammar is skeptical with regard to whether we are dealing with "real" complex nominal expressions in these cases (cf. Jones 2011: 245–246).

Thirdly, both in the Vedic Sanskrit material and in the literature on Australian languages, examples can sometimes be found where the nominals are separated not only by elements typical for the 2nd position, but by additional unrelated material. This is illustrated with the Vedic ex. (27) and the Gooniyandi ex. (28).

(27)   Vedic Sanskrit

     ubháyīr   *vā́*   *agnihotríṇi*   **devátā**
     both_kinds.NOM.PL.F   PRT   Agnihotr_sacrificer.LOC.SG.M   deity.NOM.PL.F
     ā́śaṁsante
     hope.MID.3PL
     '**Both kinds of gods** hope for the Agnihotr-sacrificer...' (MS I 5,11(2))

(28)   Gooniyandi (non-Pama-Nyungan)

     Wanyirri   *ngarragi-ngga yoowooloo gardbini* **jamarra**
     river:kangaroo   my-ERG   man   he:hit:it male:kangaroo
     'A male river kangaroo was killed by my son.' (McGregor 1997: 95)

A note is in order here on the different types of modifiers involved. Several of the examples quoted involve demonstratives, quantifiers and other operators, and such elements are frequent also in the Vedic Sanskrit corpus (see also remarks in Louagie & Verstraete 2016: 51–52). Still, Left-Edge Discontinuity is attested also with content nominals, as seen in some of the examples above. This point suggests that the flexibility shown here extends significantly beyond more widely attested types of flexibility, such as relating to the behaviour of quantifiers ("quantifier float").

    Turning to functional correlates, I begin with the seemingly most robust form-function correlate that can be identified cross-linguistically, namely where a modifier placed at the left clausal edge bears a certain type of focus interpretation, typically one invoking contrast. This correlate has been described for a number of Australian languages (Schultze-Berndt & Simard 2012: 1038–1039 with references) and is also found in the Vedic data.[22] A further characteristic that has been associated with modifier focus is that the modified element is given, as has been described for Jaminjung (Schultze-Berndt & Simard 2012), a pattern also found in Vedic Sanskrit. Consider once more the following example from Jaminjung where the 'dog' was already mentioned in the preceding conversation. Then the speaker says to the dog owner:

---

[22] A correlation with modifier focus is also attested for discontinuity in languages with nominal expressions whose syntax is otherwise more rigid (see e.g. Siewierska 1984; Fanselow & Caroline 2006).

(29) Jaminjung (=ex. (2)) (non-Pama-Nyungan)
      ^***mulanggirr*** *ngantha-ma-ya*   ***wirib*** \
      fierce   2SG>3SG-have-PRS dog
      'You have a ***dangerous*** **dog**!' (Schultze-Berndt & Simard 2012: 1035,
      emphasis in original)

In the next example from Vedic, the discontinuous nominal expression presents a new topic (the Wackernagel particle *vái* being associated with new topics, whether paragraph-initially or paragraph-internally). Context suggests an identificational focus interpretation of the demonstrative, picking out a previously introduced referent, i.e. 'in *that* place (not another place), the sacrificer ... '.

(30) Vedic Sanskrit
      agnír   *vái* *bhū́yāṁsaṃ*   *prádahaty.*
      Agni.NOM.SG.M   PRT more.ACC.SG.M   burn.3SG
      etáṁ   *vái* ***lokáṁ***   *yájamāno*   *'nvátimucyate.*
      DEM.ACC.SG.M   PRT place.ACC.SG.M   sacrificer.NOM.SG.M   liberate.PASS.3SG
      'Agni [i.e. the fire god] burns (s.o.) who is too much [i.e. who comes too close to the fire]. **In *that* place** [i.e. the boundary of the fire] the sacrificer is liberated.' (MS I 4,6(2))

The focus reading of the modifier can also be suggested by the following context rather than by the preceding context. An example is the next paragraph-initial one from Vedic Sanskrit where the 'two types of gods' are elaborated on subsequently:

(31) Vedic Sanskrit
      dvayā́   *vái* ***devā́***   *yájamānasya*   *gr̥hám*
      twofold.NOM.PL.M PRT god.NOM.PL.M sacrificer.GEN.SG.M house.ACC.SG.M
      ā́gacchanti   *somapā́*   *anyé*
      come.PRS.3PL   Soma_drinker.NOM.PL.M   other.NOM.PL.M
      'somapā   *anyé*   *hutā́do*
      non_Soma_drinker.NOM.PL.M   other.NOM.PL.M sacrificed_eater.NOM.PL.M
      'nyé   *'hutādo*   *'nyé*
      other.NOM.PL.M   non_sacrificed_eater.NOM.PL.M   other.NOM.PL.M
      '***Two types* of gods** come to the house of the sacrificer. Some are Soma-drinkers [i.e. drinkers of an intoxicating substance], some aren't Soma-drinkers, some eat what has been sacrificed, some do not eat what has been sacrificed.' (MS I 4,6(3))

Focus interpretations are also found on modified elements in utterance-initial position, as illustrated by the following Bardi example (Bowern 2012; cf. Louagie & Verstraete 2016: 53).

(32)　Bardi (non-Pama-Nyungan)
　　　aarli *i-na-m-boo-n*　　　　　　　***gooyarra***
　　　fish　　3-TR-PST-spear-REMPST　two
　　　'He speared two *fish*' (Bowern 2012: 329; translation added)

That the initial position may be associated with a focus reading, no matter whether it is filled by a modifying or modified element, fits well with analyses of the initial position as hosting highly newsworthy elements in flexible word order languages (Mithun 1987; see Simpson & Mushin 2008 on four Australian languages). Note that this correlation between form and function seems largely independent of whether the elements are separated only by typical 2nd position elements or also by further material, and of how the remainder of the domain is structured. This is a primary reason for positing an overarching Left-Edge Discontinuity irrespective of the formal variation discussed.

At the same time, however, the nominal that is not in initial position has also been associated with a another kind of focus interpretation. McGregor's (1997) "Type A" discontinuity in Gooniyandi involves cases where "the first part of the discontinuous NP is a theme of the clause; and ... the second part of the discontinuous NP is an unmarked focus of the information unit." (1997: 94). The following example illustrates this pattern:

(33)　Gooniyandi (=ex. (28)) (non-Pama-Nyungan)
　　　Wanyirri　　　　*ngarragi-ngga yoowooloo gardbini* ***jamarra***
　　　River:kangaroo　my-ERG　　　man　　　he:hit:it　male:kangaroo
　　　'A male river kangaroo was killed by my son.' (McGregor 1997: 95)

Variation between the two patterns of a modifier either preceding or following, while bearing a focus interpretation, is illustrated with examples from Wardaman in Merlan (1994: 241–2).[23]

In yet other cases, we seem to be dealing with a focus interpretation of the entire construction rather than of one element only. Cases where modifier and modified are reversed in Jaminjung are analysed by Schultze-Berndt and Simard

---

23 Some of the examples quoted by Merlan may possibly involve secondary predicates, however, judging by the contents.

(2012) as involving "thetic" or "sentence-focus" interpretations where no element is given, which are found in spontaneous exclamations (see the previous descriptions in Siewierska 1984 on possibly similar cases in Polish; McGregor 1997: 95–96 on Gooniyandi):

(34) Jaminjung (non-Pama-Nyungan)
      jarndu *ga-ram*   **luba** *mangurn=mij!*
      boat   3SG-come.PRS big   white.person=COM
      'There comes **a big boat** with white people!' (Schultze-Berndt & Simard 2012: 1041)

In terms of the theoretical status of Left-Edge Discontinuity, one might consider it simply another functional template, a possibility raised above also for cases of non-default continuity. It is of course desirable to spell out how the discontinuous template(s) relate(s) to the (default) continuous one in terms of the association of functional roles such as "qualifying", "classifying" etc. with syntagmatic slots. However, as with non-default continuity, it is not clear how to accommodate the "prominent" information-structural interpretations that are typically associated with discontinuity, if one operates only on the level of functional templates. It seems to be a more adequate solution at this point to include a level of information structure besides the level of functional templates in accounting for the different patterns (see also Simpson 2006a). In doing so, one would need to distinguish between several different information-structural interpretations for the same formal pattern if different types of focus interpretations are found mapping onto one and the same slot and functional role.[24] For cases of discontinuity, it will furthermore be important to take into account that not all cases show clear evidence of a "focus" or other kind of prominent information status playing a role, a point which I turn to now.

   In some cases in Vedic Sanskrit, nominal expressions are repeated in full, and are thus neither contrastive nor new. This pattern is frequent in the Vedic Sanskrit corpus, which is not entirely surprising given the assumed origin of the texts as oral literature (cf. the role of redundant or copious style in primarily oral

---

24 An example of the default continuous order might of course also involve an informationally "prominent" interpretation of one or all of the elements involved (which may be marked, for instance, by prosodic means). However, the default pattern does not appear to be as closely linked to "prominent" information-structural interpretations as deviations from the default ordering seem to be.

societies, Ong 1982). In the example below, 'all seasons' is taken up twice, the second time placed discontinuously:

(35) Vedic Sanskrit (quoted in abbreviated form in ex. (21))
  tád asmai   **sárva**   **r̥távaḥ**   *śivā́*
  so   DEM.DAT.SG.M all.NOM.PL.M season.NOM.PL.M auspicious.NOM.PL.M
  bhavanti   **sárva**   *enam*   **r̥távo**
  become.PRS.3PL all.NOM.PL.M DEM.ACC.SG.M season.NOM.PL.M
  jinvanti
  quicken.PRS.3PL
  '**All seasons** are thus becoming auspicious for him; **all seasons** invigorate him' (MS I 6,9(2))

Ex. (35) is relevant in light of the discussion of the general range of functional correlates of discontinuity. Schultze-Berndt & Simard (2012: 1033) argue that the low frequency of discontinuity is in accordance with the fact that it signals special, non-default information structure. However, ex. (35) from Vedic Sanskrit, and possibly also examples such as ex. (28) from Gooniyandi, do not so obviously involve "prominent" information status. Thus, while Left-Edge Discontinuity indeed seems to frequently involve focus interpretations, this does not seem to be the case in all instances. Possibly, a detailed study of genre effects could shed more light on the different information-structural interpretations associated with this discontinuity type, language-internally as well as cross-linguistically.

4.2 Minor types

The following two sections present two minor types of discontinuity which are more restricted in terms of the languages in which they appear and in terms of the specific types of elements involved.

4.2.1 Wackernagel Pronoun Discontinuity

As already pointed out in the context of Left-Edge Discontinuity, in some languages, pronominal forms are regularly placed in 2nd position (see e.g. Mushin 2006 on some Australian languages). If these pronominal forms qualify as syntactically independent elements, and if they convincingly seem to be used adnominally, they may be analysed as forming a discontinuous complex expression with nominals elsewhere in the clause. I outline this type first for Vedic

Sanskrit and then discuss similar patterns in certain Australian languages, which may or may not fall within the realm of "discontinuity" depending on empirical specifics as well as theoretical assumptions.

One of several phenomena discussed by Schäufele (1991a) as giving rise to "discontinuous constituents" (1991a: 84) in Vedic Sanskrit is what he calls "pronominal fronting" (1991a: 74–77; see also Delbrück 1878: 59–60; Hock 1982).[25] In ex. (36), a demonstrative is placed in the Wackernagel position and the nominal with which the demonstrative forms a complex nominal expression follows later in the clause. On the assumption that the pronominal and the nominal form a complex nominal expression, we are here dealing with a type of discontinuity.

(36) Vedic Sanskrit

 káti   ***ayám***   *adyá*   ***udgātā́***   *asmin*

 how_many   DEM.NOM.SG.M   today   Udgātṛ.NOM.SG.M   DEM.LOC.SG.M

 yajñé   *stotríyāḥ*   *stoṣyati*

 sacrifice.LOC.SG.M   stotriya.ACC.PL.M   sing.FUT.3SG

 'How many stotriyas [verses] will **this Udgātṛ** [a type of priest] sing in this sacrifice today?' (ŚaB 14.4.6.12, cf. Schäufele 1991a: 76)

Arguably, cases where a pronominal is placed in 2nd position and a co-referential nominal is directly adjacent to it (whether in initial or in third position) can also be counted as Wackernagel Pronoun Discontinuity. In the following example, which illustrates this sub-type, we see a case where the Wackernagel position hosts three elements *vā́ asya ayám*.[26] The last element in the string, *ayám*, forms a nominal expression with nominals following directly after the Wackernagel position, *manuṣyò 'gníḥ*.

(37) Vedic Sanskrit

 yó   *vā́*   *asyāyáṃ*   *manuṣyò*

 yáḥ   vā́   asya   **ayám**   **manuṣyàḥ**

 REL.NOM.SG.M   PRT   DEM.GEN.SG.M   DEM.NOM.SG.M   of_men.NOM.SG.M

 'gníḥ

 agníḥ

 fire.NOM.SG.M

 '**This fire** of his, **belonging to man**, which ...' (MS I 6,5(1))

---

25 Wackernagel elements do not occur in every Vedic clause, but it is a common feature especially in certain prose texts.

26 The typical template of Wackernagel elements in Vedic prose texts is (P) (Ṕ) (D) (D́) (P = (discourse) particle, D = pronoun; the acute indicates word accent), where one or several of the positions can be filled, and some positions may be filled by two elements (Hock 1982).

While it presents a significant terminological stretch to include cases such as ex. (37) under the label of Wackernagel Pronoun Discontinuity, there are good reasons for taking this approach. In terms of what gives rise to this type, i.e. the constraint on the pronoun to appear in 2nd position, we are dealing with the exact same phenomenon as in cases like ex. (36). Thus, "surface" continuity as in ex. (37) should not be taken as being motivated by the pronoun and the following nominal forming a discourse-functional unit. The surface continuity only arises because the pronoun happens to be either the only one in Wackernagel position or, alternatively, the first or last one in a string of Wackernagel elements, with the other nominal(s) directly adjacent.

As for the information-structural profile of Wackernagel Pronoun Discontinuity, it depends on the properties of the pronominal placed in P2. Thus, the different types of demonstratives in Vedic Sanskrit each come with their own semantic (i.e. proximative or obviative) properties and functions of reference tracking (i.e. anaphoric or cataphoric) (see remarks in Hock 1982; Amano 2009: 55–58, 67–70).

Shifting our focus to Australian languages, it is relevant to take note of the key role that pronominals in 2nd position have played in scholarly debates over how to theoretically accommodate the "non-configurational" grammars of languages such as, e.g. Warlpiri, Nunggubuyu, or Kalkatungu. As briefly mentioned in fn. 5, Jelinek (1984) famously proposed that, in Warlpiri, pronominal clitics in 2nd position fill the argument slots of verbs, while other nominals are mere adjuncts to these "pronominal arguments" (see Austin 2001: 37 for references to similar proposals predating Jelinek's paper). This analysis was claimed to explain the flexible phrase as well as word order and zero anaphora, the defining characteristics of "non-configurationality" as proposed in the seminal paper Hale (1983). While Jelinek's approach has long been shown not to account for several facts of Warlpiri (Simpson 1991; Austin & Bresnan 1996) or of flexible word order languages more generally (Austin 2001; Austin & Bresnan 1996), the relevant question remains whether "pronominal" elements in Wackernagel position may have a status that can be characterised as adnominal.

There are, as far as I can see, at least four criteria that bear on this question. These are (i) whether the "pronominals" in the language in question are optional in 2nd position (rather than obligatory) either in all clause types or in specific clause types,[27] (ii) whether they can also appear elsewhere in the clause

---

27 In the languages at issue here, the pronominal form in P2 may be the only realisation of an argument in the clause. In some languages including Vedic Sanskrit, the argument may also remain completely unexpressed, not even by an element in 2nd position (nor, in Vedic, by verb agreement in the case of non-subject arguments).

including adjacent to the other element(s) they form a complex expression with, (iii) whether they instantiate the same case system as free nominals or pronominals, and (iv) whether they stand independently in 2nd position, or are hosted by some kind of a "base" which expresses, e.g. TAM values. Vedic Sanskrit pronominals in 2nd position meet these criteria: they are syntactically optional, at least some of them may also appear elsewhere in the clause including adjacent to a nominal they form a complex expression with, they inflect for the same case values as free (pro-)nominals, and they do not require any kind of "base". Thus, they qualify as syntactically independent elements in all four respects, and may function adnominally, which as a result allows for an analysis of examples such as ex. (36) (and possibly also ex. (37)) above in terms of discontinuity. By contrast, if the 2nd position pronouns in a language fail one or more of these criteria of syntactic independence, it is up for debate whether we are dealing with discontinuity, as this would require at least two independent and syntagmatically at least somewhat flexible words.

    The following example from Ngiyambaa fulfils the above criteria like Vedic Sanskrit with the partial exception of criterion (ii).

(38)   Ngiyambaa (Pama-Nyungan)
     *miri-gu=**na**   **bura:y**   gadhiyi*
     dog-ERG=3ABS   child.ABS   bite.PST
     'The dog bit **the child**.' (adapted from Donaldson 1980: 128)

The pronominal is syntactically optional in 2nd position,[28] it shows the same case marking as free forms, and it is not attached to a "base". However, the absolutive form in 2nd position is only partially identical to demonstratives that appear elsewhere in the clause (*na* vs. *ŋina/ŋana*, see Donaldson 1980: 135–136). Nevertheless, Donaldson analyses the "pronominal" in second position as adnominal, matching the analysis adopted here for Vedic Sanskrit:

> When a third person pronoun is adnominal, i.e. part of a NP containing a nominal, it is enclitic on the first word or constituent of the sentence in exactly the same way as when it serves as an anaphoric pronoun. (1980: 128)

As the data come from spoken fieldwork corpora, clauses tend to be shorter than in the Vedic Sanskrit sample, so that we often find a pattern of surface adjacency as in ex. (37) above. However, Donaldson writes that "[t]he nominal(s) in

---

28 See the contrasting examples in Donaldson (1980: 128, 5–6 and 5–7), which indicate a difference in definiteness.

apposition to it [i.e. to the pronominal in 2nd position] may be in any subsequent position in the sentence".

However, in at least some other languages with pronominal Wackernagel elements, it seems that these enjoy less syntactic independence than in either Vedic Sanskrit or Ngiyambaa. Mushin (2006) outlines how pronominals in Wackernagel position are often obligatory in Australian languages that do have such a position, a feature that has been accounted for as being an areal feature, and how some of these languages combine the pronominals with, e.g. TAM markers, as was mentioned above (2006: 299–300). Warlpiri differs from Vedic Sanskrit with regard to all four criteria. Pronominal-like forms in 2nd position are generally obligatory, at least in specific clause types (but see Austin & Bresnan 1996: 242–243), they do not appear elsewhere in the clause, nor do they display the same case distinctions as free forms, a point famously discussed in Hale (1983), and they are typically hosted by an auxiliary base (Nash 1986[1979]; Simpson 1991). Compare the encoding of both arguments in the next example:

(39) Warlpiri (Pama-Nyungan)
   maliki-jarra-rlu *ka-**pala**-jana*   **puluku-patu** *wajilipi-nyi*
   dog-DU-ERG   PRS-33SUBJ-333OBJ bullock-PL   chase-NONPAST
   'The **two dogs** are chasing **the several bullocks**.' (Hale 1983: 18)

The reduced syntactic freedom of Wackernagel "pronominals" in Warlpiri and in certain other Australian languages is reason for excluding example types such as ex. (39) from discussions of discontinuity. (And, terminologically, one might reject the label "pronominals" for the Warlpiri cross-referencing elements on the same grounds.) Still, even though these cases fall outside the realm of discontinuity in the sense of the present paper, it is worth noting the structural resemblances with regard to the behaviour of pronominals regularly appearing in 2nd position in other languages. With regard to the criteria formulated above, Vedic Sanskrit pronominals are at one end of the spectrum with regard to their syntactic independence, while the elements in Warlpiri are found at the other end, showing little syntactic independence. The 2nd position pronominals in Ngiyambaa, for instance, while relatively independent, take an intermediate position.

4.2.2 Conjunction-related Discontinuity

A further minor syntagmatic pattern, which may be classified as discontinuity and which is found in Vedic Sanskrit, involves the placement of a clitic conjunction in-between elements that form a complex nominal expression, with the conjunction

having scope over the entire expression (cf. e.g. Gillon & Shaer 2005, ex. 23b). The next two examples illustrate how the conjunction *ca* 'and' may either follow a complex expression or its first element. While the conjunction is of course directly connected to the nominal expression by virtue of having scope over it, it is not itself part of the expression, so that this type may be characterised as discontinuity. Despite its minor status, it is included here for completeness's sake. (In Australian languages, coordination can be syntactically unmarked; see Sadler & Nordlinger 2006, which is also an option in Vedic Sanskrit.)

(40) Vedic Sanskrit
  héḷāṃsi   *daívyā ...*   ***ádevāni*   *hvárāṃsi* *ca**
  wrath.ACC.PL.N heavenly.ACC.PL.N un_godly.ACC.PL.N trick.ACC.PL.N CONJ
  'the heavenly wraths and **the ungodly tricks**' (RV 6.48.10, cf. Klein 1985: 53)

(41) Vedic Sanskrit
  kṣapaḥ ...   ***śárádaś*   *ca** ***pūrvī́ḥ***
  night.ACC.PL.F season.ACC.PL.F CONJ many.ACC.PL.F
  'nights ... and **many seasons**' (RV 4.16.19, cf. Klein 1985: 53)

It is an open question whether assuming further types or sub-types of discontinuity would allow deeper insights into ordering patterns and their correlates. In fact, there are some examples which may be classified as one of the three types of discontinuity proposed, or may alternatively be categorised as yet another type (and see also fn. 16). The following example illustrates this point.

  An example in the Vedic Sanskrit corpus (ex. (42)) could be classified as either yet another type of discontinuity, as an unusual version of Left-Edge Discontinuity, or as a similarly unusual combination of Left-Edge Discontinuity and Conjunction-related Discontinuity. This example also involves *ca* 'and' like ex. (41) above, which here however coordinates clauses. In addition, the "focus" particle *evá* follows *ca*, and the adverb *átra* 'there' concludes the string of elements interrupting the complex nominal expression in question.

(42) Vedic Sanskrit
  yā́ṃ   *caivā́tra* *yajñásya*   *prā́yaścittìṁ*
  yā́m ca evá átra **yajñásya**   **prā́yaścittim**
  REL.ACC.SG.F and FOC there sacrifice.GEN.SG.M compensation.ACC.SG.F
  vidmá
  vidmá
  know.PERF.1PL
  'and **whichever compensation for the sacrifice** we may know' (MS I 4,8(7))

An argument could be made for considering *ca evá átra* a string of "Wackernagel elements" as at least two out of the three elements are somehow "grammatical" or "discourse-functional", as is characteristic for 2nd position elements. This would suggest Left-Edge Discontinuity. However, none of the three elements regularly occurs in 2nd position irrespective of scope as do normal Wackernagel elements. *Ca* and *evá* attach to a host irrespective of absolute position in the clause, while *átra* as an adverb occurs in various environments. Furthermore, only *átra* has any morphological connection to normal Wackernagel elements, as it is formed from the pronominal stem *á-*. A variant of this approach would be to say that Conjunction-related Discontinuity, involving *ca*, combines here with Left-Edge Discontinuity, involving the pronominally derived element *átra*, as well as *evá*, cliticising to the host it has scope over. A third alternative would be to decide that this example instantiates yet another type of discontinuity altogether. Examples such as this notwithstanding, this paper takes a conservative stance on categorisation and only posits such categories as are comparatively well substantiated. The vast majority of examples are accounted for by the three types established here.

5 A spotlight on focus interpretations

Different types of focus interpretations associated with non-default word order of nominal expressions are a thread running through the literature and through this paper. This section summarises the different uses encountered so far and goes into some more detail with regard to the types of focus interpretations that can be identified.

Summarising, we have seen analyses of focus interpretations of modifiers and of modified elements. The former seems the most frequent type across languages, and also the most frequent one in terms of token numbers within languages. Both focus interpretations of modifying and of modified elements occur in Vedic Sanskrit and in various Australian languages. These focus interpretations have been linked to a reversal of the default order in continuity, where the element drawn to the left bears the focus interpretation, as well as to Left-Edge Discontinuity, where focus interpretations are frequently associated with the initial element. In cases of Left-Edge Discontinuity, we have also seen focus readings associated with the element on the right, or with the utterance as a whole in "sentence-focus" interpretations.

A closer look at examples reveals that some of the focus interpretations found in the literature and in this paper involve a contrastive reading that invokes

alternatives. An interpretation of contrast can be found with modifiers (e.g. ex. (29) 'a *dangerous* dog', ex. (30) '(in) *that* place') or with modified elements (e.g. ex. (16) 'four *Brahmins*', ex. (32) 'two *fish*'). As examples ex. (16) and ex. (32) illustrate, this contrastive interpretation is found both with reversed continuity and with discontinuity. In other cases, contrast does not seem to be involved, but rather newness, as with sentence-focus readings illustrated in ex. (34), 'There comes a big boat with white people!'. Newness may also be involved in cases such as the Gooniyandi ex. (33), 'A male river kangaroo … ' where the focus interpretation is said to be on the right-most element in modifying function. The notion of highlighting, in turn, seems orthogonal to contrast and newness, as it can be connected to contrastive interpretations as well as to sentence-focus interpretations, but perhaps not so clearly to the type illustrated by the Gooniyandi ex. (33).

The fact that different examples suggest that we are dealing with different types of focus interpretations fits squarely with the approach advocated by Matić & Wedgwood (2013) that "focus" is not a linguistic category in the sense of a stable form-function pairing. Instead, it is a label for a family of interpretations that involve notions including the ones just mentioned, which may be variously expressed by a range of structural patterns, not only cross-linguistically, but also within one and the same language. The formal means may include different types of word order patterns, but of course also other ones such as, e.g., prosodic cues or particles. At the same time, a specific structural pattern does not seem to be linked to only one specific information-structural interpretation.

That we are indeed dealing with a family of interpretations rather than with a monolithic concept becomes even clearer in a more detailed analysis. While several of the focus interpretations of the Vedic Sanskrit examples appear to involve contrast, they differ with regard to finer variables, such as whether a closed or open set of alternatives is involved (see Krifka 2008), and whether the example involves a semantic or an "identificational" focus interpretation (i.e. where a previously introduced discourse referent is picked out). Thus, ex. (43), 'distribute *all* oblations at the same time', can be said to involve an open (or less restricted) set of alternatives in comparison to ex. (44), where 'unlimited (enlivenment)' is contrasted only with the alternative 'limited(ness)'.

(43) Vedic Sanskrit
    tásya   **sárvāṇi**   *sākáṁ*   **havī́ṁṣi**
    DEM.GEN.SG.M   all.ACC.PL.N at_the_same_time oblation.ACC.PL.N
    nírvapet
    distribute.OPT.3SG
    'for him he should distribute ***all*** **oblations** at the same time' (MS I 6,8(1))

(44) Vedic Sanskrit
    kṣeṣṇú   *vái*   *párimitam*
    perishable.NOM.SG.N   PRT   limitedness.NOM.SG.N
    áparimitam   *evā́smai*   *jī́vanam*
    áparimitam   evá asmai   **jī́vanam**
    un_limitedness.ACC.SG.M   FOC   DEM.DAT.SG.M   enlivenment.ACC.SG.M
    ávarunddhe
    ávarunddhe
    attain.MID.3SG
    'Limitedness is perishable; he attains ***unlimited* enlivenment** for it [i.e. the fire]' (MS I 6,10(2))

While in ex. (43) and ex. (44) the modifier makes a semantic contribution, ex. (45) involves an "identificational" use where a referent (i.e. the 'abode') previously introduced is picked out by way of a reflexive pronoun in focus.

(45) Vedic Sanskrit
    svá   *eváinā*   *yónau*   *své*
    své   evá enāḥ   **yónau**   své
    REFL.LOC.SG   PRT DEM.ACC.PL.F   abode.LOC.SG.M/F   REFL.LOC.SG
    goṣṭhé
    goṣṭhé
    cowshed.LOC.SG.M
    sáṁveśayati
    accommodate_together.PRS.3SG
    'in ***their* abode**, in their cowshed, (he) accommodates them together' (MS I 5,9 (4))

Note that in ex. (44) and ex. (45) we find not only Left-Edge Discontinuity, but also the "focus" particle *evá* (which attaches directly to the element that it has scope over, irrespective of absolute position). The question of the exact interpretational outcomes of such combined means of marking prominent (focus) information is subject to further research.

    More research is needed to reach a deeper understanding of the range of focus interpretations associated with deviations from the default order in the languages investigated here. What does clearly emerge, however, is that deviations from the default, particularly changes of the default continuous order and Left-Edge Discontinuity, show a substantial correlation with the family of focus interpretations across languages.

6 Conclusion

This paper has examined continuous and discontinuous word order patterns of complex nominal expressions in flexible (or "free") word order languages along with any information-structural and other correlates. While main constituent order in flexible word order languages has long attracted attention, the present paper is the first comprehensive cross-linguistic study that looks at both continuous and discontinuous patterns of complex nominal expressions. Important sources were McGregor's work on continuous (1990) and discontinuous (1997) nominal expressions in Gooniyandi, Schultze-Berndt & Simard's (2012) work on discontinuity in Jaminjung, the recent survey in Louagie & Verstraete (2016), and results from a corpus study of Vedic Sanskrit prose conducted for this paper. A further important starting point was the discussion in Schultze-Berndt & Simard (2012) and Reinöhl (To appear) about how to delimit "real" complex nominal expressions from "apparent" ones in the languages at issue here. Challenges in this regard are not only word order flexibility, but also weak or non-existent noun-adjective distinctions, and the ubiquitous use of nominals in a broad range of grammatical roles.

First of all, default word order in nominal expressions in numerous of the languages under discussion appears to be determined primarily by functional templates rather than by slots tied closely to specific parts of speech. This is one of several reasons why the label "noun phrase" (think of an NP being "headed" by or "projected" from an N), does not seem an appropriate label and "nominal expression" is preferred.

Secondly, when it comes to deviations from the default pattern, the most prominent thread running through the literature and through this paper is the association with different types of focus interpretations, whether we are dealing with deviations from the default relative order in continuity or with discontinuity. The label "focus" is used as a cover term for notions such as contrast, newness or highlighting in the same way as in the literature more generally, but it is often possible to be more specific with regard to the type of focus reading at issue in a given example. Importantly, these focus interpretations are associated not only with modifying elements, but also with modified ones, or even with entire utterances in "sentence focus" interpretations.

In principle, this study corroborates the argument in Schultze-Berndt & Simard (2012) that non-default word order of nominal expressions in flexible word order languages typically signals specific, non-default interpretations. However, it is also clear that we are only dealing with correlates and not with one-to-one mappings between form and interpretation (Matić & Wedgwood 2013). Importantly, not all

deviations from the juxtaposed default necessarily map onto some kind of focus or otherwise "prominent" information-structural interpretation. For instance, some cases of non-default order in continuity seem to be due to heaviness effects instead. Furthermore, some minor types of discontinuity arise in connection with the behaviour of specific function words, rather than being associated with non-default information structure. In yet other cases, there is no clearly identifiable trigger for the deviating word order, as in some cases of Left-Edge Discontinuity in Vedic Sanskrit, where we may be dealing with a genre effect. Overall, it emerges that there are striking cross-linguistic similarities both formally and functionally. At the same time, language-specific characteristics such as the behaviour of Wackernagel elements constrain the formal patterns and functional correlates attested.

References

Amano, Kyoko. 2009. *Maitrāyaṇī-Saṁhitā I-II. Übersetzung der Prosapartien mit Kommentar zur Lexik und Syntax der älteren vedischen Prosa*. Bremen: Hempen Verlag.

Anderson, Stephen R. 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language* 69(1). 68–98.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Austin, Peter. 2001. Word order in a free word order language: The case of Jiwarli. In Jane Simpson, David Nash, Mary Laughren, Peter Austin & Barry Alpher (eds.), *Forty years on: Ken Hale and Australian languages*, 205–223. Canberra: Pacific Linguistics.

Austin, Peter & Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language & Linguistic Theory* 14(2). 215–268.

Baker, Brett & Ilana Mushin. 2008. Discourse and grammar in Australian languages. In Ilana Mushin & Brett Baker (eds.), *Discourse and grammar in Australian languages*, 1–23. Amsterdam: John Benjamins.

Baker, Mark C. 2001. Configurationality and polysynthesis. In Martin Haspelmath & Ekkehard König (eds.), *Language typology and language universals* (Handbücher zur Sprach- und Kommunikationswissenschaft 20.2), 1433–1441. Berlin/New York: de Gruyter.

Bowern, Claire. 2012. *A grammar of Bardi*. Berlin/Boston: de Gruyter.

Casaretto, Antje & Uta Reinöhl. In prep. Identifying discourse functions without clear formal clues – Secondary predicates and related functions in Vedic Sanskrit.

Chafe, Wallace. 1994. *Discourse, consciousness, and time*. Chicago: Chicago University Press.

Delbrück, Berthold. 1878. *Die altindische Wortfolge aus dem Çatapathabrāhmaṇa dargestellt*. Halle: Verlag der Buchhandlung des Waisenhauses.

Delbrück, Berthold. 2009[1888]. *Altindische syntax*. Cambridge: Cambridge University Press.

Devine, A. M. & Laurence D. Stephens. 2000. *Discontinuous syntax. Hyperbaton in Greek*. Oxford: Oxford University Press.

Donaldson, Tamsin. 1980. *Ngiyambaa. The language of the Wangaaybuwan*. Cambridge: Cambridge University Press.

Evans, Nicholas. 1995. *A grammar of Kayardild – with historical-comparative notes on Tangkic*. Berlin/New York: de Gruyter.

Fanselow, Gisbert & Caroline Féry. 2006. *Prosodic and Morphosyntactic Aspects of Discontinuous Noun Phrases: A comparative perspective*. Unpublished manuscript. University of Potsdam. http://user.uni-frankfurt.de/~cfery/publications/Prosodic_and_morphosyntactic_aspects_of_discontinuous_NPs.pdf (accessed 9 December 16).

Gillon, Brendan & Benjamin Shaer. 2005. Classical Sanskrit, 'wild trees', and the properties of free word order languages. In Katalin Kiss (ed.), *Classical languages and generative linguistics*, 281–317. Amsterdam/Philadelphia: Benjamins.

Hale, Kenneth L. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1(1). 5–47.

Halpern, Aaron L. & Arnold M. Zwicky (eds.). 1996. *Approaching second. Second position clitics and related phenomena*. Stanford: CSLI.

Harvey, Mark. 1992. The noun phrase in Australian languages: A comment. *Australian Journal of Linguistics* 12(2). 307–319.

Heath, Jeffrey. 1986. Syntactic and lexical aspects of nonconfigurationality in Nunggubuyu (Australia). *Natural Language & Linguistic Theory* 4(3). 375–408.

Himmelmann, Nikolaus & Eva Schultze-Berndt (eds.). 2006. *Secondary predication and adverbial modification – The typology of depictives*. Oxford: Oxford University Press.

Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer.

Hock, Hans Henrich. 1982. Clitic verbs in PIE or discourse-based verb fronting? Sanskrit *sá hovāca gárgyaḥ* and congeners in Avestan and Homeric Greek. *Studies in the Linguistic Sciences* 12(2). 1–38.

Hock, Hans Henrich (ed.). 1991. *Studies in Sanskrit syntax. A volume in honor of the centennial of Speijer's Sanskrit Syntax (1886–1986)*. Delhi: Motilal Banarsidass.

Hock, Hans Henrich. 1997. Chronology or genre? Problems in Vedic syntax. In Michael Witzel (ed.), *Inside the texts — beyond the texts: New approaches to the study of the Vedas*, 103–126. Cambridge, MA: Harvard University.

Hudson, Joyce. 1978. *The core of Walmatjari grammar*. Canberra: Australian Institute of Aboriginal Studies.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language & Linguistic Theory* 2(1). 39–76.

Jones, Barbara. 2011. *A grammar of Wangkajunga*. Canberra: Australian National University.

Klein, Jared S. 1985. *Toward a discourse grammar of the Rigveda. Volume I: Coordinate conjunction*. Heidelberg: Winter.

Klokeid, Terry. 1969. *Thargari phonology and morphology*. Canberra: Australian National University.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55. 243–276.

Laughren, Mary. 1984. Some focus strategies in Warlpiri, Paper presented to the annual meeting of the Australian Linguistics Society, Alice Springs.

Laughren, Mary. 1989. The configurationality parameter and Warlpiri. In Laci K. Maracz & Pieter Muysken (eds.), *Configurationality: The typology of asymmetries*, 319–353. Dordrecht: Foris.

Lee, Jennifer. 1987. *Tiwi today: A study of language change in a contact situation*. Canberra: Australian National University.

Legate, Julie Anne. 2000. Discontinuous noun phrases and the Warlpiri scope-marking construction. https://people.umass.edu/roeper/711-05/Legate–partial%20move00 (accessed 16 November 18).

Lindqvist, Urban. 2015. *Rigvedic formulas and metres*. US: Books on Demand.

Louagie, Dana & Jean-Christophe Verstraete. 2016. Noun phrase constituency in Australian languages: A typological study. *Linguistic Typology* 20(1). 25–80.

Lühr, Rosemarie. 2015. Traces of discourse configurationality in older Indo-European languages? In Carlotta Viti (ed.), *Perspectives on historical syntax*, 203–232. Amsterdam/Philadelphia: Benjamins.

Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49. 127–163.

McGregor, William. 1990. *A functional grammar of Gooniyandi*. Amsterdam: Benjamins.

McGregor, William B. 1997. Functions of noun phrase discontinuity in Gooniyandi. *Functions of Language* 4. 83–114.

Merlan, Francesca. 1994. *A grammar of Wardaman*. Berlin: Mouton de Gruyter.

Mithun, Marianne. 1987. Is basic word order universal? In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 281–328. Amsterdam: John Benjamins.

Mushin, Ilana. 2006. Motivations for second position: Evidence from North-Central Australia. *Linguistic Typology* 10. 287–326.

Nash, David. 1986[1979]. *Topics in Warlpiri Grammar*. New York/London: Garland Publishing Inc.

Nordlinger, Rachel. 1998. *Constructive case. Evidence from Australian languages*. Stanford: CSLI.

Nordlinger, Rachel. 2014. Constituency and grammatical relations in Australian languages. In Harold Koch & Rachel Nordlinger (eds.), *The languages and linguistics of Australia*, 215–262. Berlin/New York: de Gruyter.

Oertel, Hanns. 1940. *Zu den Wortstellungsvarianten der Mantras des Atharvaveda in der Śaunaka- und Paipalāda-Rezension und des Sāmaveda in der Kauthuma- und Jaiminīya-Rezension*. München: Verlag der Bayerischen Akademie der Wissenschaften.

Ong, Walter. 1982. *Orality and literacy. The technologizing of the world*. London: Methuen.

Parry, Milman. 1971[1928]. L'epithèt traditionnelle dans Homère. In Milman Parry & Adam Parry (eds.), *The making of Homeric verse. The collected papers of Milman Parry*. Oxford: Clarendon Press.

Parry, Milman (ed. Adam Parry). 1971. *The making of Homeric verse. The collected papers of Milman Parry*. Oxford: Clarendon Press.

Pullum, Geoffrey K. 1981. Evidence against the 'AUX' node in Luiseño and English. *Linguistic Inquiry* 12. 435–463.

Reinöhl, Uta. 2016. *Grammaticalization and the rise of configurationality in Indo-Aryan*. Oxford: Oxford University Press.

Reinöhl, Uta. In prep. Discontinuous nominal expressions in Vedic Sanskrit – Forms and functions.

Reinöhl, Uta. To appear. What are and what aren't complex nominal expressions in flexible word order (or 'non-configurational') languages. *STUF – Language Typology and Universals*.

Sadler, Louisa & Rachel Nordlinger. 2006. Apposition as coordination: Evidence from Australian languages, *Proceedings of the LFG06 conference*. CSLI Publications, 437–454.

Sadler, Louisa & Rachel Nordlinger. 2010. Nominal juxtaposition in Australian languages: An LFG analysis. *Journal of Linguistics* 46. 415–452.

Schäufele, Steven William. 1991a. A note on the term 'scrambling'. *Natural Language & Linguistic Theory* 9(2). 365–368.

Schäufele, Steven William. 1991b. *Free word-order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory*. Urbana-Champaign: University of Illinois dissertation.

Schultze-Berndt, Eva & Candide Simard. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics* 50(5). 1015–1058.

Shih, S. S. & K. Zuraw. 2017. Phonological conditions on variable adjective and noun word order in Tagalog. *Language* 93(4). e317–e352.

Siewierska, Anna. 1984. Phrasal discontinuity in Polish. *Australian Journal of Linguistics* 4. 57–71.

Simpson, Jane. 1991. *Warlpiri morpho-syntax. A lexicalist approach*. Dordrecht: Kluwer.

Simpson, Jane. 2006a. Expressing pragmatic constraints on word order in Warlpiri. In Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson & Annie Zaenen (eds.), *Architectures, rules, and preferences: A Festschrift for Joan Bresnan*. Stanford: CSLI Publications.

Simpson, Jane. 2006b. Depictives in English and Warlpiri. In Nikolaus Himmelmann & Eva Schultze-Berndt (eds.), *Secondary predication and adverbial modification – The typology of depictives*, 69–106. Oxford: Oxford University Press.

Simpson, Jane & Ilana Mushin. 2008. Clause-initial position in four Australian languages. In Ilana Mushin & Brett Baker (eds.), *Discourse and grammar in Australian languages*, 25–57. Amsterdam/Philadelphia: Benjamins.

Speas, Margaret. 1990. *Phrase structure in natural language*. Dordrecht: Kluwer.

Staal, J. F. 1967. *Word order in Sanskrit and Universal Grammar*. Dordrecht: D. Reidel.

Taylor, Ann. 1990. *Clitics and configurationality in Ancient Greek*. University of Pennsylvania dissertation.

Tsunoda, Tasaku. 2011. *A grammar of Warrongo*. Berlin, Boston: de Gruyter.

Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1. 333–436.