Chapter 5

HPSG in understudied languages

Douglas L. Ball
Truman State University

Work within HPSG has explored typologically-different and genetically-diverse languages, though the framework is not well-known for such explorations. This chapter details some of that work, focusing on the phenomena of argument indexing (pronoun incorporation or agreement), non-accusative alignment, and VSO constituent order. Examination of proposed and possible analyses within these areas reveals that HPSG can flexibly handle a wide range of languages all while maintaining a certain uniform "underlying structure" within the analyses.

1 Introduction

To date, the most intensely studied language within the HPSG framework has been English; this follows the trend in modern syntactic theorizing at large: English is currently the best described language in the world. Still, there has been plenty of work within HPSG on languages other than English (ISO 639-3 code: eng); in fact, substantial work has occurred within the framework[1] on German (ISO: deu) (Crysmann 2003; Müller 2013), Danish (ISO: dan) (Müller & Ørsnes 2015), Norwegian (ISO: nor) (Hellan & Haugereid 2003), French (ISO: fra) (Abeillé & Godard 2000; 2002; 2004; Abeillé et al. 2006), Spanish (ISO: spa) (Marimon 2013), Portuguese (ISO: por) (Branco & Costa 2008), Mandarin Chinese (ISO: cmn) (Müller & Lipenkova 2013; Yang & Flickinger 2014), Japanese (ISO: jpn) (Siegel et al. 2016), and Korean (ISO: kor) (Kim et al. 2011), among others. However, work within HPSG is not particularly well-known for exploring a wide range of typologically- and genetically-diverse languages, certainly not to the degree of

---

[1]Citations in this paragraph are to works, if available, whose focus is on the entire morphosyntax of the language in question rather than on particular issues in these languages.

its constraint-based lexicalist cousin, Lexical Functional Grammar. Nevertheless, there has been work within HPSG on such languages; this chapter will discuss some of this work as well as suggesting some further avenues for HPSG work within these languages.

The term I will employ for these typologically- and genetically-diverse languages is *understudied languages*. Which languages qualify as understudied languages, though? Does the term just cover languages that have not previously been investigated, syntactically? Or maybe all the languages without any previous HPSG work? Or maybe the term encompasses any language that is not the most described language, English? Though I reject all these (somewhat jocular) definitions, I do grant that *understudied language* is surely a fuzzy category, with boundaries that are difficult to demarcate and with conditions for inclusion that could be controversial. As a working benchmark for this chapter, I will suppose that the term *understudied languages* includes those languages that have a combined native and non-native speaker population of 1.2 million or fewer (roughly 0.01% of the world's population at present), that are spoken currently or have gone extinct within the last 120 years, that are generally spoken in a smaller, contiguous part of the globe, and that are not usually employed in international diplomacy or commerce. With this benchmark, languages[2] like Tongan (Polynesian, Austronesian; Tonga; ISO: ton), Kimaragang (Dusunic, Austronesian; Sabah, Malaysia; ISO: kqr), Warlpiri (Ngumpin-Yapa, Pama-Nyungan; west central Northern Territory, Australia; ISO: wbp), Burushaski (isolate; Gilgit-Baltistan, Pakistan; ISO: bsk), Lezgian (Lezgic, Nakh-Dagestanian; southern Dagestan, Russia; ISO: lez), Maltese (Semitic, Afro-Asiatic; Malta; ISO: mlt), Basque (isolate; Basque Country, Spain & France; ISO: eus), Welsh (Celtic, Indo-European; Wales, UK; ISO: cym), Oneida (Iroquoian; New York & Wisconsin, USA; Ontario, Canada; ISO: one), Coast Tsimshian (Tsimshianic; NW British Columbia, Canada & SE Alaska, USA; ISO: tsi), Yucatec Maya (Mayan; Yucatan Peninsula, Mexico & Belize; ISO: yua), and Macushi (Cariban; Roraima, Brazil, E Venuzuela, & SE Guyana; ISO: mbc) would all be included, while the ten languages mentioned in the first paragraph would not.[3]

---

[2]The locations and genetic affiliations of the languages listed here were checked at Hammarström et al. (2018) (accessed 2018-07-31).

[3]Some of the languages listed above will be discussed further in this chapter. Others from the above list are well-known understudied languages from the linguistics literature. A few of these languages have HPSG work that is not mentioned elsewhere in this chapter: on Tongan, see also Dukes (2001); on Warlpiri, see Donohue & Sag (1999); on Maltese, see Müller (2009); on Basque, see also Crowgey & Bender (2011); on Oneida, see also Koenig & Michelson (2010); on Yucatec Maya, see Dąbkowski (2017).

The denotation of the term *understudied languages* is intended to be different from *endangered language*; there is nothing in the above supposition of what an understudied language is that says that an understudied language may (or may not) potentially cease to be spoken within the next five to seventy years (see discussion in Krauss 1992 and Simons & Lewis 2013 for more on endangered languages and the crisis they face). However, the two terms do, in actuality, overlap: many understudied languages are endangered languages. While the use of HPSG (or other formal syntactic framework) or not likely has no direct bearing on the continued viability of a particular language, the practitioners of HPSG join with other linguists in seeing the importance of documenting such languages and supporting the rights of communities of endangered languages to continue to speak these languages.

Understudied languages exhibit a great variety of syntactic behaviors – some of them quite similar to "well-studied languages," some of them quite different – and these languages do not form an obvious natural class, syntactically. Due to space limitations, I will focus on just a very small portion of the syntactic phenomena of understudied languages: argument indexing, non-accusative alignment (chiefly ergativity), and VSO constituent order. These phenomena and their analyses will give the reader a sense of how HPSG has been or could be applied to understudied languages. Unfortunately, this means that a collection of phenomena made famous by understudied languages – including, among others, noun incorporation (but see Malouf 1999; Runner & Aranovich 2003; Ball 2005a,b; 2008), serial verbs (but see Muansuwan 2001; 2002; Kropp Dakubu et al. 2007; Müller & Lipenkova 2009; Lee 2014), clause-chaining, evidentiality systems (but see Lee 2012), object-initial word order, and applicatives (but see Runner & Aranovich 2003; Ball 2008; 2010) – will not be discussed.

In going through the phenomena to be discussed, it will become clear that HPSG can flexibly handle a wide range of languages even while keeping its core characteristics. In fact, in most areas of analytic interest, several different approaches within the framework are equally viable at the outset. Relatedly, the analysis of many areas, especially from a cross-linguistic perspective, is far from settled. This seems to me to be an advantage: it allows competing analyses to be modeled clearly and precisely, while allowing empirical facts to better adjudicate between approaches.[4]

In my discussion, I will move through the three areas of argument indexing (in Section 2), non-accusative alignment (Section 3), and VSO constituent order (Section 4), which corresponds to decreasing pervasiveness – roughly estimated

---

[4]This point is also made in Fokkens (2014), especially in Chapter 1.

– for each phenomenon across the world's languages.

2 Argument indexing

Widespread among all sorts of natural languages – understudied or not – is what Haspelmath (2013) terms argument indexing. In argument indexing, morphologically dependent elements – that is, affixal elements usually located within or near the verb and with denotations (seemingly) similar to pronouns – either occur in place of arguments of the main semantic predicate of the clause or alongside them.[5] While this phenomenon occurs even a bit in English and throughout other European languages, argument indexing in understudied languages tends to be more "rampant": that is, all (or most) of the verb's arguments are indexed, rather than just the subject being indexed, as is the most common pattern in Europe (Siewierska 2013b). When the argument indexing is "rampant": its treatment within the syntax (and within the morphology-syntax interface) of a language becomes a key question. HPSG analyses offer several possible answers how the syntax of argument indexing works, all while maintaining the framework's surface orientation. Empirically, it is clear that not all argument indexes behave in quite the same way in all languages, so I will explore the analysis of two subtypes of indexes in the sections to follow: first, indexes that do not co-occur with external, role-sharing noun phrases, and, second, indexes that can co-oocur with external, role-sharing noun phrases.[6]

2.1 Indexing in complementary distribution with conominals

Some argument indexes in some languages have the property that they do not – and cannot – appear with a non-pronominal element sharing their same syntactic/semantic role in the same narrow clause (I will refer to these non-pronominal elements as conominals, following Haspelmath 2013). The term that Haspelmath suggests – and I will use – for such argument indexes is pro-index. One language showing pro-index behavior with its argument indexes is Macushi, as revealed from the examples in (1):

---

[5]Thus, this area includes what has been considered to be predicate-argument agreement as well as what some consider to be "pronoun incorporation", though one of the key points of Haspelmath (2013) is that the pre-existing terminology – if not also the pre-existing analyses – in this domain has been misleading.

[6]See also Saleem (2010) for a similar – though not identical – analysis of the same analytical domain.

(1) Macushi [mbc] (Abbott 1991 via Siewierska 1999 and Corbett 2003)

    a. i-koneka-'pî-u-ya
       3SG.ABS-make-PAST-1SG-ERG
       'I made it'

    b. * uurî-ya i-koneka-'pî-u-ya
       1SG-ERG 3SG.ABS-make-PAST-1SG-ERG
       Intended: 'I made it'

The example in (1a) is just a verb with all its arguments realized as argument indexes. The example in (1b) clearly reveals the pro-index behavior of the argument indexes: the affixed verb is incompatible with an independent pronoun, such as *uurîya* '1SG.ERG'.

The pro-index phenomenon has a straightforward (and, as a result, commonly assumed) analysis within HPSG. The analysis was originally proposed by Miller & Sag (1997) for French "clitics", but could equally be applied to the Macushi case above, among others. Key to this analysis is the idea found in most versions of HPSG (emerging in the mid-to-late 1990s) that there are separate kinds of lists for the combinatorial potential(s) of heads. In fact, not only are there these separate lists, but there can be (principled) mismatches between them (see Abeillé & Borsley 2020, Chapter 1 of this volume and Wechsler, Koenig & Davis 2020, Chapter 9 of this volume). The first of these lists is the ARGUMENT-STRUCTURE (ARG-ST) list. This list handles phenomena related to the syntax-semantic interface, like linking (Davis 2001), case assignment (Przepiórkowski 1999), and binding restrictions (Manning & Sag 1998; Wechsler & Arka 1998). The other lists are the two valence lists, the SUBJECT (SUBJ) list and the COMPLEMENTS (COMPS) list. These are concerned with the "pure" syntax and mediate which syntactic elements can combine with which others.

On the Miller & Sag-style analysis of pro-indexes, the verb's ARG-ST list contains feature descriptions corresponding to all of its arguments. For the examples in (1), the verb's ARG-ST list would include a feature description for both the semantic maker and the semantic element that is made (as will appear in (3)). However, the same verb's SUBJ and COMPS lists would contain no elements corresponding to any affixed arguments. What prompts this disparity? The arguments realized by affixes correspond to a special kind of feature description on the ARG-ST list, typed *non-canonical*.[7] (Intuitively, these arguments are realized

---

[7]In the version of HPSG of Ginzburg & Sag (2000), *non-canonical* was an immediate subtype of *synsem*; and the relevant feature descriptions were thus seen as syntactico-semantic complexes. In the latter-day version of HPSG known as Sign-Based Construction Grammar (Sag 2012),

in a non-canonical way.) Feature descriptions of the *non-canonical* type differ from their sibling type *canonical* in how they interact with the SUBJ and COMPS lists. Governing the relationship between the ARG-ST and these valence lists is the Argument Realization Principle, which is stated in (2):[8]

(2) *word* $\Rightarrow$ $\begin{bmatrix} \text{SS|LOC|CAT} \begin{bmatrix} \text{SUBJ} & \boxed{\text{B}} \\ \text{COMPS} & \boxed{\text{C}} \end{bmatrix} \\ \text{ARG-ST } \boxed{\text{A}} \;\; list(canonical)(= \boxed{\text{B}} \oplus \boxed{\text{C}}) \bigcirc \boxed{\text{D}} \; list(non\text{-}canonical) \end{bmatrix}$

(cf. Ginzburg & Sag 2000: 171)

The constraint in (2) says that only the list of *canonical* feature descriptions on the ARG-ST can (and must) appear on the SUBJ and COMPS lists. Thus, the principle disallows any *non-canonical* argument from being mapped to either of these valence lists. This, then, captures the idea that such affixal arguments are generally inert in the combinatorics of the syntax proper: they saturate an argument slot and that argument slot is no longer available for other (at least direct) syntactic combination.

So, returning to the Macushi word *ikoneka'pîuya* 'I made it' from (1a), it will have the partial lexical description in (3):

(3) $\begin{bmatrix} \text{SS|LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & verb \\ \text{SUBJ} & \langle \rangle \\ \text{COMPS} & \langle \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \text{RELS} \left\langle \begin{bmatrix} make\text{-}rel \\ \text{ARG1} & \boxed{1} \\ \text{ARG2} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST } \left\langle \begin{bmatrix} non\text{-}canonical \end{bmatrix}_{\boxed{1}}, \begin{bmatrix} non\text{-}canonical \end{bmatrix}_{\boxed{2}} \right\rangle \end{bmatrix}$

In (3), the word's ARG-ST list is comprised of two *non-canonical* feature descriptions (corresponding to the maker argument with index $\boxed{1}$ and the made argu-

---

the *non-canonical* type was re-christened *covert* and was an immediate subtype of *sign*, the relevant feature descriptions being entire signs. In spite of these differences, which may seem significant, the analysis is very similar in both versions of the framework. Several subtypes of *non-canonical*/*covert* have been recognized, the subtype relevant for this example would be *aff*. But I will just use the *non-canonical* type here.

[8]The append operator $\oplus$ allows two lists to be combined, preserving the pre-existing order of the elements on the new list. Thus, $< a, b > \oplus < c, d >$ will yield $< a, b, c, d >$. The shuffle operator $\bigcirc$ allows the elements to be re-ordered provided the relative order of the elements in the individual lists stays constant. Thus, $< a, b > \bigcirc < c, d >$ results in a disjunctive specification of the following six orders:

$< a, b, c, d >, < a, c, b, d >, < a, c, d, b >, < c, a, b, d >, < c, a, d, b >, < c, d, a, b >$. See Reape (1994) for more, especially on the shuffle operator.

Draft of 29th September 2020, 16:53

ment, with index ②). Yet, by the Argument Realization Principle, the SUBJ and COMPS lists are empty. Consequently, *ikoneka'pîuya* 'I made it', can be a clause by itself – as it is in (1a) – because it requires no other valents (that is, it has a "saturation" level on a par with a clause) and it is headed by a verb (just like a clause is).

The specification of an empty list also contributes to ruling out examples like (1b) with a conominal. On the standard HPSG view of how valence is managed, no element with an empty valence list can combine with any possible valence-saturating syntactic entity, like an NP.[9] Thus, the grammar would correctly not license a tree like in Figure 1.[10]
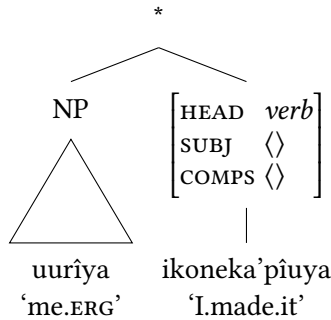

Figure 1: A tree of an illicit conominal–pro-index combination

Overall, the Argument Realization Principle-mandated non-mapping of the *non-canonical* ARG-ST list members to either the SUBJ or the COMPS list captures the key behavior found in the pro-indexing type of argument indexing: the argument indexes occur in complementary distribution with any conominal.

---

[9]To truly rule out the NP from combining with the verb in Figure 1, the NP would also need to not match any non-local requirements of the verb.

[10]The tree in Figure 1, as well as other trees in this chapter, only provides the relevant attribute-value pairs, suppressing the geometry of features found in more articulated feature descriptions.

2.2 Indexing co-occurring with conominals

Even though the pro-index type of argument index has a more straightforward analysis, this type is not the most common in the world's languages. Rather, the most common type of argument indexing appears to be the one where the argument indexing affix(es) can co-occur with a conominal, but do(es) not have to. In Haspelmath's (2013) terms, this is the cross-index type.[11] A language exhibiting this type of behavior is Basque, as evident from the example in (4):

(4)   Basque [eus] ([Laka 1996]: 98)
      Zuk   niri   liburua   saldu d-i-da-zu.
      2SG.ERG 1SG.DAT book.DEF sold   3SG.ABS-AUX-1SG.DAT-2SG.ERG
      'You have sold me the book.'

Though *zuk* 'you', *niri* 'me', or even *liburua* 'the book' would not need to be present for the grammaticality of this sentence, this sentence (and language) exhibits cross-index behavior because, even though these conominals are present, the argument indexing affixes on, in this case, the auxiliary *didazu* '3SG.ABS:AUX:-1SG.DAT:2SG.ERG' still occur.

Unlike the pro-indexes, there is no current standard HPSG analysis of cross-indexes.[12] Nevertheless, there are some possible approaches. I detail two in some depth here – what I will call the underspecification analysis and what I refer to as the direct syntax approach – and mention some other options near the end of the section.

2.2.1 Underspecification for Cross-Indexes

On the underspecification analysis, the lexical descriptions of argument index-containing words would have underspecified feature descriptions on their ARG-ST lists, corresponding to their argument indexes. These would then resolve depending on the syntactic context. Which portions of the feature description would be underspecified is a bit flexible (at least, in the abstract) and depends on whether the analyst thought the "agreement" (argument indexing) was more

---

[11]The behavior of cross-indexes is canonical for so-called "pro-drop" languages, a term arising from the transformational syntax tradition (particularly from [Chomsky 1981]), but now with wider currency.

[12]This may, in part, be a consequence of the standard way of managing predicate-argument relations in the syntax in HPSG: this management strategy is resource-sensitive – namely, once something is "cancelled" off a SUBJ or COMPS list, it no longer appears on any subsequent (higher) lists and cannot be used for other syntactic purposes. However, Section 2.2.3 will discuss some HPSG approaches where the management strategy is not so resource-sensitive.

Draft of 29th September 2020, 16:53

formal or semantic in nature (see Wechsler 2020, Chapter 6 of this volume for a more thorough discussion of what is involved here). For the sake of illustration purposes, I will employ a more semantic approach below.

Let us consider a word, like the Basque auxiliary *dut* 'aux:3.abs:1sg.erg', that has a third-person singular absolutive argument index. Such a word might just be specified, by the constraints on the various lexical types of Basque, as in (5):[13]

(5) $\left[\text{ARG-ST} \left\langle \left[\textit{synsem} \right], \begin{bmatrix} \textit{synsem} \\ \text{SS|LOC|CONT|IND } \textit{3sg} \end{bmatrix}, \cdots \right\rangle \right]$

To be consistent with (5), the second ARG-ST list member just needs to be something that is semantically a third person. Therefore, the second ARG-ST list member could ultimately resolve to a *non-canonical* feature description, as in (6):

(6) $\left[\text{ARG-ST} \left\langle \left[\textit{synsem} \right], \begin{bmatrix} \textit{non-canonical} \\ \text{SS|LOC|CONT|IND } \textit{3sg} \end{bmatrix}, \cdots \right\rangle \right]$

This resolution would be forced when no conominal is present (if this *synsem* on the ARG-ST list resolved to the *canonical* type and a conominal was not present, the COMPS list would illicitly not be emptied). The analysis would, in this condition, be identical to that of the pro-indexes provided in Section 2.1.

However, the second ARG-ST list member could also ultimately resolve to a *canonical* feature description, as in (7):

(7) $\left[\text{ARG-ST} \left\langle \left[\textit{synsem} \right], \begin{bmatrix} \textit{canonical} \\ \text{SS|LOC|CONT|IND } \textit{3sg} \end{bmatrix}, \cdots \right\rangle \right]$

This resolution would be forced when a conominal is present (otherwise, the conominal could not be syntactically licensed). Thus, the analysis, in this condition, is like an instance of obligatorily co-present conominal and argument index (a gramm-index in Haspelmath's terms).

As the discussion above indicates, there is a certain portion of this analysis that is not lexically mandated: the precise resolution of the argument depends on the specific syntactic expressions appearing in a particular clause. This analysis is also of the dual-nature type discussed by Haspelmath (2013): the argument index is treated as a pro-index when it has no conominal and it is treated as gramm-index when a conominal is present. Other frameworks employ a similar analysis

---

[13]The lexical descriptions associated with *dut* in (5)–(7) all have a further argument – the verbal expression associated with the auxiliary – suppressed in these descriptions (with ellipses) because such a verbal argument (and its interaction with the other arguments) is not the focus of the analysis here.

(LFG does, for instance – see Bresnan et al. 2016: Chapter 8). Haspelmath criticizes this approach for positing two distinct structural types for a single kind of affix; though, in the analysis above, it does not seem that the structural types are that radically different (observe that just one underspecified lexical description is associated with a given affixed form). Still, we might want to at least consider other options – and, in keeping with the tendency for multiple different approaches to be found within HPSG, there are some.

2.2.2 The "direct syntax" approach to cross-indexes

Another approach to cross-indexes, proposed for Oneida in Koenig & Michelson (2015), takes the view that, in at least some languages, argument indexes always stand for arguments and the combination of a conominal and a verb with argument indexing is more purely semantically mediated and akin to a nominal expression combining with an already saturated narrow clause.[14] This approach Koenig & Michelson have called the "direct syntax approach" (it is direct in the sense that the combinatorics are not mediated by any valence lists, which, arguably, is a bit more "indirect").

As Koenig & Michelson (2015) discuss in detail, it appears that Oneida exhibits some interesting properties that make treating its argument indexing patterns in a (seemingly) rather different way much more plausible. For one, as shown in (8), the verb indexes all its arguments morphologically (except inanimates, like 'his axe' in (9)) – often with portmanteau affixes, as in (8) – making the case that the argument indexes are the actual arguments much stronger.

(8)   Oneida [one] (Koenig & Michelson 2015: 5)
      wa-hiy-até·kw-a-ht-eʔ
      FACT-1SG>3.M.SG-flee-LNK.V-CAUS-PNC
      'I chased him away.'

Second, the evidence is equivocal about whether the language has any selection that cannot be treated as semantic selection.

Thus, on Koenig & Michelson's view (and in keeping with the terminology of the previous discussion): all the arguments correspond (at best) to *non-canonical* elements on the ARG-ST list and thus there is never any head-argument combinations in the syntax. Any and all conominals then are licensed via index sharing of a nominal and an element on a NON-LOCAL feature that Koenig & Michelson

---

[14]This analysis is perhaps the closest any HPSG analysis comes to the so-called Pronominal Argument Hypothesis (Jelinek 1984).

call DISLOC (see [Koenig & Michelson 2015](): 39 for discussion of why they consider this the best way to deal with the NON-LOCAL feature), as shown in Figure [2](), a tree of [(9)]():

(9) Oneida [one] ([Koenig & Michelson 2015](): 17)
Λ-ha-hyoʔthi·yát-eʔ   laoto·kΛ·,
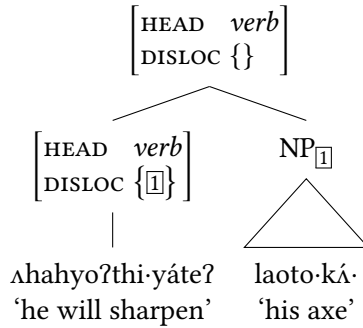FUT-3M.SG.A-sharpen-PNC his.axe
'he will sharpen his axe,'

Figure 2: Licensing conominals on the direct syntax approach

Koenig & Michelson's (2015) discussion suggests that the direct syntax type might represent an extreme, occurring only in the most polysynthetic and non-configurational of languages, like Oneida and its Iroquoian kin. However, this claim remains an open question. Perhaps further study will reveal that this sort of analysis could profitably employed in other kinds of languages.

2.2.3 Other possibilities for cross-indexes

In addition to the analyses discussed in the previous two subsubsections, there are a few more conceptual avenues that might be explored for the analysis of cross-indexes, though it is not clear that they have been fully explored in the literature yet (which might raise some questions as to their viability).

One route to explore would be to use lexical rules to create ARG-ST or the valence lists with feature descriptions corresponding to both the argument indexes *and* the conominals (something similar was explored for "clitics" in various Romance varieties by [Monachesi 2005]()). Such a lexical rule might look as in [(10)]():

$$(10) \quad \begin{bmatrix} \text{ARG-ST} \; \boxed{A} \; \left\langle \begin{bmatrix} canonical \end{bmatrix}, \begin{bmatrix} non\text{-}canonical \\ \text{SS}|\text{LOC}|\text{CONT}|\text{IND} \; \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \mapsto$$
$$\begin{bmatrix} \text{SS}|\text{LOC}|\text{CAT}|\text{COMPS} \; list \oplus \left\langle \begin{bmatrix} canonical \\ \text{SS}|\text{LOC}|\text{CONT}|\text{IND} \; \boxed{1} \end{bmatrix} \right\rangle \\ \text{ARG-ST} \qquad\qquad \boxed{A} \end{bmatrix}$$

This approach might be a way to loosen the resource sensitivity of the usual valence regime, though a proposal along these lines would need to take care in considering whether any changes would be needed in the statement of the Argument Realization Principle (and if so, what form they should take) and if there would be any undesirable consequences to allowing single semantic arguments to correspond to more than one syntactico-semantic element.

Another route to consider would be to relax the resource sensitivity in the syntax, instead of in the information associated with single words. Given proposals like the one in Bender (2008) for non-cancellation of arguments (more detailed discussion of this proposal is in Section 6 of Müller 2020, Chapter 10 of this volume[15]) to deal with apparent cases of discontinuous constituency, maybe something similar should also be explored for the cross-index type of argument indexing.

Overall, there seems to be a need for more explorations into cross-index behavior cross-linguistically within HPSG. Certainly, the above discussion shows that, in fact, there is no shortage of possible analyses, but work remains to further determine which of these might be the best analysis, overall, or which of these might be best for which languages.

3 Non-accusative alignments

Another area (in fact, not so distant from argument indexing in function) where understudied languages have enriched the general understanding of natural language morphosyntax is in the area of (morphosyntactic) alignment. Alignment concerns how the morphology of a language (if not also its syntax) groups together (or "aligns") different arguments together into (what seem to be) particular grammatical relations (see Bickel & Nichols 2009 for an overview of alignment).[16] The most widespread alignment is the accusative one – familiar from ancient Indo-European languages and conservative modern day ones – where

---

[15]Also see like-minded proposals in Meurers (1999) and Müller (2008).

[16]Alignment can be explored both in head-marking and dependent-marking (Nichols 1986); however, having already focused on a kind of head-marking strategy in the previous section, I will focus on the corresponding dependent-marking strategy in this section.

subjects of transitive and intransitive verbs are treated differently from the objects of transitive verbs. Other recognized alignments include ergative (where subjects of transitive verbs are treated differently from the subjects of intransitives, which, in turn, pattern with the direct objects of transitives) (see Comrie 1978; Plank 1979; Dixon 1979; 1994, among others, for further discussion), split-S/active (where semantic agents and patients are treated differently) (see Klimov 1973; 1974; Dixon 1994: Chapter 4; Mithun 1991; Wichmann & Donohue 2008, among others, for further discussion), tripartite (where subjects of transitive verbs, subjects of intransitive verbs, and objects of transitive verbs are each treated differently) (Dixon 1994: 39–40), Austronesian alignment[17] (where arguments of various semantic roles can flexibly hold a privileged syntactic slot) (see Schachter 1976; Ross 2002; Himmelmann 2005 for more discussion), and hierarchical alignment (where elements of higher discourse salience are treated differently from elements of lower discourse salience) (see Jacques & Antonov 2014 for a good overview of what is involved).

Surveys from WALS (Comrie 2013a,b; Siewierska 2013a) indicate that accusative alignment is common worldwide,[18] and this seems to be even more true of languages with large numbers of speakers. Of the top 25 most widely spoken languages at present, arguably only a collection of languages from the Indian subcontinent (Hindi-Urdu, Marathi, and Gujarati) have non-accusative alignments, and even those are restricted to certain portions of their respective verbal systems (see Verbeke 2013: Chapter 7 for more on the patterns in these and other Indo-Aryan languages). Impressionistically, it seems that understudied languages do have a much stronger propensity for non-accusative alignments.

Because the non-accusative alignments at least seem to be rather different than accusative alignment, it is an interesting question how a given framework might handle these kinds of systems. In the majority of this section I will focus on the analysis of ergative systems, as a proof of concept (see, however, Drellishak 2009 for analyses of each of the non-accusative alignments, including the hierarchical type).[19]

---

[17]This kind of system is known by various different names other than Austronesian alignment, including a symmetrical voice system, a Philippine-type voice system, or an Austronesian focus system.

[18]However, since the surveys focus more on coding patterns rather than behavioral patterns (coding and behavioral in the sense of Keenan 1976 – "coding" related to morphological patterns or function words that signal a particular grammatical relation category; "behavioral" related to reference properties or patterning across clauses), it is possible that they underreport behavioral accusative patterns, even among languages that have so-called "neutral" coding patterns.

[19]There is also some discussion of an HPSG analysis of the ergative-aligned case system of the

In dealing with the analysis of ergative systems, it will be useful to divide the discussion into two parts. First, I will consider how particular morphological forms within NPs are licensed in instances when they co-occur with their governing verb – I will call this "the licensing of case in the syntax" (see also Przepiórkowski 2020, Chapter 7 of this volume). Second, I will considered how particular arguments come to be associated with particular morphological forms (whether realized or not) – I will call this "the licensing of case in linking" (see also Wechsler, Koenig & Davis 2020, Chapter 9 of this volume). This division is not commonly recognized in most other frameworks; however, it does present itself as a possible division within HPSG, due to the separate ARG-ST and valence lists.

3.1 The licensing of case in the syntax

The licensing of case in the syntax within HPSG is as straightforward in non-accusative alignments as it is in an accusative alignment system; the fundamentals are the same, regardless of alignment. This comes about due to the use of feature value matching (also known as "feature unification") for case licensing in the syntax.[20] The simple premise of feature value matching is that a value for a particular feature possessed by an argument and a feature value required by its head (for that same argument) must match. The nature of this analysis makes case licensing nearly identical – excepting the different values involved – to the selection of part-of-speech categories.[21]

To actually license case in the syntax with an ergative system, the key elements are (1) a feature for nominal expressions (call it CASE) and (2) appropriate values for CASE, like *ergative* and *absolute*. Note that the values of *ergative* and *absolute* are typed, so they can be potentially grouped with other case values into supertypes, like *structural* cases or *semantic* cases, if such groupings are relevant (as was done for the first time in Heinz & Matiasek 1994). With those features in place, the rest of the analysis falls out through the larger theories of

---

Caucasian language Archi – similar, in some respects, to the Lezgian examples I consider further on – in Borsley (2016), though the focus of that paper is much more on Archi's argument indexing system rather than its case system.

[20]Feature value matching does have some conceptual similarity to the "feature checking" approach to case found in more recent Minimalist work (Chomsky 1991; 1993; Adger 2000; 2010; Frampton & Gutmann 2006; Pesetsky & Torrego 2007), though there are notable differences between the approaches, particularly that features are deleted in feature checking, but not in feature value matching.

[21]Thus, to use terms more commonly associated with Mainstream Generative Grammar, HPSG views case licensing as (a specific kind of) c-selection (in the sense of Grimshaw 1979).

syntactic selection, featural identities, and syntactic combination: certain heads will require [CASE *erg(ative)*] and [CASE *abs(olutive)*] of their arguments. If certain potential arguments are just single words, the values for CASE of these words will straightforwardly match or not. If certain potential arguments consist of multiple words, independent constraints on HEAD value identity will ensure that the value for CASE will be identical between the head daughter and overall phrase; constraints on the syntactic combination then ensure that the CASE values of the nominal expressions and the head requirements match.

To see this with an actual example, let us consider the Lezgian sentence in (11):

(11) Lezgian [lez] (Haspelmath 1993: 287)
      Aburu   zun   ajibda.
      3PL.ERG 1SG.ABS shame.FUT
      'They will shame me.'

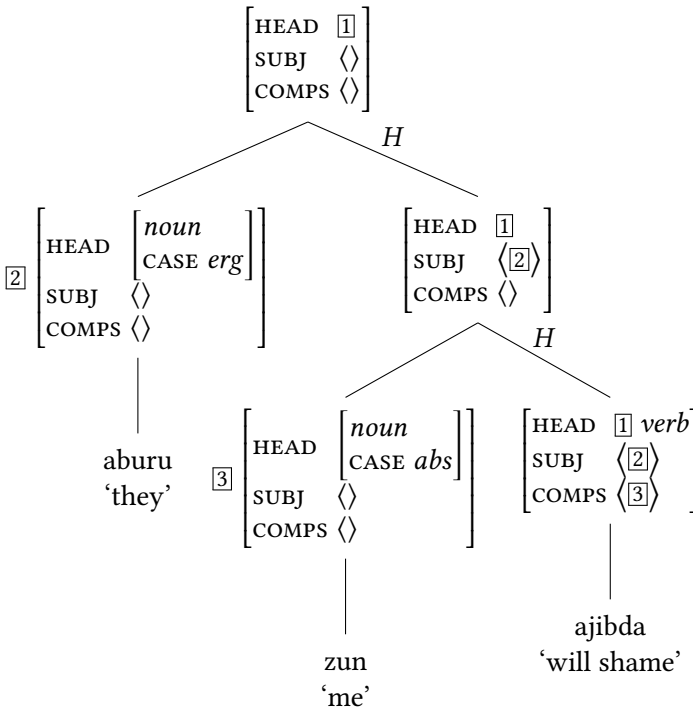The example in (11) could be analyzed with the tree in Figure 3.

Figure 3: A tree of (11)

The tree in Figure 3 consists of two head-argument combinations, and in fact, the tree has the same geometry as an accusative verb-final language (on standard assumptions about the constituency) – indeed, as the HPSG analysis does not intrinsically tie the analysis of case with constituency, the geometry of clauses would, all else being equal, not differ based on alignment alone. The most idiosyncratic aspect of Figure 3 is that the verb *ajibda* 'will shame' is one that requires an ergative–absolutive combination of arguments. Because the HPSG framework is feature-rich and formally rigorous in how feature values must be constrained within constituent structures, the licensing of case in the syntax in an HPSG analysis is very straightforward.

3.2 The licensing of case in linking

Lurking behind the most idiosyncratic aspect of Figure 3 is the question of how particular heads come to have their particular argument requirements. This is, in fact, the question of how case is licensed in linking. As with other matters of non-accusative alignments, it seems that different alignments need not be treated in a wholly different fashion from each other: thus, the same kinds of analytic moves for accusatively aligned systems could be used for non-accusatively aligned systems. That being so, it is probably too hasty to assume that there is a one-size-fits-all solution for linking of case across all languages (regardless of alignment), as quite a few different factors appear to be important in different languages, among them at least verb class (that is, the classes related to the verbal lexical semantics), the semantic nature of the argument itself, the morphological form of the verb, and the subordination status of the clause headed by the verb (see, for example, discussion in Dixon 1994).

In all known ergative languages, the ergative–absolutive case pattern – clearly indicating that the subjects of transitive verbs are not encoded like the subjects of intransitive verbs – appears with "primary transitive verbs" (a term from Andrews 1985; 2007): predicates with the canonical meaning associated with transitive verbs where an initiating entity causes change in an undergoing entity. Given this basic generalization, a possible analysis of the arguments' case requirements with these "primary transitive verbs" would be through the constraint in (12):

(12) $trans\text{-}v\text{-}lxm \Rightarrow$
$$\begin{bmatrix} \text{SS|LOC|CONT|KEY} & \begin{bmatrix} act\text{-}und\text{-}rel \\ \text{ARG1} \;\boxed{1} \\ \text{ARG2} \;\boxed{2} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{CASE } erg \end{bmatrix}_{\boxed{1}}, \begin{bmatrix} \text{CASE } abs \end{bmatrix}_{\boxed{2}} \right\rangle \oplus list \end{bmatrix}$$

In (12), the *transitive-verb-lexeme* (*trans-v-lxm*) has an ARG-ST list with both an ergative and an absolutive argument. Key to this result is that the verb lexeme is associated with an *actor-undergoer-relation* (*act-und-rel*), and, in fact, this is the value of the KEY feature in (12), encoding the designated semantic relation relevant for case and linking (see Koenig & Davis 2006 for more on the KEY feature). The *act-und-rel* type designates semantic predicates with precisely the denotation behind the notion of "primary transitive verb" (see Davis 2001: 75–134 for discussion of this type, other related types, and how these types fit into a hierarchy of semantic relations). Provided that the constraint in (12) is the only argument realization constraint to mention the ergative and absolutive cases, the ergative–absolutive collection of arguments would only be available–with verbs with this particular sort of meaning.[22] The overall linking constraint is placed on a *trans-v-lxm*, so that (a) the case requirements are stated once for all the different inflected forms of the verb and (b) so these case requirements could also be inherited by other semantically appropriate verbs with even more arguments than the bivalent predicate illustrated in (12).

As alluded to above, in other case "assignment" situations, though, other factors beyond just the semantics of the verb can be relevant (certainly with any instance of "split ergativity", among others). These could be still be treated with constraints with a similar format to (12), but if they needed to refer to, say, just past tense verb forms, the relevant linking constraint would almost assuredly need to reference information from the morphology (perhaps encoded as part of a MORPH(OLOGY) attribute). Given the known claims about what non-accusative alignment can be sensitive to, it seems likely that the sign-based architecture (where all linguistic areas of structure can interact in parallel) would enable the straightforward statement of case constraints based on the previously claimed generalization. And, in fact, having the possibilities of morphological form, semantics, and various syntactic properties easily available for an analysis could be useful as a means of testing and modeling which areas might be relevant in particular examples.

Overall, there is a lot still to be done to better understand the intricate details of case and linking generally, but given the toolbox available in the HPSG framework (again, see Przepiórkowski 2020, Chapter 7 of this volume and Wechsler, Koenig & Davis 2020, Chapter 9 of this volume), it seems like HPSG offers a lot of flexibility for better figuring out what linguistic elements are crucial for particular patterns and for encoding analyses that directly reference the interaction

---

[22]Though to achieve more generality with the licensing of absolutive case, one might follow Ball (2008: Chapter 7) and have separate linking constraints for absolutive and ergative case.

of these elements across different levels of structure.

4 Verb-subject-object constituent order

Let us turn to another interesting phenomenon of understudied languages: Verb-Subject-Object (VSO) constituent order.[23] VSO order appears to be the rarest of the more common orders. Various typology surveys (like Dryer 2013) indicate that it is only found in about 8–10% of the world's languages. Interestingly, a greater number of examples of languages with this order do come from the realm of understudied languages. Of the twelve understudied languages (a non-random sample[24]) mentioned in the introduction, five of them have VSO (or verb-initial with no strict ordering of S and O) order. Perhaps a bit more telling of the (apparent) understudied language bias to VSO order is that only one of the top 50 languages by native speakers – Tagalog – reasonably clearly has VSO order.[25] Interestingly, VSO order does occur in a number of languages as a non-dominant word order: for instance, it is found in a great many western European languages (English, German, French, Spanish, among others) as a common order in questions.

In spite of its relative rarity as a dominant order, VSO order (as well as verb-initial order with flexible ordering of verbal dependents) poses some interesting challenges for frameworks that place some importance in constituency (as HPSG has done). Since the V and the O are not (normally) adjacent in VSO clauses, it is less than obvious that there is a constituent that groups them together (as a VP or a V′) in these languages. This contrasts with the more common Subject–Verb–Object and Subject–Object–Verb orders where a constituent that groups the V and O together is much more plausible, on surface adjacencies alone. A long-standing question across constituency-based frameworks is how to best characterize VSO order, both on its own and in the context of the other cross-linguistically attested and common order patterns.

Analyses within Mainstream Generative Grammar have generally analyzed

---

[23]It would probably be clearer to refer to this order as Predicate–Agentive–Patientive order, because, as noted in the previous section, alignment and constituent order are, to a degree, disjoint. However, I will bow to tradition and use the terms verb, subject, and object and their abbreviations, V, S, and O.

[24]See fn. 3 for the rationale behind the choice of those twelve languages.

[25]Tagalog's distant Austronesian relatives Indonesian, Javanese, and Sundanese, along with Arabic – all four of these languages are also in the top 50 – had VSO, historically, and each of these languages preserves some instances of VSO order. As is often the case when looking at word orders and languages, things are rarely as cut and dry as they might otherwise seem.

Draft of 29th September 2020, 16:53

VSO as a derived order; all (or nearly all) of them (especially after the 1970s) have viewed VSO as a derived permutation of some constituent (or more) from a covert SVO order (see Clemens & Polinsky 2017 for an overview of the analyses within this tradition). Some of the suggested HPSG analyses follow a similar line of analysis, and I will briefly touch on those proposals below. However, more HPSG analysts have generally taken VSO order as is, and so I spend more of this section discussing two surface-oriented VSO analyses in HPSG: what I call the flat structure analysis and what I call the binary branching head-initial analysis.

4.1 The analogues of verb movement in HPSG

Interestingly, there is not one, but two styles of HPSG analyses that are roughly analogous to Mainstream Generative Grammar's verb movement, commonly employed to derive VSO order. Both are discussed in greater detail in Müller (2020), Chapter 10 of this volume, so my comments here will be somewhat superficial and will center around verb-initiality. The first of the two uses the DOUBLESLASH feature (see Section 4.1 of Müller 2020, Chapter 10 of this volume) and so treats the initial verb as involved in a long-distance dependency. Thus, in keeping with other long-distance dependencies in HPSG, information related to the initial verb is passed through the constituent structure to the verb's downstairs position (a trace, with semantic and syntactic structure, though no phonological realization). While this analysis has been explored for some Germanic languages (see Figure 2 in Müller 2020, Chapter 10 of this volume for a pictorial depiction of an analysis of an English verb-initial clause), I am not aware that it has (yet) been seriously explored in the HPSG literature for any particular verb-initial language (let alone for an understudied verb-initial language).

The other verb-movement-like analysis uses constituent order domains and linearization (see Section 5 of Müller 2020, Chapter 10 of this volume. On this analysis, the verb, while combined with its complements at a low level, is constrained at the clausal level to be initial (see Borsley 2006 for more discussion of this in a verb-initial context). This style of analysis has been closely and carefully considered for Welsh in work by Borsley (for example in Borsley 1989; 1995; 2009), but time and again, it seems that Borsley suggests that an analysis (at least for the basics of clausal structure) more in line with what is discussed in Section 4.2 is to be preferred for Welsh.

Given the rarity (and perhaps reluctance) – noted above – of HPSG researchers to analyze VSO order as covertly SVO (or, more to the point, to recognize a constituent that groups together the V and O within VSO structures), one might wonder why this has (hitherto) been so. Probably, HPSG's surface orientation

has played a role, as well as the fact that HPSG-internal considerations do not force or strongly suggest positing a VP constituent. Furthermore, HPSG analysts have also carefully considered how constituency tests might inform such structures. In exploring these, various HPSG researchers (such as Borsley 2006 for Welsh and Ball 2008: Chapter 3 for Tongan) have not found compelling evidence for positing a VP constituent in particular VSO languages.[26] For instance, Ball, in looking at Tongan, found that: putative VP-coordination "over" a subject is not possible; no auxiliary or verb obviously subcategorizes for a verbal constituent that obviously excludes its subject, nor do adverbial elements obviously select for such a constituent; and, while "VP-fronting" and "VP-ellipsis" are possible, they seem to involve NPs rather than VPs. While these facts do not definitively rule out a VP (it is difficult to argue that anything is clearly absent), they suggest that not positing a VP does not complicate the grammar of this kind of language. Undoubtedly, it would be interesting to see what further explorations like these with more verb-initial language might reveal. Still, the VSO-as-covert-SVO analysis may lie on shakier grounds empirically than analyses within Mainstream Generative Grammar have generally acknowledged and this, explicitly or implicitly, has led HPSG analysts to explore other avenues in the analysis of VSO languages.

---

[26]The undermotivated VP in Welsh is probably just a VP headed by a finite verb, as Welsh does give evidence for non-finite VPs (Borsley et al. 2007). In other languages, like Tongan, the undermotivated VPs might include both finite and non-finite VPs. As has emerged in the study of verb-initial languages in several frameworks, these languages might not be as structurally uniform as the term "verb-initial languages" suggests.

4.2 The flat structure analysis

The seemingly most common analysis of VSO languages in HPSG is the flat structure analysis.[27] As its name suggests, the proposed structure is flat, with the verb, subject NP, and any complement NPs all being sisters within the same constituent. To license such a structure, one has to depart from rules that put just heads and complements or just heads and subjects together. The flat structure analysis instead makes use of what I call the Head-All-Valents Schema (also sometimes called the Head-Subject-Complements Schema), given in (13):

(13)   *head-all-valents-schema* ⇒

$$
\begin{bmatrix}
\text{SS|LOC|CAT} & \begin{bmatrix} \text{SUBJ} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\
\text{HD-DTR} & \begin{bmatrix} \textit{word} \\ \text{SS|LOC|CAT} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, ..., \boxed{n} \rangle \end{bmatrix} \end{bmatrix} \\
\text{NON-HD-DTRS} & \langle \boxed{1}, \boxed{2}, ..., \boxed{n} \rangle
\end{bmatrix}
$$

Per its name, it licenses a fully saturated phrase comprising a head – a single word – and all its valents (subject, object, and whatever else). This schema has not just been used for canonical VSO clauses within HPSG, but other clause-level head-initial structures, including polar questions in English. Thus, this schema has a long pedigree in the HPSG literature (compare the schema in (13) with Schema 3 from Pollard & Sag 1994: 40; *sai-ph* from Ginzburg & Sag 2000: 36; and *aux-initial-cxt* from Sag 2012: 188).

To see an example using the Head-All-Valents Schema, let us consider example (14) from Kimaragang:

---

[27]There are, in fact, several alternative flat structure analyses, differing slightly in how the head's valence features relate to the structure. Besides having the head combine with its subjects and complements simultaneously, as in the main text, one variant has all the arguments as complements and, thus, VSO order arises out of a head-complements structure. Borsley (1995) suggests that different languages might utilize different variants: in particular, Borsley suggests that Syrian Arabic uses the head-subject-complements combination while Welsh uses the head-complements combination. Still other analysts (such as Ball 2008; 2017) utilize a mixture of these two: having all arguments on just one valence feature, but combining all arguments with the head at once.

(14) Kimaragang [kqr] (Kroeger 2010: 7)
      Minangalapak it   kogiw   do ratu.
      PST.AV.TR.split NOM orangutan GEN durian
      'The orangutan split (open) a durian.'

By the Head-All Valents Schema (and appropriate inherited constraints concerning the featural identities of HEAD values), a tree for (14) would be as in in Figure 4.
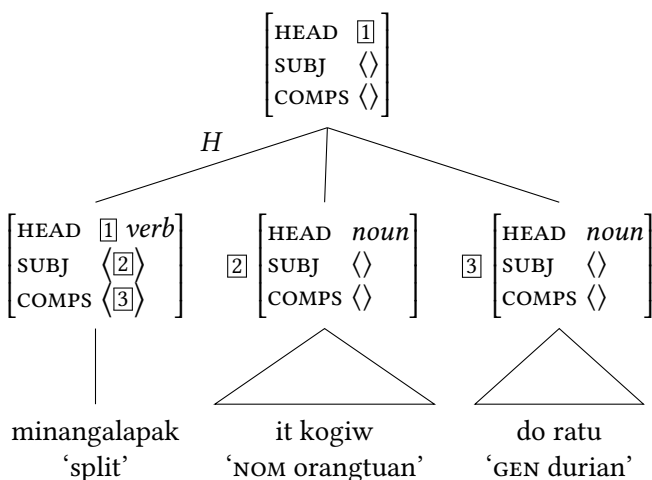

Figure 4: Tree of (14) in the flat structure analysis

To license the tree in Figure 4, we first should observe that the verb *minangalapak* 'split' appears to require both a nominative and a genitive argument. With two such nominal expressions fitting those requirements available, the Head-All-Valents Rule can put all three of these elements – the verb and two NPs – together, and the resulting mother node's SUBJ and COMPS lists would be empty.

In spite of the flatness of Figure 4, the structure is like all head-nexus combinations in HPSG: a head and (at least some of) its dependents. In fact, Figure 4 is identical to certain verb phrases headed by a ditransitive verb (on some HPSG analyses) – just a verb and two NPs. Furthermore, the flat nature of the structure is less of a concern than it would be under c-command-based proposals (which are the off-the-shelf analyses in Mainstream Generative Grammar): binding relations in HPSG are not calculated from the configurations within the tree, but from configurations on the ARG-ST list (see Müller & Branco 2020, Chapter 20 of this volume). Other subject-object and agent-patient asymmetries (to the extent

they exist) are likewise encoded in HPSG analyses using non-configurational data structures and do not seem to be relevant for determining constituency.

Additionally, assuming a flatter structure for VSO/verb-initial languages eases the analysis of several other phenomena (especially versus a treatment of the same data with a VP constituent). In verb-initial languages where the order of elements following the verb is flexible (as in Tongan, among others), having all arguments together with the verb as part of a single constituent allows for such "scrambling" to be analyzed with simple linear precedence constraints within that constituent, rather than having to deal with different orders across a VP boundary (see the analysis in Ball 2008: Chapter 3 for Tongan and Müller 2020, Chapter 10 of this volume for other HPSG approaches to "scrambling"). There are also a few languages like Coast Tsimshian, where morphological marking (somewhat surprisingly) on one syntactic item refers to the next constituent over. An example of this phenomenon is given in the Coast Tsimshian sentence in (15), where the second line employs brackets to better show which elements are related to which others:

(15)   Coast Tsimshian (Sm'algya̲x) [tsi] (Mulder 1994: 32)
       Yagwat huumda duusa hoon.
       Yagwa-t   huum-[da   duus]-[a   hoon]
       cont-3.erg smell-[erg.cn cat]-[abs.cn fish]
       'The cat is sniffing the fish.'

It is far more straightforward to analyze the apparent sideways relationships when the interacting elements are sisters, rather than to manage the relationships across a VP boundary (and possibly other constituent boundaries) (see Ball 2011 for an in-depth look into this syntactic phenomenon in Coast Tsimshian and an analysis of it).

4.3 The binary branching head-initial analysis

Another approach in HPSG to VSO structures takes the view that all verb-headed structures within the clause are maximally binary branching, but strongly head-initial. This approach still has a strong surface-orientation – so it does not take the VSO order to be covertly SVO or SOV – but does posit that more structure is present within a clause than on the flat structure analysis.

On the binary branching head-initial analysis, VSO clauses are built out of several instances of a single rule. The rule, which I call the Head-Valent Schema, is given in (16):[28]

----

[28]The Head-Valent Schema here is designed to implement the Categorial Grammar analysis of

(16)   *head-valent-schema* $\Rightarrow$

$$
\begin{bmatrix}
\text{SS|LOC|CAT|VAL } \boxed{A} \\
\text{HD-DTR} \qquad \begin{bmatrix} \text{SS|LOC|CAT|VAL } \langle\boxed{1}\rangle \oplus \boxed{A} \end{bmatrix} \\
\text{NON-HD-DTRS } \langle\boxed{1}\rangle
\end{bmatrix}
$$

The rule in (16) allows a head to combine with just one of its valents; in particular, the first one on its VAL list. This aspect of the ordering is crucial to ensure that the subject-NP-before-object-NP sequence is licensed.

Returning to the Kimaragang example of (14), we can see how a structure licensed by the Head-Valent Schema in (16) differs from a structure licensed by the Head-All-Valents Schema. The structure licensed by the Head-Valent Schema (and relevant inherited constraints) is given in Figure 5.
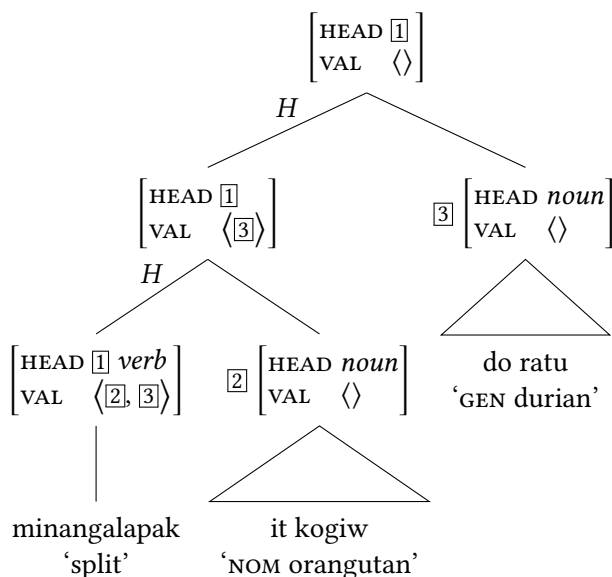
Figure 5: Tree of (14) in the binary branching head-initial analysis

Like in Figure 4, in Figure 5, the head verb requires a nominative and a genitive argument. However, instead of combining with both of these at the same time,

Draft of 29th September 2020, 16:53

the verb just combines with the initial nominative argument (②), leaving the genitive argument (③) to be passed up to the mother. At this second level of structure, the Head-Valent Schema again applies – because there is still at least one element on the relevant head's VAL list – integrating ③ into the structure. The rule is barred, correctly, from applying to the root node of the tree in Figure 5, as this root node has an empty VAL list and the Head-Valent Schema requires the head daughter to have at least one valent.

A noteworthy feature of the VSO binary branching head-initial analysis is its grouping of verb and the subject NP into a constituent. Exactly this sort of thing has been reported to occur in some verb-initial languages, like Malagasy (Keenan 2000), suggesting the binary branching head-initial analysis might be preferable for such languages. In VSO languages without such evidence, it would seem that either the flat structure analysis or the binary branching head-initial analysis would be possible, all else being equal.

If one is accustomed to seeing the trees from Mainstream Generative Grammar, the structure in Figure 5 may still seem strange (notably, the structural prominence relationships between what seems to be the subject NP and what seems to be the object NP are reversed). Nevertheless, many of the same kinds of comments made for the flat structure analysis hold here as well. The structure in Figure 5 is a just a series of head-argument structures, the most common kind of structure in HPSG. And, once again, the non-configurational approach to binding in HPSG renders any issues related to tree configuration and binding as irrelevant.

Both approaches to VSO order discussed above do raise interesting questions about whether there are any underlying grammatical principles, processing preferences, or historically-driven outcomes behind the patterns. For the binary branching head-initial analysis, there is a question as to why the required order of combination goes from from least oblique to most oblique. A similar set of question can be leveled to the flat structure analysis: what inhibits a more constituent-rich structure? Why are flat structures licensed here and not elsewhere? To my knowledge, these questions have yet to be tackled within the HPSG literature, but they do seem to be reasonable next steps, in addition to better seeing which analyses are appropriate for which verb-initial languages.

5 Wrapping Up

In general, HPSG practitioners have been fairly conservative in what they assume to be universal in syntax: since there is no core assumption in HPSG that

particular rich, innate, and universal class of structures help children learn any language (Mainstream Generative Grammar's Universal Grammar), proposals can be (and are) made that are agnostic as to universality. Even so, the brief trip made in this chapter through argument indexing, non-accusative alignments, and verb-initial constituent order found in understudied languages reveals that the more dependency-oriented portions of the framework – in particular, the areas encoded in the SUBJ, COMPS, and ARG-ST lists – are useful for the analysis of all three of these areas, across different languages, and, thus, are candidates for universality[29] (though the current level of understanding does not clearly point to them originating from either a rich language-specific part of cognition or from general cognition). Furthermore, the explorations above show that the rich and precise modeling using attribute-value matrices also allows for uniform sorts of analyses, even though the details may differ.[30] While the precise attributes and feature values may not completely be candidates for universality, they certainly aid in the enterprise of exploring different analyses and determining what precisely must be said to capture certain linguistic phenomena.

In addition to revealing some of the more uniform aspects of HPSG, the above discussion also reveals a certain flexibility in how the framework can be deployed – several analyses might be possible and certain ones might be more appropriate for certain languages and not for others. Thus, on top of a uniform foundation, various languages and phenomena are open to be analyzed in their own terms, dependent on what the specific empirical facts reveal. This mesh of uniformity and parochiality in HPSG analyses seems to strike a good balance as grammarians try to capture the two (somewhat paradoxical) realities one finds when comparing across languages: languages are both surprisingly similar and surprisingly different.

Acknowledgements

---

[29]Or in the case of the SUBJ and COMPS lists, a candidate for near-universality, as Koenig & Michelson (2015) argue that Oneida does not require such lists.

[30]Two projects within the HPSG community have explored in-depth how uniform particular HPSG analyses of different languages might be. The Grammar Matrix project (Bender et al. 2010) just starts from a common core and adds language-specific elements as needed; The CoreGram project (Müller 2015) actively tries to use the same sorts of data structures for as many languages as possible within the project.

     Draft of 29th September 2020, 16:53

representative when I was first approached to write this chapter), for suggesting that a chapter like this one exist in this handbook.

References

Abbott, Miriam. 1991. Macushi. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, 23–160. Berlin: Mouton de Gruyter.

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2006. The syntax of French *à* and *de*: An HPSG analysis. In Patrick Saint-Dizier (ed.), *Syntax and semantics of prepositions* (Text, Speech and Language Technology 29), 147–162. Springer Verlag. DOI: 10.1007/1-4020-3873-9_10.

Abeillé, Anne & Robert D. Borsley. 2020. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 3–44. Berlin: Language Science Press. DOI: ??.

Abeillé, Anne & Danièle Godard. 2000. French word order and lexical weight. In Robert D. Borsley (ed.), *The nature and function of syntactic categories* (Syntax and Semantics 32), 325–360. San Diego: Academic Press.

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Abeillé, Anne & Danièle Godard. 2004. The syntax of French adverbs without functional projections. In Martine Coene, L. Tasmowski, Gretel de Cuyper & Yves d'Hulst (eds.), *Current studies in comparative romance linguistics: Proceedings of the international conference held at the antwerp university (19-21 september 2002) to honor liliane tasmowski* (Antwerp Papers in Linguistics), 1–39. University of Antwerp.

Adger, David. 2000. Feature checking under adjacency and VSO clause structure. In Robert D. Borsley (ed.), *The nature and function of syntactic categories*, 79–100. New York: Academic Press.

Adger, David. 2010. A Minimalist theory of feature structure. In Anna Kibort & Greville G. Corbett (eds.), *Features: perspectives on a key notion in linguistics* (Oxford Linguistics), 185–218. Oxford: Oxford University Press.

Andrews, Avery D. 1985. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description: clause structure*, 1st Edition, 62–154. Cambridge: Cambridge University Press.

Andrews, Avery D. 2007. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description: clause structure*, 2nd, 132–223. Cambridge: Cambridge University Press.

Ball, Douglas L. 2005a. Phrasal incorporation in Tongan. In Jeffrey Heinz & Dimitrios Ntelitheos (eds.), *Proceedings of the Twelfth Annual Conference of Austronesian Formal Linguistics Association (AFLA)*, vol. 12, 19–33. UCLA Working Papers in Linguistics. http://phonetics.linguistics.ucla.edu/wpl/issues/wpl12/papers/Ball_AFLA12.pdf.

Ball, Douglas L. 2005b. Tongan noun incorporation: lexical sharing or argument inheritance? In Stefan Müller (ed.), *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 7–27. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2005/ball.pdf.

Ball, Douglas L. 2008. *Clause structure and argument realization in Tongan*. Stanford University. (Doctoral dissertation).

Ball, Douglas L. 2010. A unified analysis of Niuean *Aki*. In Raphael Mercado, Eric Potsdam & Lisa deMena Travis (eds.), *Austronesian and Theoretical Linguistics*, vol. 167 (Linguistik Aktuell/Linguistics Today), 249–270. Amsterdam/Philadelphia: John Benjamins.

Ball, Douglas L. 2011. Morphology in the 'wrong' place: the curious case of Coast Tsimshian connectives. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 25–45. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2011/ball.pdf.

Ball, Douglas L. 2017. ''VP" adverbs without a VP: the syntax of adverbs in Tongan. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 61–81. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-ball.pdf.

Bender, Emily M. 2008. Radical non-configurationality without shuffle operators: An analysis of Wambaya. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 6–24. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/ (18 August, 2020).

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation* 8(1). 10.1007/s11168-010-9070-1, 23–72. http://dx.doi.org/10.1007/s11168-010-9070-1.

Bickel, Balthasar & Johanna Nichols. 2009. Case marking and alignment. In Andrej L. Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 304–321. New York: Oxford University Press.

Borsley, Robert D. 1989. An HPSG approach to Welsh. *Journal of Linguistics* 25. 333–354.

Borsley, Robert D. 1995. On some similarities and differences between Welsh and Syrian Arabic. *Linguistics* 33(1). 99–122.

Borsley, Robert D. 2006. On the nature of Welsh VSO clauses. *Lingua* 116(4). 462–490. DOI: 10.1016/j.lingua.2005.02.004.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language & Linguistic Theory* 27(2). 225–265. DOI: 10.1007/s11049-009-9067-3.

Borsley, Robert D. 2016. HPSG and the nature of agreement in Archi. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective* (Oxford Studies of Endangered Languages), 118–149. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780198747291.003.0005.

Borsley, Robert D., Maggie Tallerman & David Willis. 2007. *The syntax of Welsh* (Cambridge Syntax Guides 7). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486227.

Branco, António & Francisco Costa. 2008. *A computational grammar for deep linguistic processing of Portuguese: LXGram, version A.4.1.* Tech. rep. TR-2008-17. Universidade de Lisboa, Faculdade de Ciências, Departamento de Informática.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax.* 2nd edn. (Blackwell Textbooks in Linguistics 16). Oxford: Wiley-Blackwell. DOI: 10.1002/9781119105664.

Chomsky, Noam. 1981. *Lectures on government and binding* (Studies in Generative Grammar 9). Dordrecht: Foris Publications. DOI: 10.1515/9783110884166.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in Generative Grammar*, 417–454. Reprint as: Chomsky (1995: 129–166). Cambridge, MA: MIT Press.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Clemens, Lauren Eby & Maria Polinsky. 2017. Verb-initial word orders (primarily in Austronesian and Mayan languages). In Martin Everaert & Henk van

Riemsdijk (eds.), *The Wiley-Blackwell companion of syntax*, 2nd edn. Hoboken, NJ: Wiley-Blackwell.

Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: studies in the phenomenology of language*, 329–294. Austin: University of Texas Press.

Comrie, Bernard. 2013a. Alignment of case marking in full noun phrases. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/98 (1 August, 2018).

Comrie, Bernard. 2013b. Alignment of case marking of pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/99 (1 August, 2018).

Corbett, Greville G. 2003. Agreement: The range of the phenomenon and the principles of the Surrey Database of Agreement. *Transactions of the Philological Society* 101(2). 155–202.

Crowgey, Joshua & Emily M. Bender. 2011. Analyzing interacting phenomena: Word order and negation in Basque. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 46–59. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2011/crowgey-bender.pdf (18 August, 2020).

Crysmann, Berthold. 2003. On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, 112–116. Borovets, Bulgaria.

Dąbkowski, Maksymilian. 2017. Yucatecan control and lexical categories in SBCG. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 162–178. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-dabkowski.pdf.

Davis, Anthony R. 2001. *Linking by types in the hierarchical lexicon*. Stanford, CA: CSLI Publications.

Dixon, R.M.W. 1979. Ergativity. *Language* 55(1). 59–138.

Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Donohue, Cathryn & Ivan A. Sag. 1999. Domains in Warlpiri. Manuscript, Stanford University.

Drellishak, Scott. 2009. *Widespread but not universal: Improving the typological coverage of the Grammar Matrix*. University of Washington. (Doctoral dissertation).

Dryer, Matthew S. 2013. Order of subject, object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/81 (18 August, 2020).

Dukes, Michael. 2001. The morphosyntax of 2P pronouns in Tongan. In Dan Flickinger & Andreas Kathol (eds.), *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, 63–80. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/1/hpsg00dukes.pdf.

Fokkens, Antske Sibelle. 2014. *Enhancing empirical research for linguistically motivated precision grammars*. Department of Computational Linguistics, Universität des Saarlandes. (Doctoral dissertation).

Frampton, John & Sam Gutmann. 2006. How sentences grow in the mind: agreement and selection inefficient minimalist syntax. In Cedric Boeckx (ed.), *Agreement systems*, 121–157. Amsterdam: John Benjamins.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10(2). 279–326.

Hammarström, Harald, Robert Forkel & Martin Haspelmath. 2018. *Glottolog 3.3*. published online by Max Planck Institute for the Science of Human History. Jena. DOI: 10.5281/zenodo.1321024.

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntax of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: Mouton de Gruyter.

Heinz, Wolfgang & Johannes Matiasek. 1994. Argument structure and case assignment in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 199–236. Stanford, CA: CSLI Publications.

Hellan, Lars & Petter Haugereid. 2003. Norsource – An excercise in the Matrix Grammar building design. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the ESSLLI 2003 Workshop "Ideas and Strategies for Multilingual Grammar Development"*. Vienna, Austria.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Him-

melmann (eds.), *The Austronesian langauges of Asia and Madagascar*, 110–181. London: Routledge.

Jacques, Guillaume & Anton Antonov. 2014. Direct/inverse systems. *Language and Linguistics Compass* 8(7). 301–318.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2(1). 39–78.

Keenan, Edward L. 1976. Towards a universal definition of subject. In Charles N. Li (ed.), *Subject and topic*, 303–333. New York: Academic Press.

Keenan, Edward L. 2000. Morphology is structure: a Malagasy test case. In Ileana Paul, Vivianne Phillips & Lisa Travis (eds.), *Formal issues in Austronesian linguistics*, 27–47. Dordrecht: Kluwer.

Kim, Jong-Bok, Jaehyung Yang, Sanghoun Song & Francis Bond. 2011. Deep processing of Korean and the development of the Korean Resource Grammar. *Linguistic Research* 28(3). 635–672.

Klimov, Georgij A. 1973. *Očerk obščej teorii ergativnosti [Outline of a general theory of ergativity]*. Moscow: Nauka.

Klimov, Georgij A. 1974. On the character of languages of active typology. *Linguistics* 12(131). 11–26.

Koenig, Jean-Pierre & Anthony R. Davis. 2006. The key to lexical semantic representations. *Journal of Linguistics* 42(1). 71–108. DOI: 10 . 1017 / S0022226705003695.

Koenig, Jean-Pierre & Karin Michelson. 2010. Argument structure of Oneida kinship terms. *International Journal of American Linguistics* 76(2). 169–205.

Koenig, Jean-Pierre & Karin Michelson. 2015. Invariance in argument realization: The case of Iroquoian. *Language* 91(1). 1–47.

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68(1). 4–10.

Kroeger, Paul R. 2010. The grammar of *hitting*, *breaking*, and *cutting* in Kimaragang Dusun. *Oceanic Linguistics* 49(1). 1–20.

Kropp Dakubu, Mary Esther, Lars Hellan & Dorothee Beermann. 2007. Verb sequencing constraints in Ga: Serial verb constructions and the extended verb complex. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 99–119. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/ (18 August, 2020).

Laka, Itziar. 1996. *A brief grammar of Euskara, the Basque language*. Manuscript, University of the Basque Country. http://www.ehu.es/grammar (1 August, 2018).

Lee, Juwon. 2012. The direct evidential *-te* in Korean: Its interaction with person and experiencer predicate. In Stefan Müller (ed.), *The Proceedings of the 19th*

International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon, 251–270. http://cslipublications.stanford.edu/HPSG/2012/lee.pdf.

Lee, Juwon. 2014. Two types of serial verb constructions in Korean: subject-sharing and index-sharing. In Stefan Müller (ed.), *The Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo*, 135–155. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2014/lee.pdf.

Malouf, Robert. 1999. West Greenlandic noun incorporation in a mono-hierarchical theory of grammar. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 47–62. Stanford, CA: CSLI Publications.

Manning, Christopher D. & Ivan A. Sag. 1998. Argument structure, valence, and binding. *Nordic Journal of Linguistics* 21. 107–144.

Marimon, Montserrat. 2013. The Spanish DELPH-IN grammar. *Language Resources and Evaluation* 47(2). 371–397. DOI: 10.1007/s10579-012-9199-7.

Meurers, Walt Detmar. 1999. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html (18 August, 2020).

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language & Linguistic Theory* 15(3). 573–639. DOI: 10.1023/A:1005815413834.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67(3). 510–546.

Monachesi, Paola. 2005. *The verbal complex in Romance: a case study in grammatical interfaces.* Oxford: Oxford University Press.

Muansuwan, Nuttanart. 2001. Directional serial verb constructions in Thai. In Dan Flickinger & Andreas Kathol (eds.), *The proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar: University of California, Berkeley, 22–23 July, 2000*, 229–246. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2001/ (18 August, 2020).

Muansuwan, Nuttanart. 2002. *Verb complexes in Thai.* University at Buffalo, the State University of New York. (Doctoral dissertation).

Mulder, Jean Gail. 1994. *Ergativity in Coast Tsimshian (Sm'algyax).* Berkeley: University of California Press.

Müller, Stefan. 2008. Depictive secondary predicates in German and English. In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary*

predicates in Eastern European languages and beyond (Studia Slavica Oldenburgensia 16), 255–273. Oldenburg: BIS-Verlag.

Müller, Stefan. 2009. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: Papers from the 1st International Conference on Maltese Linguistics (Bremen/Germany, 18–20 October, 2007)* (Studies in Language Companion Series 113), 83–112. Amsterdam: John Benjamins Publishing Co.

Müller, Stefan. 2013. *Head-Driven Phrase Structure Grammar: Eine Einführung.* 3rd edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI: 10.15398/jlm.v3i1.91.

Müller, Stefan. 2020. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 349–393. Berlin: Language Science Press. DOI: ??.

Müller, Stefan & António Branco. 2020. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 841–851. Berlin: Language Science Press. DOI: ??.

Müller, Stefan & Janna Lipenkova. 2009. Serial verb constructions in Chinese: An HPSG account. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 234–254. Stanford, CA: CSLI Publications.

Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In Huei-ling Lai & Kawai Chui (eds.), *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei, Taiwan: Department of English, National Chengchi University.

Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). In Preparation. Berlin: Language Science Press.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62(1). 56–119.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: syntactic derivation and interpretation, in honor of Joseph E. Emonds*, 262–294. Amsterdam: John Benjamins.

Plank, Frans (ed.). 1979. *Ergativity: Towards a theory of grammatical relations.* London: Academic Press.

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Przepiórkowski, Adam. 1999. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach.* Eberhard-Karls-Universität Tübingen. (Doctoral dissertation). https : / / publikationen . uni - tuebingen.de/xmlui/handle/10900/46147 (18 August, 2020).

Przepiórkowski, Adam. 2020. Case. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 229–257. Berlin: Language Science Press. DOI: ??.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems*, 17–62. Canberra: Pacific Linguistics.

Runner, Jeffrey T. & Raúl Aranovich. 2003. Noun incorporation and rule interaction in the lexicon. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University*, 359–379. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2003/runner-aranovich.pdf.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Saleem, Safiyyah. 2010. *Argument optionality: a new library for the grammar matrix customization system.* University of Washington. (MA thesis). http://www.delph-in.net/matrix/saleem-thesis.pdf.

Schachter, Paul. 1976. The subject in Philippine languages: topic, actor, actor-topic, or none of the above. In Charles N. Li (ed.), *Subject and topic*, 317–363. New York: Academic Press.

Siegel, Melanie, Emily M. Bender & Francis Bond. 2016. *Jacy: An implemented grammar of Japanese* (CSLI Studies in Computational Linguistics). Stanford, CA: CSLI Publications.

Siewierska, Anna. 1999. From anaphoric pronoun to grammatical agreement marker: Why objects don't make it. *Folia Linguistica* 33(2). 225–251.

Siewierska, Anna. 2013a. Alignment of verbal person marking. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/100 (1 August, 2018).

Siewierska, Anna. 2013b. Verbal person marking. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/102 (1 August, 2018).

Simons, Gary F. & M. Paul Lewis. 2013. The world's languages in crisis: a 20-year update. In Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Responses to language endangerment. in honor of Mickey Noonan*, vol. 142 (Studies in Language Companion Series), 3–19. Amsterdam: John Benjamins.

Verbeke, Saartje. 2013. *Alignment and ergativity in New Indo-Aryan languages*. Berlin: Walter de Gruyter.

Wechsler, Stephen. 2020. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 205–228. Berlin: Language Science Press. DOI: ??.

Wechsler, Stephen & I. Wayan Arka. 1998. Syntactic ergativity in Balinese: An argument structure based theory. *Natural Language and Linguistic Theory* 16(2). 387–442.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2020. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 299–347. Berlin: Language Science Press. DOI: ??.

Wichmann, Søren & Mark Donohue (eds.). 2008. *The typology of semantic alignment*. Oxford: Oxford University Press.

Yang, Chunlei & Dan Flickinger. 2014. ManGO: Grammar engineering for deep linguistic processing. *New Technology of Library and Information Service* 30(3). 57–64.

Draft of 29th September 2020, 16:53

Part II

Syntactic phenomena