Chapter 10

Constituency in Hup: Synchronic and diachronic perspectives

Patience Epps

The University of Texas at Austin

This chapter provides a fine-grained description of the result of constituency diagnostics applied to Hup (Naduhup family, NW Amazonia). It describes the planar structures in the verb and noun complex, the set of morphosyntactic constituency diagnostics considered and their outcomes as applied over the verb and noun complexes, and the phonological domains applied over the noun and verb complexes. This chapter is based on dozens of hours of naturalistic speech and a large number of sentences from elicitation (see ailla.utexas.org). As this chapter explores, a feature of particular typological and theoretical relevance concerning Hup constituency is the fact that the criterion of non-interruptability does not apply straightforwardly in the Hup verb – a challenge for proposals that non-interruptability is a key test for wordhood cross-linguistically. The degree of mismatch among morphosyntactic and phonological criteria as defining particular spans as units also has practical implications, in that it creates difficulties in establishing principled conventions for representing the orthographic word. These mismatches are also implicated in the varying assessments of the Naduhup languages as isolating vs. polysynthetic.

1 Introduction

Languages of the western Amazon have been observed to have a relatively fuzzy distinction between morphology and syntax, challenging the view that a discrete divide separates these two areas of grammar (Payne 1990, Tallman 2020). A more precise prediction that emerges from this generalization is that diagnostics of constituency should be relatively non-convergent, such that one diagnos-

tic might not align closely with another. This expectation may apply on a synchronic level, relating to diagnostics applied both across and within languages, but also on a diachronic one: It is grounded in the observation that processes of grammaticalization lead to independent elements becoming more bonded (i.e. forming tighter units of constituency) over time;[1] but it also allows for the possibility that elements may become less bonded – a process that has to do in part with a mismatch in constituency criteria across domains, such that escaping one can lead to escaping others.

This chapter considers these questions through an investigation of constituency in Hup, a member of the small Naduhup language family of the northwest Amazon, following the procedure developed by Tallman (2021). Through the application of various constituency diagnostics to verb and noun structure, I show that the different measures of constituency in this language are notably non-convergent, especially in verbal constructions. Among particular challenges to assumptions relating to wordhood, I note the failure of the non-interruptability test for Hup verbs, which allow particular etyma to intervene and/or switch positions among other morphological elements; another is the lack of convergence among different phonological domains relevant to assessments of wordhood in Hup. As I argue below, an account of these mismatches is further illuminated by a diachronic perspective, which underscores the motivations behind particular diagnostic outcomes, and highlights the way in which developments that might be construed as relatively minor can have major implications for wordhood. The discussion in this chapter is informed by original research in collaboration with Hup speakers, and draws on dozens of hours of naturalistic speech and a large number of sentences from elicitation.

The chapter is organized as follows. Section §2 introduces the Hup language within the context of the Naduhup family, and §3 describes the planar structures in the verb and noun complex, and the categories of elements that make them up. Section §4 considers a set of morphosyntactic constituency diagnostics and their outcomes as applied over the verb and noun complexes, and §5 does the same for phonological domains. Section §6 offers some diachronic and comparative considerations; §7 concludes.

---

[1]The use of the term *bonded* in this chapter refers to relative tightness of constituency, while *bound* is reserved for nouns that require a preceding nominal element within a compound construction (see §3.2).

2 Hup and the Naduhup language family

Hup is spoken in the Upper Rio Negro region, in the border area of Brazil and Colombia. Like its Naduhup sister-languages Yuhup, Dâw, and Nadëb, the speakers of Hup traditionally inhabit the interfluvial zones of the Rio Negro region (see Figure 1). Hup has approximately 2500 speakers (according to a 2017 regional census) – the most of any of the Naduhup languages, with Dâw comprising the smallest population at about 130 speakers. Hup is still robustly transmitted to children, and while most Hup speakers today are bilingual in Tukano (and probably have been fluent in a range of East Tukanoan languages over time), only a few are competent in Portuguese.

Figure 1: Hup and the Naduhup languages

While the Naduhup languages were formerly lumped together with Kakua and Nukak (and, by some accounts, Puinave to the north) as the "Makú" family, comparative evidence indicates that they constitute an independent grouping (Epps & Bolaños 2017). According to our current understanding of relationships within the family, Hup and Yuhup are quite closely related, while Dâw is a more distant sister, and Nadëb occupies a distinct primary branch (Figure 2). Despite a clearcut signal of genetic relationship in the lexicon, the languages are typologically

divergent, due in large part to grammatical restructuring driven by contact with their respective neighbors – in particular, East Tukanoan languages on the part of Hup and Yuhup, and (probably) Arawakan languages on the part of Nadëb (Epps 2007, Epps & Bolaños 2017).
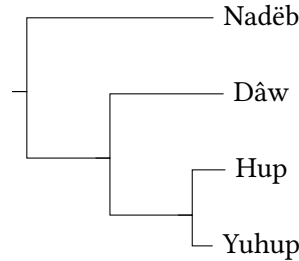

Figure 2: The Naduhup language family

Like its sister languages, Hup received only minimal description through the early 21st century. The discussion in this chapter draws on my work with the language, informed by approximately 18 months of fieldwork carried out between 2000 and 2016. A comprehensive grammatical description of Hup is provided in Epps (2008a); an extensive corpus of natural discourse and elicited material is housed in the Archive for the Indigenous Languages of Latin America (Epps 2001), and a dictionary of Hup is also available (Ramirez 2006). The Hup examples in this chapter are all drawn from my corpus, and most of them are also represented in Epps (2008a). More information on all aspects of Hup grammar addressed here can be found in Epps (2008a).

I turn here to a brief overview of Hup grammar, as context for the following discussion. Hup has nine vowels (i, ɨ, u, e, ə, o, æ, a, ɔ) and nineteen consonants, including a glottalized series (p, t, c, k, ʔ, b, d, ɟ, g, b', d', ɟ', g', ç, h, w, j, w', j'). Hup also has two contrastive tones, realized on stressed syllables, and prosodic nasalization, by which most morphemes are entirely nasal or entirely oral.[2] Constituent order is strongly verb-final, although with some flexibility associated with information structure (see §3.1.1 below); core arguments

---

[2]The orthographic conventions used in this chapter (which follow those in Epps 2008a) are the following: The representation of consonants and vowels is consistent with the IPA values, with the exception of <j> for /ɟ/ and <y> for /j/, and the use of /m, n, ŋ/ to represent /b, d, g/ in nasal contexts. With reference to the latter convention, morpheme-level nasalization is conveyed orthographically by the presence of nasal segments within the morpheme (i.e. either consonants or vowels, although all are in fact targets for nasalization within the relevant domain). Finally, tone is indicated via diacritics above the vowel in the relevant syllable: v́ indicates high tone (or its falling allophone); v̌ indicates rising tone.

are frequently dropped. Alignment is robustly nominative-accusative and favors dependent-marking, with case-marking on some core and oblique nominal arguments. Grammatical categories marked within the verb complex include tense, aspect, mood, evidentiality, associated motion, and others; some of these elements can also associate with nouns, as explored below, while nouns also may receive number marking and classifiers. Many of these features of Hup grammar match those seen in Tukanoan and Arawakan languages of the same region, and at least some are undoubtedly the outcome of contact-driven restructuring (see Epps 2007, 2008c, *inter alia*).

In Hup, verbs in particular tend to occur in complex serial expressions, in which as many as five or six conjoined roots may be followed by multiple grammatical formatives (example (1)). These formatives can be distinguished on phonological and morphosyntactic grounds into several categories, as discussed in more detail in §3 below. They are given the following labels: *inner suffixes* (such as telic *-yɨʔ* in example (1)), *boundary suffixes* (of which one, and normally only one, occurs obligatorily on verbs in nearly all contexts; e.g. the *dynamic* suffix *-V́y*), *enclitics* (e.g. reported evidential =*mah*), and associated *particles* (effectively enclitics as well, but which tend to occur further toward the end of the verb complex and are more phonologically independent). As discussed below, Hup also has a restricted set of prefixes/proclitics.

(1) *ʔapɨd   nutkán   puhu-hi-cɨ̆p-kəd-cak-yɨ́ʔ-ɨ́y=mah*
     right.away here.DIR swell-FACT-complete-pass-climb-TEL-DYNM=REP
     'Right away it had already swelled up and spread quickly up to here [on her leg], it's said.'

While there has been very little evaluation of constituency in Hup beyond my own work, the representations of the orthographic word – particularly for verbs – are notably variable across the different sources that do exist, a fact that reflects the relative lack of convergence among the constituency diagnostics explored below. In Epps (2008a) and the examples provided here, phonological criteria relating to stress rules have been taken as a key guide for the orthographic representation of the word, together with domains defined by obligatory morphological combinations (particularly the boundary suffix position). These morphological domains also relate to constraints on free occurrence and interruptability (i.e. whether certain classes of morphemes cannot occur independently, and what can intervene between them), as explored below.

On the other hand, in the practical orthography that has recently been adopted in Hup communities (and also used for Ramirez' 2006 dictionary), consonant-

initial morphemes in both verbal and nominal constructions are normally represented as separate orthographic words. This convention reflects the fact that CVC/CVV is both a minimal phonological word and the minimal freely occurring morphological unit in the language, alongside the fact that Hup strongly prefers single-syllable morphemes (see §5 below). As such, the verbal construction in (1) above would be orthographically represented as (2) in these materials. Similar criteria may have been applied in the brief description of Hup by SIL missionaries Moore & Franklin (1980), referenced in Payne's (1990: 219) comment that "the two dominantly isolating families in the [Lowland South American] region are Jê and Makú [Naduhup]".

(2) *puhu hicĩp   kəd cak yíʔ-ɨy   mah*
     swell FACT.complete pass climb TEL-DYNM REP
     '...already swelled up and spread quickly up to here, it's said.'

3 Planar structures

This section presents the planar structures for the verb and noun complexes, following the approach presented in Tallman (2021). These are based on flattening out and elaborating the template representations and/or phrase structure rules of Hup, as described in detail in Epps (2008a). The discussion provided in this and the following sections attends most closely to verb structure, in light of the overall complexity of its morphological composition and implications for constituency diagnostics. Noun structure is included for comparative purposes and provides context for the diachronic considerations addressed in §6 below.

The majority of the positions in the planar structures are slots, in which only one element can occur at a time; among these slots, relatively fixed ordering applies. A few positions are zones, in which multiple elements may co-occur with variable ordering. The relative ordering of elements within zones is primarily determined by scope, as in the case of the evidentials in position 23 of the verb structure below (Table 1), or the stacking of multiple possessors in the nominal construction (position 2, Table 2). The order of compounded verb roots (position 6, Table 1) tends to reflect temporal iconicity, and the order of NP arguments of verbs (positions 1 and 30, Table 1) is sensitive to information structure. The precise assignment of ordered slots and zones in the tables below is informed by extensive work on this language, as noted above, but some points may be refined by further testing.

3.1 Verbal structure

Table 1 introduces the verbal structure.

Table 1: Verbal planar structure in Hup

Position	Type	Elements
(1)	Zone	NP{A,S,O,Oblique}
(2)	Slot	Subject Proclitic {S, A} (marginal)
(3)	Zone	Valence: (causative root), Interactional ʔũh-, Reflexive/passive hup-, O (simplex)
(4)	Slot	Causative root
(5)	Slot	Factitive hi-
(6)	Zone	**Verb base (one or more compounded roots)**
(7)	Slot	Telic -yɨʔ
(8)	Slot	Venitive -ʔay
(9)	Slot	Applicative -ʔũh
(10)	Slot	Completive -cĩp / -cĩw
(11)	Slot	Emphasis -pog
(12)	Slot	Counterfactual -tæ̃ʔ
(13)	Slot	Perfective -ʔeʔ
(14)	Slot	Clausal negative -nɨh
(15)	Slot	Habitual -bɨg , Distributive -pɨd, Future -teg
(16)	Slot	evidentials -hɔ̃, -cud, -mah, frustrative -yæ̃h, repetitive -bʼay
(17)	Slot	Inchoative -ay
(18)	Slot	Inferred -ni
(19)	Slot	Filler -Vw
(20)	**Slot**	**Boundary**
(21)	Slot	Counterfactual =tih
(22)	Slot	Emphatic Coordinator =nih
(23)	Zone	Evidentials =cud, =hɔ
(24)	Slot	Repetitive =bʼay
(25)	Slot	Emphasis =pog
(26)	Slot	Reportative =mah
(27)	Slot	Habitual bɨg, Distributive pɨd
(28)	Slot	Frustrative yæ̃́h
(29)	Slot	Contrast jʼám / jʼã́h, páh, tán

(30) ZONE INTENSIFIER MÚN, ADVERSATIVE CONJUNCTION KǍH, PERSISTIVE TÆ, EPISTEMIC MODALITY ʔŨH, ETC.

(31) ZONE NP{A,S,O,OBLIQUE}

(32) SLOT DECLARATIVE *-V́h*

Below, I discuss the positions in the template according to the principal categories of verbal morphology that were introduced in §2 above, and as schematized in (3). Position 6 is occupied by the verb stem, which as noted above may be composed of multiple serialized roots. The bolding of positions 6 and 20 reflects their obligatory status within (most) verb words. The labels used here should be understood primarily as heuristics reflecting language-specific distinctions, as the discussion below explores.

(3) *hup-*   *yəd*   *-cɨ́w*   *-ɨ́y*   *=cud*
     Preformative[3]- Stem -Inner.Suffix -Boundary.Suffix =Enclitic
     (3-5)-   (6)   -(7-19)   -(20)   =(21-26)
     REFL-   hide -COMPLETIVE -DYNAMIC[4]   =INFER.EVID
     yǽh
     Particle
     (27-30)
     frustrative
     'had already hid himself, apparently, in vain'

3.1.1 Positions 1-2: Nominal arguments

Nominal arguments in Hup are largely independent of the verbal complex, according to the constituency diagnostics discussed below. However, pronominal subjects are particularly likely to occur in pre-verbal position, where they are unstressed, and in one dialect of Hup (Upper Rio Tiquié) they are further phonologically reduced (from *tɨh* to *tV-*, with the vowel quality matching that of the following syllable). Subjects occurring post-verbally are subject to certain restrictions: in declarative clauses, they require the clause-level declarative suffix *-V́h* (which can function as a boundary suffix on clause-final verbs), whereas uninflected post-verbal pronominal subjects are a feature of polar interrogatives (example (3)). Simplex nominal O arguments can also intervene between a valence-adjusting proclitic and the verb, as addressed below. Otherwise, whether a nominal argument or other adjunct precedes or follows the verb appears to depend largely on information structure.

(4) a. *kɨt-d'ák-áy=mah*   **tɨh-ɨh**
        chop-be.against-DYNM=REP 3SG-DECL
        'She hit (her machete) against (the fishtrap).'

    b. *wǎed=yíʔ níh-ɨ̃y   níŋ-ǎn tíh ?!*
        food=ADV be.like-DYNM 2PL-OBJ 3SG
        'Is it just like food for you all?!'

3.1.2 Positions 3-5: Preformatives

Morphological elements that precede the verb stem are relatively few in Hup and are associated with valence-adjusting. These are the preformatives *ʔ̃h* 'interactional (reciprocal')' and *hup* 'reflexive/passive' (with an additional, more limited reciprocal function); and the prefix *hi-* 'factitive', which is relatively unproductive and semantically idiosyncratic, but tends to increase valency.[5] There is also a small set of verb roots that have been semi-grammaticalized within serial verb constructions as causativizers, of which the most productive element is *d'oʔ* 'take'.

Position 3 is identified as a zone, reflecting the fact that the interactional, reflexive, and causative preformatives may co-occur and variably order with one another. The relative ordering of these elements is primarily scope-dependent, as seen in (5–6). Factitive *hi-* almost always forms a tight unit with a single verb root – as can be seen in the fully lexicalized verb *hipãh* 'know' in 6; no semantically relevant form *pãh* exists. However, and unlike the suffixes and other post-stem verbal elements, these preformative + stem combinations can form scopally nested units, as seen in (6). In certain cases, such integrated or semi-lexicalized preformative + root combinations can occur in the midst of a serialized set of verb roots (within position 6), thus representing a minor exception to the template above; this is true of factitive *hi-* in particular (see (1) above).

(5) *hɨd **ʔ̃h-hup**-yɔ́d-ɔ́y*
    3PL INTRC-REFL-hide-DYNM
    'They are hiding from each other.'

(6) *wɔ̌h=n'ǎn(...)   **d'oʔ-[[hup-hipãh]-næn]**-ní-h*
    River.Person=PL.OBJ take-REFL-know-come-INFR-DECL
    'He brought the River People to be educated.' (lit., he caused them to come and have knowledge)

---

[5]The term *preformative* is used here as a generic term for grammatical elements that precede a root (i.e. prefixes and/or proclitics).

Another notable feature of the preformatives is that a simplex O argument can intervene between the verb and interactional *ʔũh* or reflexive *hup*, as seen in example (7). This O argument is understood as incorporated, since it cannot be inflected with any nominal morphology (whereas any potential incorporation elsewhere in Hup is obscured by the general preference for OV constituent order). However, the preformative in this context receives stress/tone and as such is more phonologically independent than it is in the immediately pre-verbal position, where it is always unstressed (see §5 below). There is also no indication in my corpus that a preformative followed by a simplex O can itself be preceded by a causative root (cf. example (6)). The separation of slots 3 and 4 for causative roots reflects this apparent co-occurrence constraint.

(7)   a.   *yã̌ʔambŏʔ=d'əh ʔũh-g'ɔ̌ç-ɔ́y*
            dog=PL   INTRC-bite-DYNM
            'The dogs are fighting.' (lit. 'biting each other')

      b.   *hɨd ʔũ̌h   nam   nɔ́ʔ-ɔ̃́y*
            3SG INTRC poison give-DYNM
            'They give poison to each other.'

3.1.3 Positions 7-19: Inner suffixes

A Hup verb can include from zero to multiple elements associated with the inner suffix category, which precede the obligatory boundary suffix. As seen in Table 1, these morphemes occupy a wide semantic range, encoding aspect, mood, negation, associated motion (the venitive *-ʔay*), and even one valence-related form (the applicative *-ʔũh*). They are for the most part templatically rather than scopally ordered, and are subject to relatively limited co-occurrence restrictions. The set of forms in slot 16 appear to be an exception, though this requires further testing. Note also that the elements in slot 16 are *fluid* formatives – i.e. they can occur *either* as members of the inner suffix or the enclitic categories, as discussed below (Sections 3.1.5 and 6). The stacking of multiple elements from positions 7-19 can be seen in example (8).

(8)   *yúp hɨd g'oʔwow'-tuʔ-y'æt-**yiʔ-pog-ʔé-w**-ăn-áh*
      that 3PL squeeze-dunk-leave-TEL-EMPH-PFV-FILL-OBJ-DECL
      '(He drank) that which they had squeezed, dunked and left.' (fish-poison vine in his drink)

The distinction between serialized verb roots (position 6) and inner suffixes in Hup is relatively non-discrete, in that there are virtually no definitive phonological or morphosyntactic cues to distinguish an element that occupies the end of a string in position 6 from one at the beginning of a string involving positions 7-19. This blurred distinction reflects the fact that many inner suffixes are quite obviously grammaticalized from verb roots, and these grammaticalization processes both facilitate and are facilitated by the lack of a clear-cut distinction between these two parts of the verb template.[6] This point is further addressed in §6 below.

3.1.4 Position 20: Boundary suffix

Verbs in nearly all predicative contexts in Hup require a boundary suffix. While apparent exceptions to this generalization appear in imperative and apprehensive moods, the bare verb stem in these contexts requires a specific tone assignment on the final syllable which may be analyzed as a boundary suffix. The most frequently encountered boundary suffixes are the principal markers of clause type: declarative *-V́h*, interrogative *-Vʔ*, dependent (subordinate) *-Vp* (and arguably imperative/apprehensive *-Ø* + tone). They, together with the dynamic aspectual suffix *-V́y* (which occurs primarily in declarative clauses, but also in some interrogatives), copy their vowel from the preceding syllable – or lose their vowel altogether when the preceding syllable ends in a vowel (see (6) above) – and are thus phonologically highly dependent on their hosts.[7] Other boundary suffixes mark TAM, various forms of subordination, etc. Boundary suffixes in Hup are lexically specified for stress/tone, and some also condition stress on the preceding verb root, such that every verbal predicate in Hup normally has either one or two syllables bearing primary stress (see §5). In most contexts, only a single member of the set of boundary suffixes can appear on a verb. Exceptions are mostly encounted in the context of the clause-level declarative marker *-V́h*, which can only occur clause-finally but is particularly promiscuous with respect to its possible hosts (as per position 32; see also example (4a) above). Several of the other clause-level boundary suffixes also may attach to clausal constituents other than verbs, and as such have properties that are often associated with clitics rather than affixes. The clausal negative *-nɨh* and the inchoative *-ay* are excep-

---

[6]See, for example, the form *cɨ̃p* 'complete' in (1) above, where it appears as a serialized verb root, but which is semantically and formally equivalent to the full form of the completive inner suffix *-cɨ̃p* of position 10 (which also has a phonologically reduced form *-cɨ̃w*).

[7]The dynamic suffix (*-V́y*), like several other suffixes in Hup, has an unspecified vowel slot, which is filled by a copy of the vowel (including its specification as nasal or oral) in the preceding syllable; see §5.

tional in that they can occur as either inner suffixes or boundary suffixes, with certain combinatory limitations.

3.1.5 Positions 21-26 and 27-30: enclitics and particles

The Hup verbal complex includes two robust categories of enclitics, the second of which are referred to as *particles* in light of their greater independence from the verb core. Many of these elements can also associate with nonverbal predicates, and a few can also occur with focused non-predicative constituents of a clause. The enclitics proper are normally encountered immediately after the boundary suffix and are unstressed, while the particles typically come later and receive stress/tone (as in example (3) above). A verbal construction may involve several of these elements. While their order is relatively fixed, it displays a certain sensitivity to scope, such that in certain cases particular scopal arrangements can override the expected templatic order. Such scope-determined variations are most evident when a large number of enclitics/particles co-occur; in (9), for example, the emphatic coordinator *=nih* actually follows the habitual particle *bɨg*.

(9)   *yɨ-dʼɔh-ăn*   *peʔ-níh=pog*   *bɨ́g=nih*   *jʼám*   *hǎʔ*
     DEM-PL-OBJ hurt-NEG=EMPH HAB=EMPH.CO DST.CNTR TAG
     'So (the insects) have never bothered those guys at all, huh?!'

3.2 Noun structure

Table 2 provides the structure of the noun phrase, focusing on elements that occur with nominal arguments of predicates. Many of the formatives that are understood here as primarily associated with the verb (see Table 1) can in fact associate with nominal (and other) predicates as well (and thus can be understood as transcategorial; see §4.3); these include some aspectual elements (e.g. perfective *-ʔeʔ*) and some evidentials. Nouns may also be associated with (and even phonologically host) still other formatives which are understood to occur at a clausal level; among others, these include the declarative suffix *-V́h* (example (4a) above). These formatives are not included in Table 2.

     In contrast to verbs, a noun phrase can consist minimally of a bare noun root. Various modifying elements can precede the noun: demonstratives, possessors, quantifiers, relative clauses, or other nouns, and further elements (classifiers and adjectives) can follow it; see (10). Most of these modifying elements can themselves head noun phrases (i.e. occupy zone 5), but usually require additional morphology to do so – typically either the plural marker *=dʼɔh* (primarily for

Table 2: Nominal planar structure in Hup

Position	Type	Elements
(1)	Slot	Demonstrative
(2)	Zone	NP{Poss}-nɨh
(3)	Slot	Quantifier, Numeral
(4)	Slot	Relative clause
(5)	Zone	**root (or compounded roots)**
(6)	Zone	Classifier
(7)	Zone	Adjective
(8)	Slot	'Respect' markers =wəd (m), =wa (f)
(9)	Slot	Deceased marker =cud
(10)	Slot	Augmentative =pog, diminutive =mæh
(11)	Slot	Plural =d'əh
(12)	Zone	Case -ăn, -an, -V́t
(13)	Slot	Intensifier =hup
(14)	Slot	Parallel marker =hin
(15)	Zone	Topic/Focus/Contrast/Evidentiality

elements in positions 1-4, preceding the root) or the dummy head *tɨh=* (for elements in positions 6-7, following the root). The stacking of some elements in the positions preceding or following the root (such as demonstratives and relative clauses, as in (11), or multiple adjectives, as in (12)) requires them to take these derivational elements. While the resulting constructions (especially those involving number, as in (11)) could be argued to involve agreement within the NP, they could alternatively be analyzed as involving multiple nominals in a compounding or appositional relationship. The function of classifiers in Hup is primarily derivational, as opposed to agreement-related, and classifiers have been shown to have developed quite transparently from nouns in compound constructions (Epps 2007, 2008a). (Compare, for example, *pɨhít=tat* [banana=FRUIT] 'banana (fruit)'; *hɔ̃=tat* [burn=FRUIT] 'light bulb'; cf. example (12).

(10) *yúp mɔy pŏg-an mah j'ám*
      DEM house big-DIR REP DST.PST
      'in that big house, it's said, long ago.'

(11) *cā-d'ɔ̆h   ʔɨd-hipāh-nɨ́h=d'ɔh   ni-bɨ́-h*
      other=PL speak-know-NEG=PL be-HAB-DECL
      'There are others who don't know how to speak (that language).'

(12) *núp=tat   tɨh=pɔ̆g   tɨh=pɑ̆y   nɔh-yɨ́ʔ-ɨ́y*
      this=fruit 3.SG=big 3SG=bad fall-TEL-DYNM
      'This big ugly fruit fell.'[8]

Grammatical formatives associated with nominal arguments include the *respect* and *deceased* markers, the augmentative and diminutive, number, case, and various other elements, many of which relate to topic, focus, and/or contrast. Many of these elements can associate with verbs as well as nouns, but tend to have somewhat non-analogous functions; for example, the augmentative *=pog* has an emphatic function in verbal predicates, and the case markers function as case-specified subordinators on verbs within headless relative clauses. Some functions are arguably still more distinct: e.g. the suffix *-Vp* marks topicality when it occurs on nouns, but on verbs it functions primarily as a marker of subordination (principally in relative and converbal clauses), and *=b'ay* is a topic-switch marker on nouns but indicates repetition of an event in verbal contexts. For some of these morphemes, of course, evidence of a diachronic relationship may not mean that they should be considered the same morpheme synchronically; however, deciding where to draw this line is often non-trivial (see Epps 2008a for discussion).

For nouns, some of the distinctions among the various sub-classes of formatives defined above for verbs do not apply. In particular, there is no inner/boundary suffix distinction, reflecting the fact that there is no obligatory position beyond the nominal root. While the morphemes occurring in positions 8-14 are labelled and segmented here as enclitics and suffixes – on a par with the corresponding categories in verbal contexts, where many of the same forms also occur – for nouns any distinction is entirely phonological (whereas for verbs it is also morphosyntactically relevant; see also §5 below): Enclitics are unstressed CVC morphemes that follow a noun, while suffixes are morphemes lacking an onset consonant (and thus violating the minimal prosodic word requirement of Hup); the latter set includes the case markers (position 12) and a few of the topic/focus markers (position 15), such as *-Vp* 'topic'. All of these -VC nominal suffix forms also occur as boundary suffixes on verbs, though sometimes with quite distinct functions, as noted above. As with verbs, some of these forms copy their vowel

---

[8]'Fruit' is a *bound* noun; i.e. it must be preceded by another nominal element.

from the preceding syllable (or delete it where the preceding syllable lacks a coda consonant), and they are lexically specified for stress/tone. However, in partial contrast to verbal contexts, the nominal suffixes tend to occur at the end of the entire noun phrase, following any post-nominal modifiers, including various clitics. Thus, any distinction between suffix and clitic for nouns is not particularly meaningful in Hup. The (unstressed) enclitic vs. (stressed) particle distinction is also only marginally relevant for Hup nouns, since most of the elements that may be identified as particles and can follow a noun have clause-level scope.

4 Morphosyntactic constituency

In this section, I explore the application of various constituency diagnostics relating to morphosyntactic domains in Hup, in light of the planar structures introduced above. Each of these tests represents a generalization over the constructions of the language that identifies a subspan in a planar structure. I consider the following diagnostics: free occurrence, non-permutability, ciscategorial selection, subspan repetition, and non-interruptability.

4.1 Free occurrence (v: 6-20, 2-30; n: 5-5, 5-15)

This variable relates to both the minimal and the maximal units that can occur as an independent utterance. For verbs, the minimal free occurrence domain spans positions 6-20, reflecting the obligatory presence of a root and a boundary suffix, as evidenced in utterances like (13) – a very frequent response to any inquiry concerning the presence or existence of a person or thing. For nouns, this domain is represented by a single root (position 5), which may form a complete utterance in contexts such as identifying or presenting someone with an object.

(13)   *ní-íy*
        v:6-20
        be-DYNM
        '(X is) present/exists.'

The maximal free occurrence domain spans the largest number of positions that can occur together as a single free unit (with the caveat that *single* here is necessarily defined according to other constituency diagnostics). For Hup verbs, this covers positions 2-30, and includes preformatives, roots, inner suffixes, the boundary suffix, enclitics, and particles, according to the language-specific categories defined above. While examples indicating the full span (2-30) within a

single construction have not been identified, example (14) illustrates a span covering positions 6-30.

(14) *tíh-ɨp*   *húp*   *ham-yíʔ-ay*   *=mah kǎh*
      v:--   -   6-7-20   =26   30
      3SG-DEP person go-TEL-INCH =REP ADVR
      'But as for him, the man, (he) got away.'

For nouns, the minimal domain is simply 5-5, as in *hup* '(it's a) person'. The maximal nominal domain spans positions 5 (root) through 15 (a zone relating to topic, focus, and contrast). Example (15) shows a relatively complex noun with elements filling multiple positions, while 16 illustrates a full span from positions 5-15. In both of these examples, the possessor plus inalienably possessed noun can be understood as a compound construction (occupying position 5), whereas an alienable possessor (with possessive morphology) is more separable from the root and appears in position 2.

(15) *ʔɨn =pǎ̧ç*   *=wəd =cud   peʔ-ní-h*
      n:5 =5   =8   =9   -
      1PL =father's.brother =RESP =DCSD sick-INFR-DECL
      'Our late uncle was sick.'

(16) *yɨ-níh-mɨ̌ʔ*   *j'ám*   *ʔɨn =b'ay, ʔɨn =tæ̌h*   *=n'ǎn*[9]   *=hin*
      n:---   -   -   -   5   =5   =11+12   =14
      that.ITG-be.like-SIM DST.CNTR 1PL =again 1PL =offspring =PL.OBJ= also
      =b'ay, "nɨŋ b'oy-ʔáy   hám!" nɔ-nɨh   ʔɨn ni-bɨ-há̧ʔ
      =15   -   -   -   -
      =again 2PL study-VEN go.IMP say-NEG 1PL be-HAB-TAG
      'Even so, we don't tell our kids "go to school!"'

4.2 Non-permutability (v: 2-10, 7-10; n: 8-11, 1-15)

The non-permutability diagnostic makes reference to spans where elements must occur in a fixed order. This order may be either templatically defined or determined by scope.

In Hup verbs, scopally defined non-permutability holds across positions 2-10, according to the properties of these elements as set out in §3.1 above. Rigid (non-scopally defined) non-permutability, on the other hand, applies only across a set

---

[9]The element *n'ǎn* is a fused morpheme composed of plural *d'əh* + object *-ǎn.*

of the inner suffixes following the verb base, from positions 7-10. The verb base itself is excluded because serialized root combinations may be nested and may include preformatives, as in example (17); see also (1) above. Accordingly, there seems to be no rigid (templatic) non-permutability in the span that overlaps the verb core (position 6).

(17)   *mɔ̆h*   *tɨh*   *yæ̃ʔ-wæd-[hi-wág]-áh*
       v:-   -   6-6-5-6-20
       tinamou 3sg roast-eat-FACT-day-DECL
       'He cooked and ate tinamou birds until daybreak.'

After slot 10, we find formatives that can occur in variable order by appearing either as inner suffixes or as enclitics (see §4.5 below), such as emphasis *pog* (positions 11 and 25; compare examples (18) and (19)) and evidentials (positions 16 and 23).

(18)   *yúp baʔtĭb'* *g'ɔ̃h-pog-ʔé-ew-ăn*   *hɨd wæd-yiʔ kəd-hám-ắy=mah*
       v:- -   6-11-13-19-20   - --   ----
       that spirit be₂-EMPH-PFV-FLR-OBJ 3PL eat-TEL pass-go-DYNM=REP
       'Then that spirit that she really had become, they ate (her) up.'

(19)   *yɨ-d'ɔ̃h-ăn* *peʔ-nɨ́h=pog*   *bɨ́g=nih*   *j'ám*   *hə̆ʔ*
       v:---   6-14=25   27=30   30   -
       DEM-PL-OBJ hurt-NEG=EMPH HAB=EMPH.CO DST.CNTR TAG
       'And (the insects) have never bothered those guys at all, huh?!'

Rigid non-permutability also appears to hold further out in the verbal planar structure (likewise in spans that do not include the core). These spans are 18-20 (inferred evidential *-ni*, filler suffix *-Vw*, and boundary suffix; see example (20)); and probably also the span represented by positions 29 -30.

(20)   *pɨ̆ŋ*   *deh=nɔ́*   *pótʔah...* *wəhə́d=d'əh*
       v: -   --   --
       tree.grape water=mouth above   old.man=PL
       j'ɔm-b'eh-ʔeʔ-ní-p
       6-6-13-18-30
       swim-cross.water-PFV-INFR-DEP
       'Above the mouth of Cucura Igarapé... the Ancestors swam across.'

For nouns, rigid permutability arguably holds across positions 8-11, although this observation bears further testing as some of these formatives rarely if ever co-occur. The domain of scopal permutability holds across the entire set of positions represented in Table 2.

4.3 Ciscategorial selection (v: 3-10, 2-18; n: 5-6, 1-13/14)

As discussed above (see §3.2 in particular), the distinction between nominal and verbal constructions in Hup is not very clear-cut. Nouns and verbs can take many of the same morphological elements, although assessing "sameness" is often complicated by the fact that some elements have developed different functions (and sometimes only subtly so) in these distinct contexts, despite being formally identical and (often) obviously historically related. In addition, nominal predicates can associate with still other formatives that otherwise are found primarily with verbs, as well as clause-level elements (such as the declarative marker *-V́h*, which also occurs as a boundary suffix on clause-final verbs); in some cases, these phenomena can be attributed historically to the extension of morphology to nominal predicates following its emergence through grammaticalization in verbal contexts. A further diachronic observation involves the reanalysis of some nominal constructions as verbal, which explains why they still retain certain features associated with noun phrases – e.g. an instrument nominalization ('thing for doing V') is the probable source of a purpose adverbial and thence a future construction, with idiosyncratic constraints on co-occurring verbal morphology, particularly negation (Epps 2008b). Finally, a subset of nouns relating to periods of time or human lives (e.g. 'day', 'night', 'child', 'old man') behave effectively as though they are intermediate between nouns and verbs; for example, *wăg* 'day' can head noun phrases without derivation (e.g. *kaʔap wăg* 'two days'), but can also head some verbal predicates and take morphology that otherwise does not occur with nouns, e.g. *wag-yɨʔ-cɨ́w-ɨ́y* (day-TEL-COMPL-DYNM) '(it is) already / has become day' (see also example (17) above).

Despite these complications, we can make a distinction between ciscategorial and transcategorial elements, here counting those morphemes that have a highly divergent function in nominal vs. verbal contexts as ciscategorial. For verbs, the minimal span (overlapping the verb base) – i.e. which contains positions that can only have verb-ciscategorial elements – spans positions 3-10 in the planar structure. Position 2 is excluded in light of the fact that pronominal elements can also occur with nouns as inalienable possessors, while position 11 is excluded because the emphatic form *pog* can appear with nouns (example (21)) as well as with verbs (see (18–19) above), with little or no difference in meaning. A maximal

span extends between positions 2 and 18 (i.e., all morphemes in the positions outside of this span are transcategorial), since the inferred evidential *-ni* (position 18) can only occur with verbs (see (20) above for an example), whereas some of the intervening elements (e.g. distributive *pɨd*, evidential *mah*) can occur with nominal arguments as well as nominal predicates.

(21) *húp=**pog***   *ʔúh*   *tɨh=ʔĩh* !
      person=EMPH EPIST 3SG=M
      'Could that be a person?!'

For nouns, the minimal span of clearly noun-ciscategorial elements is limited to positions 5-6 (root and classifier), with the understanding that when classifiers attach to verb roots they necessarily derive a nominal construction. In limited cases, relative clause constructions (position 4) can occur as main clauses (through an insubordination process), and most adjectives (position 7) can also function as adverbs. The maximal span covers positions 1 (demonstratives) to either 13 ('intensifier' *=hup*) or 14 ('parallel' marker *=hin* ['also']), given that *=hin* can associate with adverbial elements, though not with verbs. Several of the intermediate elements – most notably the oblique case marker *-V́t* and the plural marker *=d'əh* can occur with verbs to form certain types of adverbial clauses.

4.4 Subspan repetition (V: <6>, 1-32; n: 1-15, 1-15)

This diagnostic relates to "a well-defined contiguous subspan of positions that occurs more than once for a given construction" (Tallman 2021: 337), as indicated by elision in contexts of subordination or coordination, and by evidence of scope over a repeated series of subspans. For Hup verbs, subspan repetition applies at several levels.

Verb serialization and clausal subordination/coordination constructions offer domains in which to consider subspan repetition in the verb. Serialization in Hup involves the combination of verb roots within position 6. The sequence of serialized verb roots is included within a single tone/stress domain (see §5.5 below), and as a unit takes a single boundary suffix. The boundary suffix and any inner suffixes or enclitics/particles that follow position 6 scope over the entire serialized unit, as can be seen in the case of negation in (22). This scopal behavior distinguishes a serial verb construction from subordinated or coordinated clauses, in which the verbs are inflected independently, with affixes scoping only over their host root(s) (example (23)).

(22) *nu-có?*   *híd-ăn*   *tɨh*   *[ye-yæ̌h]-nɨ́h*
      --   --   -   6-16-20
      this-LOC 3PL-OBJ 3SG enter-request-NEG
      'He forbids them to come in here.'

(23) *tɨnɨ̌h*   *?ɨ́d*   *[wɨ?-níh]*   *[g'et-g'o?-tú-ay=d'əh=nih]*
      -      -   6-20   6-6-6-17-20-22
      3SG.POSS speech hear-NEG stand-go.about-want-INCH=PL=EMPH.CO
      'And we'd go about without understanding her language.'

For serial verbs, *post*posed affixes scope over the entirety of position 6, as example (22) shows; however, *pre*-posed affixes (preformatives, specifically those relating to positions 3 and 5) scope instead over individual serialized verb roots. This can be seen in 24, where the reflexive preformative *hup-* scopes over *hi-cu?* 'cover', which itself is composed of the factitive prefix *hi-* and the verb *cu?* 'grab' (see also example (17) above). In light of this scopal behavior of preformatives, then, the minimal domain of subspan repetition for the Hup verb is best understood as the single root, which itself can be a component *within* position 6 (here represented as <6>).

(24) *[hup-[hi-cu?]]-ham-tú?-ay-áh*
      3-5-6-6-6-17-20
      REFL-FACT-cover-go-immerse-INCH-DECL
      '(The crab) went and covered himself up in the water (to hide).'

One other context that may relate marginally to subspan repetition involves clauses linked via the etyma *-yó?* 'simultaneous' or *-mɨ?* 'sequential' (both of which are boundary suffixes). The 'simultaneous' construction normally involves the same subject across the two clauses, which is usually (though not obligatorily) elided, as in (25); the 'sequential' construction almost always involves different subjects (26). However, this same/different subject pattern allows exceptions; moreover, as far as elision is concerned, arguments in general may be freely elided when understood from the discourse. The same is generally true for evidentials and other elements in positions 21–30, following the boundary suffix, which may also be dropped when already activated within the discourse context. Thus these processes of elision are common in contexts of clause combination, but are not exclusive to them.

(25) *"hə̆ʔ", nɔ-yóʔ, tíh-ăn tɨh yók-ay-áh*

    -     6-20   --   1   6-17-20

    OK   say-SEQ SG-OBJ 3SG poke-INCH-DECL

    'Having said "all right", he poked him.'

(26) *j'ɔ́m-ɔ̃p   tɨh kəd-d'ŏb-mɨ̆ʔ=mah,   d'ŭç hɨd*

    --   1   6-6-20-26   -   1

    bathe-DEP 3SG pass-go.to.river-SIM=REP timbó 3PL

    tətəd-d'óʔ-óy=mah

    6-6-20-26

    beat.timbó-take-DYNM=REP

    'While she (their mother) went down to bathe, they beat the timbó (to release the poison), it's said.'

For verbs, a maximal subspan repetition domain – i.e. the largest set of positions that clause combination may target – is represented by the entire planar structure. The same is true for nouns, in which there appears to be no substantive difference between the minimal and maximal subspan domains.

4.5 Non-interruptability (v: 3-10, 2-30; n: 5-15)

The diagnostic of non-interruptability is particularly interesting for verbs in Hup, especially in light of the key role attributed to this diagnostic in prior work in morphological theory (Booij 2009, Bauer 2017: 17). Non-interruptability also relates to a related consideration, extended exponence – i.e. the deviation from biuniqueness that is often associated with morphological relations. For nouns, there seems to be little to say regarding this diagnostic, which identifies a span between positions 5 and 15. For verbs, we can identify the span between the preformatives in position 3 and the modal elements in position 30 as a domain in which a complex free form (e.g. a multi-word noun phrase) cannot intervene. However, the non-interruptability test for verbs otherwise breaks down for a number of positions and etyma.

One exception to non-interruptability relates to the valence-related preformatives in position 3 (interactional *ʔũh* and reflexive *hup*). As discussed above (§3.1.2 and example (7)), a simplex O argument can intervene between the preformative and the rest of the verb. While this O argument may be best understood as incorporated, insertion of a nominal argument within this verbal span is otherwise not attested (with the marginal exception of the 'verby' nouns mentioned in §4.3 above, which can occur as serialized roots within position 6). Moreover, the insertion of the O argument in these constructions has additional phonological

outcomes that also challenge our understanding of this span as a single word, namely the assignment of independent stress/tone to the interrupted preformative (see §5 below).

The non-interruptability test is also challenged by the set of *fluid formatives* – morphemes that may appear in more than one place within the verbal construction, as introduced in §3.1.3 above (a distinct property from that of transcategoriality). This set consists of emphasis, habitual, distributive, repetitive, and frustrative markers, and evidentials, and these etyma can occur variably as either inner suffixes, preceding the boundary suffix (positions 11, 15, and 16), or as enclitics, occurring later in the verb (positions 23 -18). Because these etyma can occur at more than one point in the verbal construction, they make the interruptibility tests ambiguous.

As shown for repetitive *b'ay* and frustrative *yæ̃h* in (27–29), these relatively bonded but positionally variable interrupting elements cut up the verb complex into distinct layers. As these examples also illustrate, the position of the fluid formative is sensitive to the type of boundary suffix present – the fluid etymon necessarily occurs in the inner suffix position when the boundary suffix is the (obligatorily) clause-final declarative form *-V́h* ((27a) and (28a)), but as an enclitic/particle in the context of other boundary suffixes ((27b) and (28b)). This situation may be compared to what happens with non-fluid formatives: For an etymon (e.g. venitive/associated motion *-ʔáy* 'go, do X, and return') that is always an inner suffix, it occurs in this position regardless of the type of boundary suffix that is taken by the verb. For one that is always an enclitic/particle (e.g. the distant past contrast marker *j'ám* in example (29)), it necessarily follows the boundary suffix (which cannot be declarative *-V́h* but can itself host the declarative marker).[10]

(27)   a.   *yúp=mah tɨh hí-**b'ay**-áh*
            v:--   -   6-16-20
            that=REP 3SG descend-again-DECL
            'Then he came down again.' (inside verb core; inner suffix status)

       b.   *yúp=ʔã́y-ăn   ʔãh b'uy-d'əh-yíʔ-ɨp=**b'ay***
            v:---   -   6-6-7-20-24
            DEM=woman-OBJ 1SG throw-send-TEL-DEP=again
            'I got rid of that woman, too.' (outside verb core; enclitic status)

---

[10]Evidence from both comparative and internal reconstruction indicates that these fluid morphemes began as verb roots in serial constructions, and developed their less bonded enclitic/particle instantiations subsequently (see §6 below and Epps 2008a for discussion).

(28) a. *núw-ăn ʔãh tuk-**yǽh**-ǽh*
      v:--   -   6-16-20
      this-OBJ 1SG want-FRUST-DECL
      'I'd like this one (but I don't expect to get it).' (inside verb core; inner suffix status)

    b. *núw-ăn ʔãh túk-úy*   **yǽh**
      v:--   -   6-20   28
      this-OBJ 1SG want-DYNM FRUST
      'I'd like this one (but I don't expect to get it).' (outside verb core; particle status)

(29) *nutǎn-ǽy=d'əh-ɔ́h,   nɨh-níh-ay   **j'ám**-ǽh,   nutǎn-ǽh*
    v:----   6-14-20   29-32   --
    today-DYNM=PL-DECL be.like-NEG-INCH DST.CNTR-DECL today-DECL
    'People of today, they don't do like this anymore, these days.'

While flexible assignment and interruptability are features that are generally considered less typical of bonded morphology, the fluid etyma actually exhibit a greater deviation from biuniqueness than is typical of most formatives in Hup, and in this sense appear more *morphological*. This deviation is evident in that one meaning-form combination is associated with multiple slots in the template; moreover, for several of these elements, the -CVC inner suffix + -*V́h* boundary suffix has an optional and/or contextually determined -CV-h variant (as in 30c; compare 30a-b). This process of phonological reduction corresponds quite closely to degree of grammaticalization, and is encountered more generally in Hup among other CV(C) root +VC suffix combinations (see §5 below).

(30) a. *ʔãh hám-ǽy **bɨ́g***
      1SG go-DYNM HAB
      'I go regularly.'

    b. *ʔãh ham-**bɨ́g**-ɨ́h*
      1SG go-HAB-DECL
      'I went regularly.' (more emphatic)

    c. *ʔãh ham-**bɨ́**-h*
      1SG go-HAB-DECL
      'I went regularly.' (more neutral)

5 Phonological constituency

Criteria associated with the prosodic word in Hup are notably non-convergent (cf. Schiering et al. 2010). As observed above, different phonological criteria are associated with different morphological units in Hup, several of which are reasonable candidates for an orthographic word. These criteria are associated with a concentric series of domains, ranging from one to multiple morphemes. This section begins with a general overview of the principal domains, and then focuses one by one on a set of particular phonological diagnostics as they apply across these domains for nouns and verbs: segmental constraints, vowel copying, final consonant deletion, and stress/tone locus.

5.1 Overview: Concentric phonological domains

Concentric domains involve particular quantitative and qualitative relationships among morphemes, syllables, segments, and stress/tone loci. These domains consist of at least five levels:

a) The canonical (and minimal, as noted below) monomorphemic prosodic word is a single CVC syllable with one stress/tone locus; e.g. *mǒy* 'house'.

b) Monomorphemic words with two syllables are much less frequent and normally take the form CV_1CV_1C (or, more rarely, $CV_1CV_1V_1$); e.g. *mɔhǒy* 'deer'. In the vast majority of such forms, the intermediate consonant is restricted to a glottal or glide, and the vowels are identical; i.e. the segmental melody tier permits just a single vowel that multiply associates when there is more than one V slot in the skeleton. Again, there is only one stress/tone locus, which occurs almost without exception on the second syllable.

c) Reduplicative words, which are morphologically complex but only marginally so, occupy the next level. These consist of a $C_1V_1[C]C_1V_1C$ structure, as in *bəbǎg* [bəʔˈbǎgⁿ] 'cubiu fruit'. In these forms, the vowel is necessarily identical between the two syllables (as in monomorphemic bisyllabic words), while the primary intermediate consonant is identical to the onset but otherwise effectively unconstrained, and an underspecified C slot forms the coda of the second syllable. Words of this kind also have a single stress/tone locus on the second syllable.

d) The next level involves units composed of two distinct morphemes, of which the second is a vowel-initial suffix. Those suffixes that copy their

vowel from the preceding syllable represent a particularly close approximation of the canonical word form (as evident in levels a-c discussed above), while other -VC suffixes are specified for a particular vowel quality. The coda consonant of the root copies to the onset of the second syllable: $CV_1(C_1)[C_1]V_{1/2}C$, e.g. *wǽd-V́y* [wǽd^n-'dǽy] 'eating'. CV roots with -VC suffixes simply appear as CVVC. Bimorphemic combinations of this type can have either one or two stress/tone loci, which are lexically conditioned by the suffix and may fall on either or both syllables (see above).

e) For units with two or more morphemes composed of syllables with onsets (which normally also have codas), there are no particular constraints on the quality of either the consonants or the vowels involved: $CV_1(C_1)C_2V_2(V/C)$, e.g. *wæd-tég* [wæd^n-'tég^ŋ] (eat-FUT) 'will eat', *bɨʔ-wæd-tég* [bɨʔ-wæd^n-'tég^ŋ] (work-eat-FUT) 'will prepare food'. Such multimorphic strings may include one to two vowel-initial suffixes; in certain contexts involving grammaticalization, combinations of CVC + VC formatives are reduced to CV + C, resulting in a new form of the canonical CVC structure. The complex combinations described here can be identified as prosodic units on the basis of stress/tone – like the strings in (d), they have maximally one to two primary stress/tone loci, which normally occur on the boundary suffix and/or on the syllable that precedes it (with one or two exceptional patterns, which are also lexically determined by the boundary suffix).

5.2 Segmental constraints: Consonant and vowel quality

As observed above, the minimal free form in Hup is a syllable with an onset and two morae, of which all possible targets for nasality must be either uniformly nasal or oral. The canonical morpheme is CVC, but a few etyma are CV and as free forms surface as CVV, with prosodically motivated vowel lengthening. For bisyllabic and reduplicated morphemes, constraints limit the quality of the intermediate consonant(s) and require identical vowels, although a few exceptions exist. With respect to the planar structures, the minimal and maximal domains of these basic segmental constraints applies to a single, simplex root, occupying position 6 in the verb structure and position 5 of the noun structure.

5.3 Vowel copying (v: 6-20, 2-32; n: 5-5, 5-12)

A parameter related to the constraints on vowel quality is seen in the morphophonological process of vowel copying, which occurs across morpheme boundaries. Out of all Hup morphemes, only a few bonded formatives lack onsets (-VC), and

a subset of these copy their vowel from the preceding syllable (as in the dynamic suffix in example 30). Hup's -VC formatives are exclusively boundary suffixes (including declarative -*V́h*), with the exception of the "filler" syllable -*Vw* and the suffix -*ay* 'inchoative' – and these two are non-canonical as inner suffixes in that -*Vw* is a semantically empty element that must directly precede a boundary suffix, and -*ay* may also occur as a boundary suffix. Vowel copying is marginally licensed in only one other context, that of the procliticized third person pronoun, which occupies position 2 in the verbal structure (but is limited mainly to one dialect of Hup).

In the noun phrase, as discussed above, phonologically bonded -VC suffixes (such as case markers) normally occur toward the end of the noun phrase, and may therefore be hosted by adjectives and other elements following the nominal root, and can also occur with nouns as an inalienable possessor or 'dummy' head for an obligatorily bound noun (see §3.2). In a few lexical items – principally the words for 'man' and 'woman' – the 'dummy' third person pronoun has undergone vowel harmony, but this is specific to these contexts and has to do with lexicalization processes. Example (31) shows lexicalized vowel copying in these words, plus the obligatory copying in the 'oblique' -*V́t* suffix.

(31) tɨyĭʔ nawyɨ́ʔɨ́y tã̃ʔã̃yã̃t
     tɨh-yĭʔ   naw-yɨ́ʔ-*V́y*   tɨh-ʔã̃y-*V́t*
     3SG-man good-TEL-DYNM 3SG-woman-OBL
     'The man got well / became fully good in the company of the woman.'

Because vowel copying can apply clause-finally when declarative -*V́h* is present, its maximal domain in the verbal construction applies from positions 2-32; that is, outside of this maximal domain no morpheme is known to undergo vowel copying. Its minimal domain spans positions 6-20; within this domain, all morphemes that satisfy the structural requirement undergo vowel copying (in practice, it is the morphemes in positions 19 and 20 that copy the vowel of whatever morpheme in position 6-18 that directly precedes them). For nouns, vowel copying may apply maximally to positions 5-12, and minimally within position 5.

5.4 Final consonant deletion (6-20)

In Hup verbs, most inner suffixes are of the form CVC. However, a subset of these undergo coda deletion when followed by a vowel-initial boundary suffix; in this context, the boundary suffix itself loses its vowel, resulting in a -CV-C form that approximates the canonical monomorphemic form in Hup, and may reflect a maximality requirement that prefers that a stem be monosyllabic.

This phonological reduction reflects a grammaticalization process: In many cases, inner suffixes display both more and less grammaticalized variants, with only the former exhibiting coda deletion. For example, the inner suffix *-teg* encodes both purpose (the historically older function) and future tense (the more recently grammaticalized function, see Epps 2008b); this suffix is almost always realized as *-te* in the context of a vowel-initial boundary suffix when it encodes future, but as *-teg* when encoding purpose (examples (32)-(33)). However, in slow, careful and/or emphatic speech, the future suffix may also be realized with the coda consonant, while the purpose reading is occasionally found without the coda consonant in fast, casual speech. The final consonant deletion process is only relevant for Hup verbs, and spans positions 6-20.

(32) *hɨd ʔŭh   kəwəg wõt-**té**-ay-áh*
      3PL INTRC eye   pull.out-FUT-INCH-DECL
      -      -   -   6-15-17-20
      'One is going to pull out the other's eyes.'

(33) *núp=yɨʔ ʔɨn ni-n'ɨh-**tég**-éh*
      this=ADV 1PL be-NMLZ-PURP-DECL
      –   -   6-20-15-20
      'This is where/how we are supposed to live.'

5.5 Stress/tone loci (v: 2-26, 3-26; n: 5-15)

In verbs, as noted above, multimorphemic strings maximally take between one and two primary stress/tone loci. These normally occur on the boundary suffix and/or on the preceding syllable, as seen in example (33) and many others above. (There are two exceptions to this generalization: the "filler" suffix *-Vw* and the 'inchoative' suffix *-ay*, which never receive stress/tone, as in 32 above; note that these suffixes are also exceptional in other respects; see 5.3.) The domain of stess/tone loci normally spans positions 2 through 26; elements following position 26 are phonologically more independent in that they receive independent stress/tone (thus the label 'particle' to distinguish them); example (34) and many others above illustrate. However, a simplex O argument that intervenes between the valence-adjusting preformatives (interactional and reflexive/passive) in position 3 is unstressed, while the preformative in this context receives independent stress/tone (see (7b) above). Thus the minimal stress/tone domain is assessed as spanning positions 3-26, the maximal as 2-26.

(34) *ye-tã̌ʔ-ã̌y*   *yã̌h*
      6-12-29   28
      enter-CNTR.FACT-DYNM FRUST
      '(It) almost went in!'

For nouns, the principal stress/tone domain spans positions 5-15, but the identity of elements as roots, adjectives, classifiers, or particular suffixes determines which syllables will attract stress. As with verbs, stress in the nominal construction is normally culminative, but certain suffixes are lexically marked to receive an additional stress.

In considering how phonological units relate to the morphosyntax, we can observe that the stress pattern in verbs is sensitive to the obligatory inflectional position (the boundary suffix). In nouns, stress makes reference to the noun phrase, such that combinations of demonstrative-noun, noun-adjective, etc. receive one primary stress, just as they behave as a unit for the purposes of case marking and other morphological processes. Different diagnostics (particularly relating to the minimal free form vs. stress/tone loci) thus yield conflicting results in defining the word in Hup, a point I return to in §7 below.

6 Diachrony

As the constituency tests in §4–5 indicate, there are relatively few domains in which different diagnostics converge in Hup, raising challenges for a clear definition of the word in this language. However, as this section briefly explores, some insights into why these mismatches exist can be gleaned from evidence of the historical processes that have shaped Hup's morphological structure. As van der Tuuk (1971 [1864]: xliii) put it, "every language is more or less a ruin" – and as such, it is not clear why we should expect a heterogeneous set of diachronic process to necessarily converge on a consistent set of outcomes (see e.g. Nichols 2008: 287-288, Cristofaro 2019, and Schmidtke-Bode & Grossman 2019 for further discussion of this question).

The failure of the non-interruptability test as a robust constituency diagnostic in Hup is one area in which diachrony can shed some light. As explored in Epps (2008a, 2010), both of the valence-related preformatives are relatively transparently grammaticalized from nouns; 'sibling' for interactional *ʔũh*, and 'person' for reflexive/passive *hup*. Their development into verbal preformatives would have involved an incorporation process, possibly via a simple reanalysis of a preverbal O argument (already the canonical order in Hup for independent clausal

arguments) as part of the verb. Since the incorporated O that intervenes between the preformative and the verb root would have undergone effectively the same set of processes, we can suppose that the {preformative + O + verb} structure is retained from an earlier stage in which the erstwhile nominal arguments were indeed independent from the verb.

Diachrony may also help us to understand the status of the 'fluid' formatives, another area in which the non-interruptability diagnostic breaks down. As noted above, serialized verb roots in Hup are a productive historical source of new inner suffixes, through processes of grammaticalization (see Epps & Ananthanarayan 2022 for further discussion). Many of the 'fluid' etyma can be traced to verb roots; e.g. *yæ̃h* (frustrative) is also a verb meaning 'order, send'; *b'ay* (repeated event) as a verb means 'return'; *hɔ̃h* (nonvisual evidential) as a verb means 'produce noise/sound'; etc. The inner suffix position in the Hup verbal template may be seen as both an outcome of and a catalyst for this grammaticalization trajectory, in light of the formal ambiguity between serialized verb roots and inner suffixes in Hup. For the etyma in the 'fluid' category, however, a widening of scope from the verb to the predicate and even the clause would have facilitated the extension of these elements to non-verbal predicates – particularly nominal predicates, where the lack of any significant distinction between inner suffixes, boundary suffixes, and enclitics would have led to a reanalysis of the new morpheme as equivalent to the enclitics that appear on verbal predicates (33). This in turn arguably facilitated the *re*-extension of this etymon back to verbal predicates as an enclitic, a position that is consistent with its new scope. Similar examples of scope-driven reorganization of morphemes can be seen in other languages (see e.g. Mithun 2000); however, in Hup both options remained, with the retention of the earlier arrangement motivated by the ambiguous identity of the *-V́h* suffix as both a verbal boundary suffix and as an obligatorily clause-final element. Thus, while synchronically the two instantiations of the 'fluid' formatives may be identified as the same morpheme (i.e. as allomorphs) in light of their formal and semantic resemblance, diachronically they represent two distinct stages of grammaticalization.

(35) *pæ̌j=**hɔ̃***
      umari=NONVIS
      'It's umari fruit.' (identifying a smashed mess by the smell)

A comparative approach provides further insights into the historical developments that gave rise to Hup's morphological structure. If we compare Hup to

its sister-language Dâw, we find a similar structure, but with several key differences. As example (34) illustrates,[11] the Dâw verb resembles the Hup verb in that it involves multiple serialized roots, followed by grammatical formatives having scope over the preceding elements. Also like Hup, the canonical morpheme (and minimal prosodic word) structure in Dâw is CVC while a small set of suffixes are -VC (e.g. negation). However, in Dâw each element in the complex verbal construction is phonologically independent, in that it receives its own stress and/or tone value, whereas in Hup the entire complex is within a single stress/tone domain. Furthermore, there is no equivalent to Hup's boundary suffix in Dâw; accordingly, any verb root can stand alone as a minimal free form, and the formatives that follow the root are not demarcated into ordered categories with particular phonological or morphosyntactic behaviors (cf. the inner suffixes, boundary suffix, and enclitics/particles in Hup). The Dâw constructions in (34) can be compared to their (constructed) Hup counterparts in (35) (in which the boundary suffixes correspond to position 20).

(36) Dâw
*ʔabɨg tɨm pôj **ʃět   dŏʔ wɨ̂d**, ʔabɨg tih **ʃět   jũt-ẽh***
thus   eye big carry take FRUST thus 3SG carry PFV-NEG
'So Big-Eyes tried to carry his basket, but he did not (succeed in) carrying it.'

(37) Hᴜᴘ
cet-d'oʔ-yɛ̌h-ɛ̌h   … cet-ham-níh
6-6-16-20   6-6-20
carry-take-FRUST-DECL carry-go-NEG
'carr(ied), in vain.' 'did not go carrying it.'

As this comparison illustrates, the verb structures in both Hup and Dâw might be described as relatively isolating or as morphologically complex, depending on which constituency diagnostics are prioritized, and furthermore on whether the domains relating to particular diagnostics are understood as word-level or rather phrase-level. If we consider the diagnostics of culminative/obligatory stress/tone and minimal free occurrence, these identify a larger, multimorphemic constituent

---

[11]The Dâw examples are transcribed in IPA. Syllables may take rising (v̌) or falling (v̂) tone, or no tone. Nasalization in Dâw is a segmental feature, not morpheme-level as in Hup. The Dâw data come from original work with speakers in Waruá community (2013, 2017); see Epps et al. (2013+). See also Martins (2004) for a description of Dâw.

in Hup but a single-morpheme constituent in Dâw. On the other hand, if we consider the diagnostics of non-interruptability by a complex free form and scopal relations, these identify a multimorphemic constituent in both languages.

At an earlier stage of Hup's development, it is likely that its morphological structure closely resembled that of contemporary Dâw. It is also probable that contact with Tukanoan languages was a key factor in directing the particular changes that led to Hup's current profile. Tukanoan-driven restructuring of Hup grammar has been wide-ranging (see e.g. Epps 2007, 2008a,c, *inter alia*), and the order and identity of elements within contemporary Hup phrase and clause structure closely mirror those seen in Tukanoan languages. The structure of the Kotiria (Wanano) finite verb provides an instructive example (Stenzel 2013: 244-245). As the template in Figure 3 illustrates, a Kotiria verb consists of a primary root (position 1), optionally followed by a series of noninitial roots (2-4), to make up the 'lexical stem' of the verb. This unit may itself be followed by nonroot stem morphemes (5-6), which together with positions 1-4 make up the full verbal stem. This unit is obligatorily inflected by one of a set of markers associated with clause modality (evidential, directive, irrealis, and interrogative). The entire verbal complex forms a single phonological unit in relation to tonal spread. The parallels with the Hup verbal structure are obvious: a Kotiria root must minimally be inflected by a suffix relating to clause modality, like the boundary suffix in Hup; elements that intervene between the initial root and this suffix include verb roots and nonroot morphemes, with a blurred distinction between these, like the serialized verb roots and inner suffixes of Hup; and finally, the entire unit in both Kotiria and Hup represents a phonological unit as defined by tone and/or stress.

In sum, the diachronic pathway to relative polysynthesis in Hup has involved several components. These include the development of a culminative/obligatory stress domain that overlaps with that defined by the minimal free form (root + boundary suffix), and the development of the boundary suffix as an obligatory verbal element, leading to a minimal free form in verbs that spans more than one morpheme. On the other hand, the fact that constituency in the nominal domain has developed differently (in particular, with no correlate to a boundary suffix), and the propensity of Hup morphology to associate with both verbal and non-verbal predicates (and even arguments), have facilitated a relatively low degree of ciscategoriality and various violations of non-interruptability.
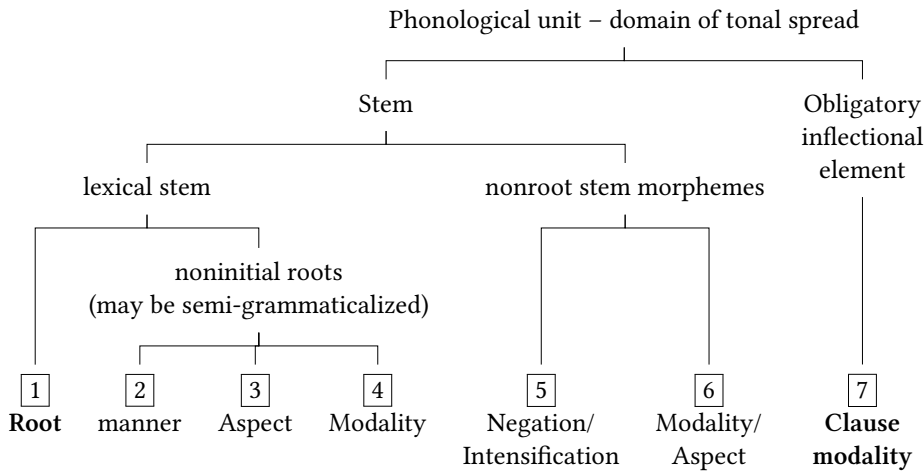
Phonological unit – domain of tonal spread

Stem		Obligatory inflectional element
lexical stem	nonroot stem morphemes	

n

noninitial roots
(may be semi-grammaticalized)

1	2	3	4	5	6	7
Root	manner	Aspect	Modality	Negation/ Intensification	Modality/ Aspect	**Clause modality**

Figure 3: Kotiria finite verb structure (adapted from Stenzel 2013: 245)

7 Conclusion

As this investigation of Hup morphological structure has explored, different diagnostics of constituency applied to the Hup verb are highly nonconvergent. Figure 4 illustrates this relative lack of isomorphism in the test results. A feature of particular typological and theoretical relevance concerning Hup constituency is the fact that the criterion of non-interruptability does not apply straightforwardly in the Hup verb – a challenge for perspectives on wordhood that prioritize non-interruptability as a cross-linguistically relevant diagnostic.

Nonetheless, the span between positions 6 (the verb root or base) and 20 (the boundary suffix) is meaningful in Hup, in that it delimits a morphosyntactic unit of minimal free occurrence, and a phonological unit relating to the minimal domain of stress (tone), as well as to the minimal domain of vowel copying and to final consonant deletion. As observed in §6 above, the properties that define this span probably emerged following Hup's divergence from its two more distant sister-languages (Nadëb and Dâw), propelled by contact with Tukanoan languages.

In comparison to the verb, the constituency diagnostics relevant to the Hup noun are more convergent (Figure 5). In particular, the span between positions 5 (the noun root) and 15 (elements relating to information structure and evidentiality) emerges as meaningful, representing the maximal unity of free occurrence, non-interruptability, and stress.
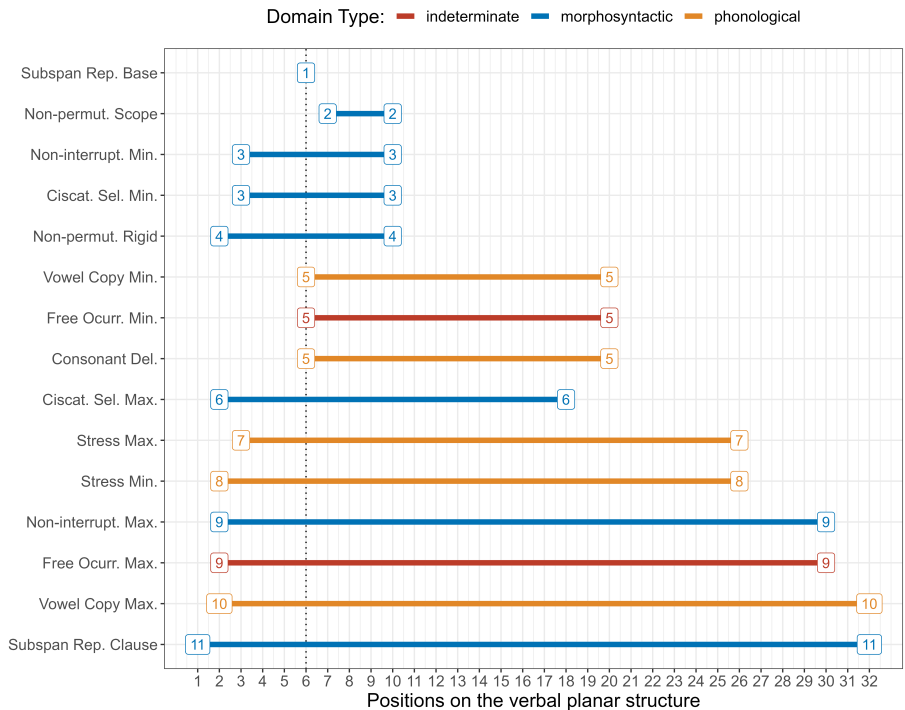
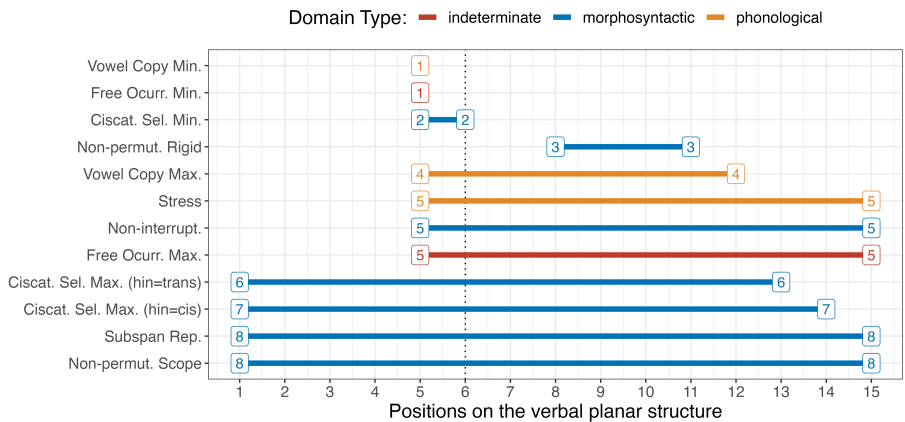Figure 4: The Hup verb: constituency diagnostics compared

Figure 5: The Hup verb: constituency diagnostics compared

The degree of mismatch seen here among Hup morphological spans, as defined by different morphosyntactic and phonological criteria, is undoubtedly behind the conflicting characterizations of Hup in the literature as relatively isolating or polysynthetic, particularly for the Hup verb. These mismatches also have practical implications, in that they create difficulties in establishing principled conventions for representing the orthographic word. Interestingly, a comparison with Dâw suggests that while these two languages might be construed as quite distinct with respect to their degree of synthesis, they actually differ only according to a few criteria, while others correspond. Ultimately, historical change relating to "degree of synthesis" involves realigning a whole set of features associated with constituency; there may be no principled reason to expect that these should all fall into line together and at the same time. Thus a view that constituency diagnostics must necessarily align across languages, or even within them, may be as untenable diachronically as it appears to be synchronically.

Acknowledgements

Abbreviations

References

Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.

Booij, Geert. 2009. Lexical integrity as a formal universal: A constructionist view. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 83–100. New York: Springer.

Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25–46. Berlin: Language Science Press. DOI: 10.5281/zenodo.583806.

Epps, Patience. 2001. *Hup collection.* Archive of the Indigenous Languages of Latin America. http://ailla.utexas.org.

Epps, Patience. 2007. The Vaupés melting pot: Tukanoan influence on Hup. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, 267–289. (Explorations in Linguistic Typology 4). Oxford: Oxford University Press.

Epps, Patience. 2008a. *A grammar of Hup* (Mouton Grammar Library 43). Berlin: Mouton de Gruyter.

Epps, Patience. 2008b. From 'wood' to future tense: Nominal origins of the future construction in Hup. *Studies in Language* 32(2). 382–403.

Epps, Patience. 2008c. Grammatical borrowing in Hup. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing: A cross-linguistic survey*, 551–565. Berlin: Mouton de Gruyter.

Epps, Patience. 2010. Linking valence change and modality: Diachronic evidence from Hup. *International Journal of American Linguistics* 76(3). 335–356.

Epps, Patience & Sunkulp Ananthanarayan. 2022. Stability in grammatical morphology: An Amazonian case-study. In Patience Epps, Danny Law & Na'ama Pat-El (eds.), *Historical linguistics and endangered languages: Exploring diversity in language change* (Routledge Series in Historical Linguistics), 121–151. New York: Routledge.

Epps, Patience & Katherine Bolaños. 2017. Reconsidering the 'Makú' family of northwest Amazonia. *International Journal of American Linguistics* 83(3). 467–507.

Epps, Patience, Karolin Obert & Luciana Storto. 2013+. *Daw collection.* http://ailla.utexas.org.

Martins, Silvana A. 2004. *Fonologia e gramática Dâw.* Amsterdam: LOT Dissertation Series.

Mithun, Marianne. 2000. The reordering of morphemes. In Spike Gildea (ed.), *Reconstructing grammar*, 231–258. Amsterdam/Philadelphia: John Benjamins.

Moore, Barbara & Gail L. Franklin. 1980. Verbal and nominal inflection in Hupda Makú. https://www.sil.org/resources/archives/79953.

Nichols, Johanna. 2008. Universals and diachrony: Some observations. In Jeff Good (ed.), *Linguistic universals and language change*, 287–294. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199298495.003.0012.

Payne, Doris L. 1990. Morphological characteristics of lowland South American languages. In Doris L. Payne (ed.), *Amazonian linguistics: Studies in lowland South American languages*, 213–241. Austin: University of Texas Press.

Ramirez, Henri. 2006. *A língua dos Hupd'äh do Alto Rio Negro: Dicionário e guia de conversação*. Sao Paulo: SSL/Libro.

Schiering, René, Balthasar Bickel & Kristine A. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709.

Schmidtke-Bode, Karsten & Eitan Grossman. 2019. Diachronic sources, functional motivations and the nature of the evidence: A synthesis. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology, diachronic sources, functional motivations and the nature of the evidence*, 223–241. Berlin: Language Science Press. DOI: 10.5281/zenodo.2583822.

Stenzel, Kristine. 2013. *A reference grammar of Kotiria Wanano*. Lincoln: University of Nebraska Press.

Tallman, Adam J. R. 2020. Beyond grammatical and phonological words. *Language and Linguistics Compass* 14(2). 1–14. DOI: 10.1111/lnc3.12364.

Tallman, Adam J. R. 2021. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal.