Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Constituency and convergence in the Americas

Edited by

Adam J.R. Tallman

Sandra Auderset

Hiroto Uchihara

Topics in Phonological Diversity

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Topics in Phonological Diversity

Editors: Natalia Kuznetsova (Università Cattolica del Sacro Cuore / Institute for Linguistic Studies, Russian Academy of Sciences) Cormac Anderson (MPI-SHH, Jena) Shelece Easterday (University of Hawai'i, Mānoa)

In this series:

1.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Constituency and convergence in the Americas

Edited by

Adam J.R. Tallman

Sandra Auderset

Hiroto Uchihara

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.). 2024.
Constituency and convergence in the Americas (Topics in Phonological Diversity).
Berlin: Language Science Press.

This title can be downloaded at:
redefine\lsURL
© 2024, the authors
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/
ISBN: no digital ISBN
no print ISBNs!
no DOI

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

redefine \publisherstreetaddress
redefine \publisherurl

Storage and cataloguing done by redefine \storageinstitution

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Contents

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Preface

This preface has no abstract and no authors, but cites **Nordhoff2018** nevertheless.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Acknowledgments

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 1

Introduction: phonological and morphosyntactic constituency in cross-linguistic perspective

Adam J.R. Tallman
Friedrich Schiller Universiät - Jena

I provide a brief history of the development of the ideas for the "constituency-convergence project", which this volume is a product of. I also motivate the project by discussing the shortcomings of Basic Linguistic Theory and Prosodic Phonology as description languages for studying constituency cross-linguistically. I briefly summarize the principles of the planar-fractal method and then provide an overview of the chapters in this volume.

1 Introduction

This volume presents a number of studies on constituency (phonological and morphosyntactic) in the languages of the Americas from a novel perspective. Constituency analyses, whether morphosyntactic or phonological, are typically conceptualized as being based on 'constituency tests.' Generally the constituency tests are used as a means to an end, a tool or a justification, to get at a particular constituency analysis - or more commonly to argue in favor of one constituency analysis over another where the constituency analyses are arrived through theoretical assumptions and intuition ("the factor of judgement" (Pike 1943: 75)). In this perspective, constituency tests might be "clues" to constituents (at best), but constituents are the units of description and comparison (e.g. (Wiltschko 2014: 44)).

Adam J.R. Tallman. 2024. Introduction: phonological and morphosyntactic constituency in cross-linguistic perspective. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 1–83. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In this volume a group of researchers consider phenomena from a variety of languages of the Americas to explore, critique and develop the notion of 'constituency *test*' as a unit for language description and comparison. Comparing languages in terms of constituency tests (or domains) is not the same as comparing languages in terms of 'constituents.' Constituents are embedded in constituency analyses which are arrived at by smoothing over (discarding, reinterpreting etc.) constituency tests to fit a set of theoretical positions or assumptions (e.g. binary branching, no-branch crossing, nesting, etc.).

The data are based on original field research by all of the authors and in some cases native speaker judgments. Participants of the volume approach linguistic phenomena from a variety of perspectives, but share the view that cross-linguistic study of constituent structure might be profitably engaged with by comparing languages in terms of 'constituency test results' themselves, rather than only abstract constituency structures proposed in the linguistics literature. This volume was also brought on by a sense that there is a gap in the literature on the relationship between 'constituent structure' and *constituency test* as a problem in cross-linguistic comparison. The vast majority of introductory syntax texts that introduce and explain the notion of constituency test rely only on examples from (standard) English, for example.

This does not mean that abstract constituency structures are rejected *per se*. Rather the volume is interested in critically engaging with the empirical basis for such constituency structures. Given the ever expanding panoply of competing morphosyntactic geometries found in the literature today, I would suggest that such a methodological orientation is helpful, if not necessary, for getting our bearings.

The notion of 'constituency test' is presented in a deceptively simple way in introductory syntax texts. When one recognizes the possibility that constituency tests can be and have been used in a biased manner in the linguistic literature (Croft 2001b, 2010), attempting to overcome this bias opens up a world of intricate complexity with competing structural analyses for language description and comparison. The intuition underlying this project is that this complexity is worth exploring and may lead us to overcome some longstanding epistemic and theoretical impasses in the field. It could lead us to abandon some longstanding, but inhibiting assumptions, and articulate new hypotheses concerning linguistic universals and diversity.

Historically, but especially since the development of (American) structural linguistics, the languages of Americas have been an important source of inspiration for understanding the nature of language variation. Languages of the Americas have not simply served as 'testing grounds' for already established hypotheses,

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

but as laboratories for the development of new perspectives on linguistic architecture. In my view the latter tradition has been attenuated in recent years, because of a strong tendency to presume that 'universal' architectures can be derived from studying a few European languages. Novel phenomena from other languages are studied as expressing deviations from the 'basic' patterns, but could not be used to challenge the fundamental *Bauplan* over which these patterns are described, compared and conceptualized.

It was a staple of the Boasian tradition to criticize traditional linguistic categories for their potential to be implicitly biased towards describing languages and cultures in the terms of languages and cultures which are dominant in European institutions (Stocking Jr. 1974, Rodseth 2022). This critical attitude was applied to traditional grammatical terminology. The issue of 'word' and 'constituent' are a classic concerns of Americanist linguistics (Boasian, Bloomfieldian, Post-Bloomfieldian) in this regard (Boas 1911, Bloomfield 1914, Hockett 1947, Pike 1972). In a sense, therefore, this volume attempts to reinvigorate the Boasian tradition of empirically based criticism of traditional categories, directing the criticism at the 'established' or 'basic' categories of general linguistics (phonological and morphosyntactic) 'word' and 'phrase.'[1] The strategy is to take a look at the 'diagnostics' for our presupposed structures and assess whether these really support the presumed grammatical architectures.

To avoid descending into a cacophony of conflicting terminologies, the multivariate autotypologizing methodology (Bickel & Nichols 2002) is leveraged and modified in service of this goal. The papers in this volume apply the "planar-fractal method", a typological description language coupled with a coding technique developed to visualize, critique, commensurate and measure constituency tests and their interrelations cross-linguistically. This method is not itself unproblematic, and it should be emphasized that it is a "tool" with its own biases and pitfalls (see Wimsatt (2007) for relevant discussion). Used in conjunction with other approaches, I think it can be a powerful technique for comparing and testing certain aspects of language structure. Moreover, for language description, the growing impression is that it has an obvious heuristic value.

Below I provide a brief history of how this volume came into being (Section 1). The papers in this volume reconceptualize some 'fundamental' notions in linguistics and I think that providing a brief history of how the perspective developed is a useful entry point.

[1]Whether the Boasians consistently approached all problems with such a critical stance is another matter (see Anderson (2019) for an important discussion of the shortcomings of the Boasian approach).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The rest of the sections provide brief discussions of the some theoretical ideas, concepts and distinctions that the approach discussed in this volume engages with. Section 3 discusses Basic Linguistic Theory. Section 4 discusses the Prosodic Hierarchy Hypothesis. Section 5 briefly discusses some methodological issues in typology.

Then I turn to providing a brief description of the planar-fractal method. A description of planar structures is described in 6. The fractal method for describing constituency tests is provided in Section 7 and a brief summary of the type of domains (constituency tests) used in this study is provided in Section 8. I then describe the papers of this volume in 9.

2 Where these ideas come from

This volume came about through ongoing collaboration and conversations between a number of researchers engaged in language description starting in about 2017. The smoothest entry point into understanding the perspective adopted in the volume might be from my own failure to analyze Chácobo, a southern Pano language of Bolivia, according to certain prescribed orthodoxies: Basic Linguistic Theory (Dixon & Aikhenvald 2002), and prosodic phonology (Nespor 2007, Anderson 2005).

Verbal word structure or word formation in Pano languages is modelled and described according to the following template (Loos 1999, Fleck 2003, Valenzuela 2003a, Fleck 2013, Neely 2019, de Souza 2020).

(1) Pano verbal 'word'
PREFIX - VERB ROOT - DERIV. SUFFIXES - INFL. SUFFIXES

Nouns follow a similar template except that inflectional and many derivational elements are understood as occurring at the end of *phrases* rather than words. In the verb complex the prefix codes the body-part (or something analogous for like a 'trunk' for a tree) of an S or P argument (typically). The derivational suffixes code a number of concepts such as valence, aspect, emotion, modality, and associated motion. Inflectional suffixes code aspect, tense, evidentiality, temporal distance and (depending on the language) person and number. An example comes from the verb *da-daiʃ-tsik-kid* 'habitually gnawing on trunks' provided in 2: *da-* 'trunk' is the prefix; *daiʃ* 'eat gnawing' is the verbal root; *tsik* 'diminutive' is a derivational suffix; *-kid* 'habitual is an inflectional suffix.

(2) Matses:

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

kwɨte *da-daiʃ-tsɨk-kid* *madu-n* *sipi-n*
dicot.tree trunk-eat.gnawing-DIM-HAB demon-GEN tamarin-ERG
'Pygmy marmosets gnaw the trunks of dicot trees' (Fleck 2003: 342)

From Fleck's description one can discern that the verbal 'word' in Matses is itself a minimal free form, cannot be interrupted by a another free form or any distributionally 'promiscuous' elements such as adverbial clitics and is the domain for stress. There is evidence that in other Pano languages the verbal word has a somewhat 'looser' constituency, however. In Shipibo, the verbal word (which has the same *basic* structure as that of Matses) can be interrupted by bound adverbial clitics and second position clitics. Valenzuela 2003a: 145-146 refers to the relevant adverbial morphemes as 'less-fixed clitics.' An example of a less-fixed clitic is provided with *=ribi* 'also' in 3.

(3) Shipibo-Konibo:
ɨ-a *ka-i-tian resto no-n* *kaibo-baon-**ribi*** *ɨ-a* *tʃiban-a*
1-ABS go-S-DS rest 1PL-GEN fellow.Shipibo-PL:ERG-**also** 1-ABS follow-PP2
iki, *onan-kas-kin-**ribi***
AUX know-want-SSSA-**also**
'When I was going (to the Salt Mountain),the rest of my fellow Shipibo follows me, wanting to know (the way) too.' (Valenzuela 2003b: 145)

Evidence for the looser constituency comes from the fact that some less-fixed clitics such as *=ribi =riba* 'also' can interrupt the verbal word. This poses problems for some definitions of wordhood, insofar as the word-internal form is regarded as the same morpheme (it is unclear why it should not be); words should be noninterruptable (Martinet 1962: 92 Bauer 2017: 17).

(4) Shipibo-Konibo:
moa *icha* *baritia pekao, Shipibo joni-bo* *moa*
already many time after Shipibo person-PL:ABS already
*kai-**ribi**-kan-a* *iki* *ja* *kimisha joni-nko-ni-a-x*
reproduce-**also**-PL-PP1 AUX that three person-LOC-ligature-ABL-S
'After many years, the Shipibo reproduced again from these three people.' (Valenzuela 2003b: 146)

Still, the verbal word in Shipibo is a stress domain and cannot be interrupted by any *free form*. It also passes the 'free utterance' or minimal free form test (Bloomfield 1933, Hockett 1958) (as far as I can discern from the available descriptions).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In Chácobo an analogous span of structure is also a minimal free form (boxed in the example below).

(5) Chácobo:
ina hoṣo tsi kiá *ta-niṣ-i-tiki(n)-yamit(t)-ki*
dog white LNK REP foot-tie-ITR-again-DST-DECL:PST
'The white dog got its feet tied up again.'

In Chácobo, however, the verbal word is interruptable, not just by a free form, but by a whole noun phrase. The prefix and root can front leaving behind the 'inflectional suffixes.' This is illustrated in the example below.

(6) Chácobo:
ta-niṣ-i *tsi kia ina hoṣo* *-tiki-yamit-ki*
foot-tie-ITR LNK REP dog white -again DST-DECL:PST
'The white dog got its feet tied up again.'

The question then arises as to how we characterize Chácobo and Matses in terms of their morphological profiles. Perhaps, we should say that Chácobo and Matses display radically different structural organizations vis-à-vis the distribution of elements within morphology or syntax. Chácobo would be isolating and Matses polysynthetic. Such a position, however, ignores the fact that Chácobo is just one step more extreme than Shipibo in terms of the looseness of the analogous span of structure from prefix to inflection. The difference between Matses, Shipibo and Chácobo is not one of drastic differences in structure from one language to the next, but rather a matter of degree regarding how well wordhood tests, or perhaps constituency tests in general, align around a particular domain of structure. Claiming that Chácobo or Matses has taken a drastic jump from polysynthetic to analytic or analytic to polysynthetic structure obscures the structural similarities between the two languages and the fact that Shipibo-Konibo stands somewhere in between.[2]

Perhaps we should claim that all Pano languages are actually like Chácobo and that the relative tightness of the Matses verbal constituent is 'superficial.' Such a move would obscure interesting typological differences between the languages, however. Yet another approach would be to claim that noninterruption is not a useful test for wordhood (Dixon & Aikhenvald 2002), but such a claim suffers

[2]It is not yet known whether Proto-Pano should be regarded as polysynthetic or not, but in a recent talk on Amawaka, another Pano language, at the Association for Linguistic Typology (2022, UT Austin), Pilar Valenzuela suggested that the proto-language was likely more analytic.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

from arbitrariness. It seems that our 'basic' categories for linguistic description obfuscate rather than clarify variation and similarity in the Pano languages.

Another take would be to claim that the above discussion focuses too much on non-convergences between specific wordhood tests, rather than looking at how diagnostics for wordhood pattern globally (Tallman 2018). The tests might show a *tendency to align* over a tendency not to. This is sometimes claimed in the case of wordhood tests (Matthews 2002) and constituency tests in general (Carnie 2010, Bennett & Elfner 2019). But the claims have been made *ex cathedra* in the absence of a systematic typological study.[3].

And a more serious problem arises for language comparison. Even if a meta-study were to be conducted showing that in case after case, researchers did not show a tendency to report 'nonconvergences', such a result could plausibly arise from 'selection bias' - picking just those results and focusing on just those constructions that illustrate convergence and discarding those that do not as irrelevant (Croft 2001b, Haspelmath 2011). If a linguist is told that all languages have words as long as one finds the right criteria (Dixon & Aikhenvald 2002), they are going to find them insofar as there are *criteria* to be found at all, under reporting, if not just missing, contradictory results so as not to provoke eye-brow raising from reviewers. *Wer suchet der findet* as the expression goes.[4]

In an attempt to assess the issue of wordhood in Chácobo more globally, I culled the literature for all wordhood tests I could find (Haspelmath 2011 for a useful, but preliminary review). But two problems became apparent. The first is that wordhood diagnostics are often stated in highly ambiguous ways in the literature. A given wordhood test is often vague such that it has multiple interpretations. For instance, when we consider noninterruption, the question arises as to what the interrupting element should be: a free-form (Haspelmath 2011), or some 'promiscuous' clitic-like element that can be bound (Bauer 2017). Insofar as these versions of the same test do not give us the same result, which one do we apply? This problem is not necessarily fatal if one rigorously reports all available interpretations of wordhood results in the literature.

The second problem is more fatal. It is not clear that there is a distinction between wordhood test and constituency test in general. The latter problem became

[3]As reviewed below, the one typological study that investigated the question provides the opposite result from what is typically claimed regarding convergence (Schiering et al. 2012, Bickel et al. 2009)

[4]The issues brought up by Croft (2001b) and Haspelmath (2011) about the possibility that diagnostics are cherry-picked just so they support a favored theory is reminiscent of debates about p-hacking and data dredging in discussions of replicability in the sciences in general (Tallman 2021b), which is why I refer to the phenomenon as 'selection bias', rather than using Croft´s term 'methodological opportunism.'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

apparent when I started 'comparing notes' with other field researchers from UT and began to take a closer look at tests for phrase-level constituency. To give one example, noninterruption as a wordhood test is actually not clearly distinct from 'movement' or 'discontinuity' as a phrasehood test. The difference appears to be one of conceptualization, not empirical reality.[5]

Putting the second problem aside, the results of Chácobo reveal very few convergences given the number of tests applied. To give the reader an idea of how tests decompose the traditional 'word', consider the example sentence from Chácobo in 7.

(7) Chácobo:
(ta) niʂ i (βiki) (βona) (tiki) ki (rá)
(foot) tie ITR (INTRC) (going:TR/PL) (again) DECL:PST (ASR)
'Again, they (e.g. the dogs) were tying each other by their feet as they went. (e.g. on a leash).'

The traditional Panoanist analysis would posit the structure in 8 (see Zingg (1998) for example) (where deriv. is 'derivational' and infl is 'inflectional').[6]

[5]In the literature on syntax, the analogue to noninterruptability is referred to as (non)displacement. A phrasal constituent is one that can be displaced with all of its elements remaining adjacent (Kroeger 2005: 25Levine 2017: 8), or a phrasal constituent is one which cannot be 'discontinuous' (Louagie 2021: 114) - distinct formulations which mean the same thing as far as I can see. To illustrate the basic empirical identity between these formulations consider a grammar with just three elements: *a*, *b* and *c*. the grammar outputs the following strings.

(i) a, b, c, ab, ac, ba, ca, bc, abc, bca

We know that all cases where *b* and *c* occur they cannot be interrupted by *a*. We can formulate the generalization in two ways.

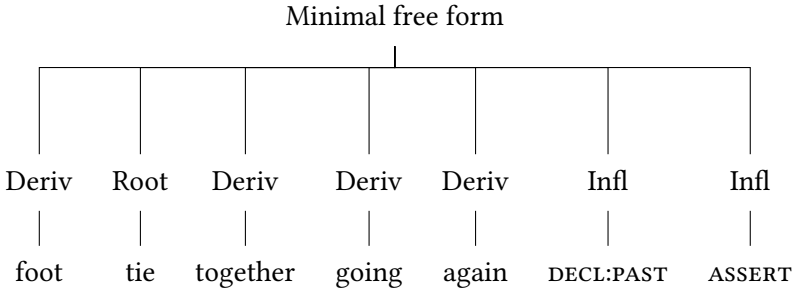
(ii) a. Noninterruption: *ab* is a constituent because it cannot be interrupted (by *a* for example)
b. Displacement: *ab* is a constituent because it can be displaced to the left (or right) side of

With some reflection, therefore, we can see that (non)displacement or (non)discontinuity can be regarded as formulations of noninterruptability, albeit with a different focus. 'Displacement' evokes a metaphor where the candidate constituent 'moves' without breaking into pieces. 'Noninterruption' evokes a metaphor where the candidate constituent does not break to pieces when subjected to the movement of extraneous elements. Likewise in displacement, extraneous elements stand still, whereas in noninterruption the pieces of the candidate constituent stand still.

[6]Note that in the following discussion trees which have straight rectangular edges are used

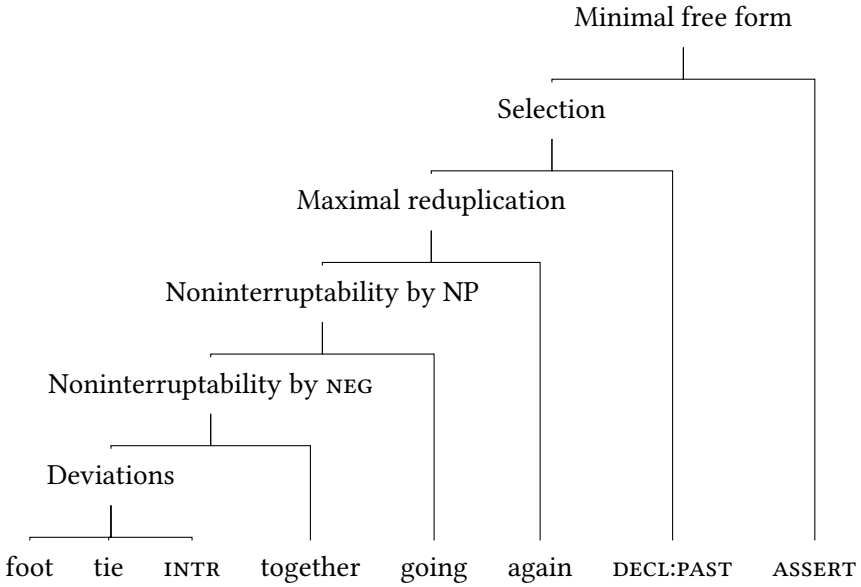
Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(8)



Wordhood (or constituency) tests parse the sentence up as in 9 (see Tallman (2021a) for relevant terminology). Thus, if one prioritizes deviations from biuniqueness, Chácobo is analytic, actually close to isolating. If one considers the minimum free form test, Chácobo is polysynthetic.

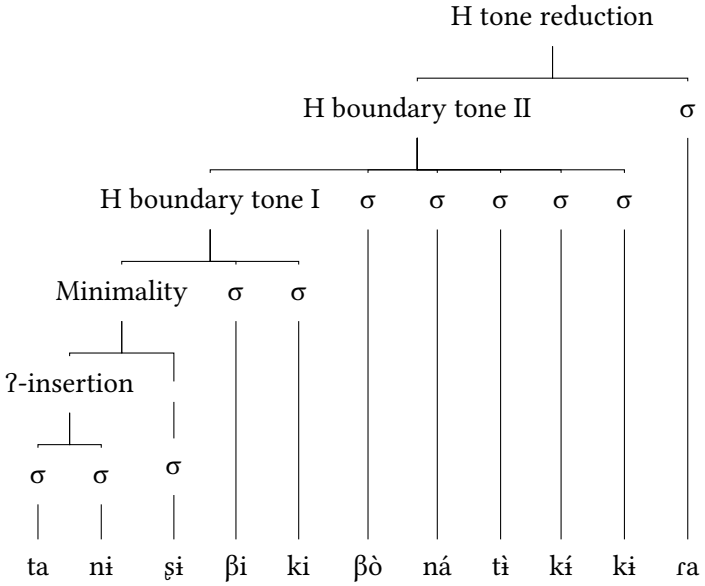
(9)



for representing constituency tests and those with triangular edges are used for representing constituency analyses.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(10)



The situation is not obviously less ambiguous with phonological domains. 10 depicts phonological domains in Chácobo showing that the language has 'small' phonological words if glottal stop insertion is chosen as word-identifying and large ones if H-tone reduction is chosen (see Tallman (2018) for a complete description of the relevant phonological processes).[7] The few convergences that can be found could be attributed to chance. With 22 constituency tests and 28 sentence level structural positions, the probability of two tests converging by accident is relatively high (Tallman 2021a).

The ambiguity here matters for linguistic theory generally. Claims about lexical integrity are not meaningfully testable, or just incoherent, if they are highly contingent on which of an open ended set of competing wordhood candidates are chosen (Tallman 2021b). We cannot discern how Chácobo data relate to the prosodic hierarchy if labeling of the relevant domains is mostly arbitrary (Tallman 2021c). Claims about the relative autonomy of morphology and what distinguishes morphology from syntax (Anderson 2015) are likewise meaningless if they rest on arbitrary choices about where to cut the division between these domains (Tallman & Auderset 2023). We cannot felicitously compare the phonetics

[7] One could question this argument on the grounds that the smallest domain should be the 'phonological word.' But then the question arises as to which domain is the phonological phrase. Domains smaller than the phonological word (the 'Pstem') have also been argued to be necessary for some languages (Downing & Kadenge 2020), which reintroduces the ambiguity. Such problems are discussed in detail in Section 4.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of boundary phenomena in cross-linguistic perspective (Sonderegger & Morgan 2018, Seifart et al. 2021), if it is not clear what level the boundaries identify.

A question arises at this point as to how general the problem of non-convergence really is cross-linguistically. This is where the larger collaboration between more researchers begins. A methodology for reporting and coding constituency test results was developed in collaboration with linguists doing fieldwork on native languages of the Americas, some of them native speakers of these languages, at the University of Texas at Austin. The collaboration began in the context of a seminar on *Morphological Typology* taught by Patience Epps and Anthony C. Woodbury. In fact, many of the tests that were applied to Chácobo in Tallman (2021a) were suggested by other fieldworkers while we attempted to operationalize wordhood tests in language after language. I did not invent the variety of tests myself, rather they gradually emerged from discussing how different linguists would apply the tests in languages they specialized in.

The notion of a planar structure and test fracturing grew out of this collaboration. A planar structure is an array of structural positions that code the relative ordering of elements in a referential (nominal) or predicate (verbal) domain. The planar structure is a hypothesis space for coding constituency test results as *spans* over adjacent positions. The hypothesis space homogenizes morphological and syntactic representations *pro tempore*. If 'words' or the word-phrase distinction are valid constituents they do not emerge from the planar structure itself, but from the patterning of constituency test results over the planar structure. The planar structure codes positions with sequential numbers and constituency test results are coded as spans over those positions.

Fracturing is the methodology employed to deal with the ambiguity of relatively abstract constituency tests or domains in their application to real empirical phenomena. When ambiguity is recognized, the researcher decomposes ('fractures') the test into multiple versions. For instance, consider the case described above with noninterruptability. Rather than choosing a single 'correct' interpretation of noninterruptability, we *fracture* the noninterruption domain into a domain not interruptable by a free form and a different domain interruptable by a 'promiscuous' element. If linguist A discovers a version of a test not identified by linguistic B, then the latter makes an attempt to apply the new version of the test to their language data as well. Thus the variables of constituenthood evolve through the reciprocal interaction of fieldworkers and become increasingly fine-grained and more comparable in the process. The research conducted in this fashion also benefits from the fact that researchers approach the issue of constituency from different intellectual traditions, further enriching the variables (see Sections 6 and 7).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

A researcher might suspect that the nonconvergences found with Chácobo would be common cross-linguistically (Bickel & Zúñiga 2017). However, application of the methodology revealed that there are apparently radical differences between languages with respect to the degree to which independent morphosyntactic and phonological principles tend to cluster. Consider the following two orthographic 'words' in Chácobo and Central Alaskan Yupik. The elements in numbers are positions in the respective planar structures (see below).

(11) Chácobo (Pano):
tɨpas$_{8}$ *wɨni*$_{16}$ *-tsa*$_{16}$ *-kas*$_{17}$ *-i*$_{24}$ *-kiá*$_{25}$
murder -before.someone immediately:ITR:SG -want -DECL -REP
'He wanted to murder him immediately before it was too late (it is said).'
(Tallman 2017: 54)

(12) Central Alaskan Yupik (Inuit-Yupik-Unangan):
quuyurni$_{2}$*-arte*$_{3}$*-llru*$_{6}$*-yaaqe*$_{8}$*-llini*$_{9}$*-u*$_{12}$*-q*$_{16}$
smile-suddenly-did-alas-evidentlyIND-3SG.S
'Evidently, s/he suddenly smiled, but alas' (Woodbury 2002: 85)

In Chácobo, the relevant orthographic word is identified only by a version of the minimal free occurrence test. The orthographic word is also identified by free occurrence in Central Alaskan Yupik. However, Yupik case is identified by stress prominence, segmental allomorphy, 'say' conjunction, selection, fixedness and is furthermore a repair domain.

Planar structures were constructed for Chácobo and Central Alaskan Yupik (see Tallman 2021a and Woodbury (This Volume) respectively). One way of displaying the results of constituency tests over the planar structure is by a convergence plot. A convergence plot is a strip plot that has the positions of the planar structure on the x-axis and the coded constituency tests on the y-axis. Figures 1 and 2 display convergence plots for Chácobo and Central Alaskan Yupik respectively. A convergence between tests is where their left and right edges align on the x-axis. We can see from these plots that while both languages display misalignments, Central Alaskan Yupik has a domain of structure from position 2 to 16 (the traditional orthographic word in this language), where a number of constituency tests align. In Chácobo, there is much less convergence overall.

Starting in 2017 at UT Austin, a number of researchers applied the planar-fractal method to a number of languages. The method travelled to the University of Ottawa and to the Laboratoire Dynamique du Langage (CNRS, Université de Lyon II), eventually diffusing to researchers at other institutions. The planar

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

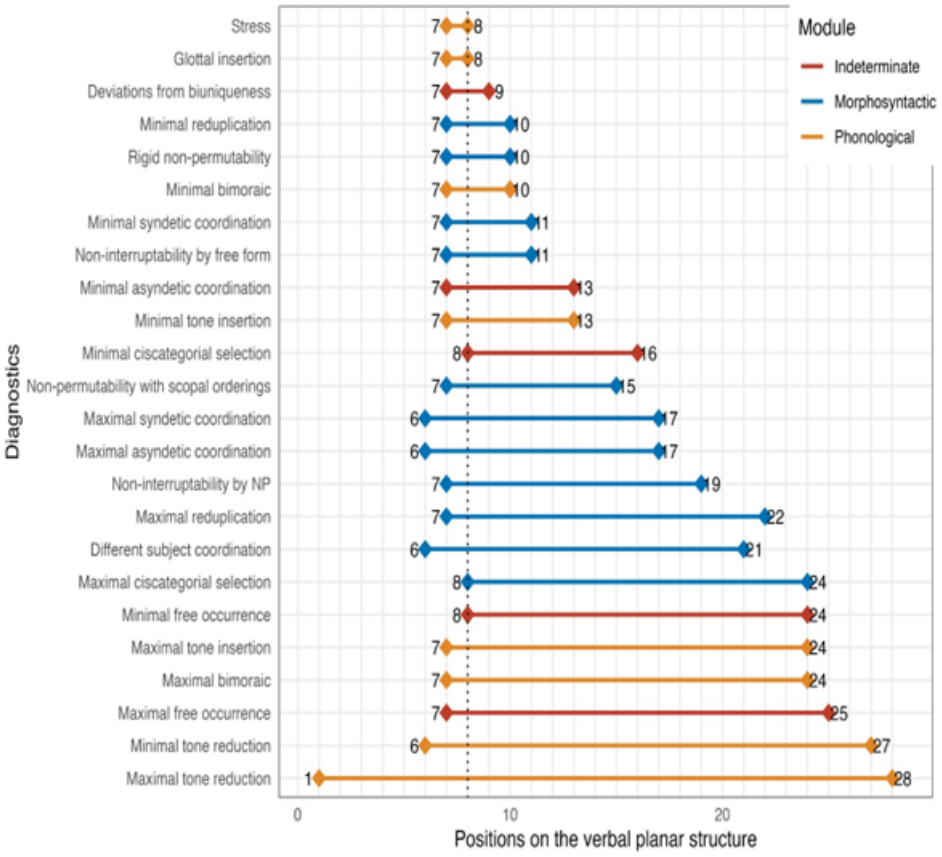


Figure 1: Constituency convergence plot for Chácobo (Pano), see Tallman 2021a for details

structure and application of the constituency tests is applied by researchers that are experts or expert native language speakers on the relevant languages. Researchers are asked to apply and critique constituency tests presented in the literature using the methodology and, where possible, reflect on how the results relate to published theoretical literature. A researcher might add a new constituency test not reported by other researchers. The other researchers in the project are then asked to apply the new test insofar as it is well formulated enough to apply without ambiguity. Researchers in the relevant project are encouraged to not just apply the methodology but critique and develop it as well. The variables for comparison are thus developed enriched through original empirical research. The idea is to pool perspectives and experiences from different

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

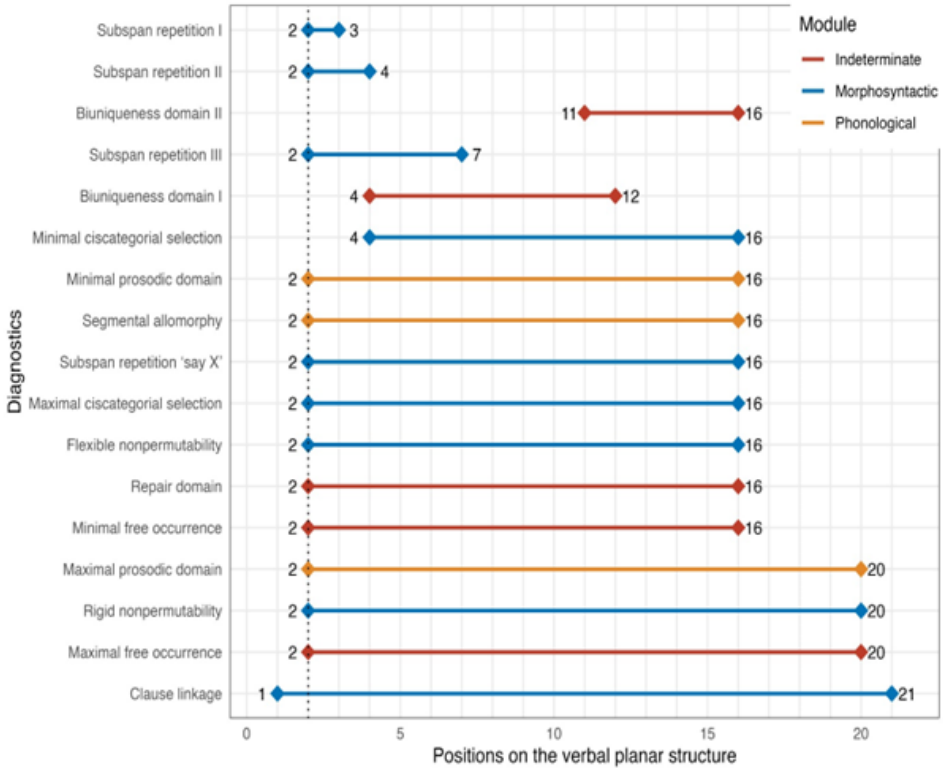


Figure 2: Constituency convergence plot for Central Alaskan Yupik, see Woodbury (This Volume) for details

researchers to enrich the variables, rather than applying them in a predefined top-down fashion or seeking to rally diagnostics here-and-there to ratify predefined formal categories such as 'word' or 'phrase.'

This book presents the ongoing results of this collaborative project. The first goal was to use the methodology to help enrich descriptions of lesser described languages. Many of the papers were written in the context of a PhD project on the documentation of the language in question. Secondly, the methodology is used to test claims about constituency and wordhood stated in the literature from a broader cross-linguistic perspective. The results suggest that there is much more cross-linguistic variation in constituency structure than would appear to be expected by the current literature. Whether the methodology can be used to test competing hypotheses about constituency structure is partially contingent on whether those hypotheses are precise enough to be testable to begin with. In

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

this respect, the methodology also provides a data structure for typological comparison that allows development of more testable hypotheses. I think the participants in this project have overall found that the methodology provides a powerful discovery procedure for the purposes of enriching linguistic description and documentation. The results have revealed that many claims about typological regularities and variation in wordhood and constituency are oversimplified and should be revised.

3 Basic linguistic theory

Basic Linguistic Theory (BLT) seeks to provide a general framework and methodology for linguistic description and typological comparison (Dixon 1997, 2010a). The framework has been the most influential in language description over the past 20 years.[8] Despite its near hegemony in descriptive linguistics, the framework is not without its critics (McGregor 2021). There is also some question as to whether all linguists interpret 'Basic Linguistic Theory' in the same way (Haspelmath 2008).

In what follows I will be concerned with the notion of BLT represented in R.M.W. Dixon's authoritative statement on the approach (Dixon 2010a,b). I will focus specifically on the approach to grammatical and phonological wordhood within BLT as articulated in Dixon & Aikhenvald 2002, Dixon 2010b and Aikhenvald et al. 2020, refer to other authors where relevant. I focus on this approach to grammatical and phonological wordhood for two reasons. First, it is my impression that it has the status of a virtual orthodoxy within linguistic description: descriptive linguists *assume* that the units and 'phonological' and 'grammatical' word and describe the languages in those terms, rather than investigate, let alone test the claim. Secondly, the methodology for this project partially developed as a critique of the BLT framework for describing and comparing grammatical and phonological words. Thirdly, many of the assumptions of BLT are commonplace across linguistics and approaches to the relationship between morphosyntax and phonology. In what follows I hope to highlight these assumptions, pointing out which of them I think are empirically unfounded.

[8] I should point out that this is a very subjective impression. It is somewhat difficult to judge how influential BLT is in grammar writing because it probably tends to depend on the domain of enquiry and analyses or assumptions can adopted in degrees rather than *in toto*. It would be hard to say that BLT has had much influence on the writing of phonology chapters in grammars over the years where the trend is to include more and more detailed phonetic information. I do not think it is too controversial though to point out that in the domain of wordhood it has become a standard.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The BLT approach recognizes that diagnostics for wordhood do not necessarily align with one another. BLT solves this by positing that grammatical words and phonological words should be distinguished. A basic statement of the how to study words in particular languages and cross-linguistically is summarized by Dixon (Dixon 2010b: 10).

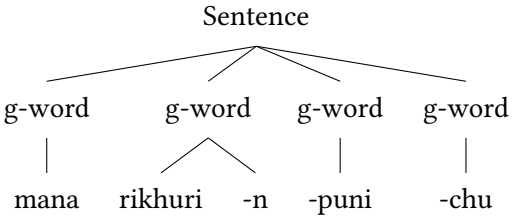
(13) "(a) Recognize 'phonological word', determined on entirely phonological principles.
(b) Recognize 'grammatical word', determined on exclusively grammatical (that is morphological and syntactic) principles.
(c) Compare the two units. In some languages, grammatical word and phonological word may coincide. In other languages, grammatical and phonological word will coincide in most cases, but with a number of instances where one grammatical word may consist of more than one phonological word, and/or vice versa." (Dixon 2010b: 10)

By phonological principles Dixon refers to phonological constraints (e.g. no coda consonants in a specific domain) and phonological processes (e.g. intervocalic voicing). It is not clear whether phonological principles also include so-called 'post-lexical' processes or phonetic modifications related to phonological constituency generally (more on this late). Grammatical principles refer to properties holding of specific domains of structure (e.g. inability to permute elements or re-curse constituents). It is unclear how grammatical principles exclude phrase identifying processes. The identification of a distinction between grammatical and phonological words, of course, represents an important advance in linguistic description. By allowing grammatical and phonological words to misalign, it allows one to capture the generalizations that hold of these constituents while capturing some of the complexities of the relationship between phonology and morphosyntax. For example, recognizing a distinction between grammatical and phonological words allows one to capture the differences and similarities between affixes and clitics in South Bolivian Quechua. The misalignment between g(rammatical)-words and p(honological)-words is represented with the labelled diagram below the example.

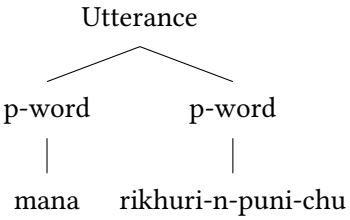
(14) g-word ⊂ p-word
South Bolivian Quechua:
mana rikhuri-n=puní=chu
NEG appear-3=certainly=NEG
'It certainly did not appear.'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(15) a.



b.



The g-words are elements or combinations of elements that can be displaced but with their internal parts in tact. The g-word *rikhuri-n* is not interruptable by a free form or clitic element and the internal parts of this constituent display little variable ordering. The clitics *=puni* and *=chu* are not part of the grammatical word because they can occur right-adjacent to a noun phrase as well (without necessarily corresponding to a difference in meaning).

However, the principles for identifying g-words (morphemes of combinations of morphemes that cannot be interrupted or split apart into pieces) do not line up consistently with phonological principles we can rally for identifying p-words. The clitics, while being independent g-words are incorporated into a pitch accent domain of the verb (projected from the verb root). The pitch accent domain is identified based on the distribution of Low-High* pitch accents on the penultimate syllable in of the relevant domain (the p-word in the domain above).[9]

BLT is not particularly clear about what phonological and grammatical principles identify 'phrases.' It is only stated that some sort of grammatical hierarchy exists (Dixon 2010b: 33). In the Quechua case above in particular it is not clear whether the p-word should instead be regarded as a phonological phrase, for instance.

Another type of misalignment warranted by BLT is where the phonological words are smaller than morphosyntactic words. An example comes from Atkan Aleut. Woodbury (2011) refers to pronominal elements that are obligatorily left-adjacent to the verb stem as 'unclitics.' (Zúñiga (2014) refers to these as 'anti-clitics'). They obey a principle of contiguity for g-words, but still have other

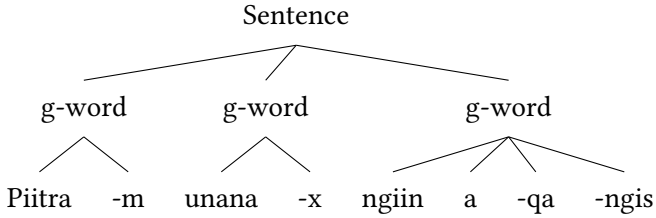
[9]This means that the high part of the tone is realized on the 'stressed' syllable and the low pitch is (typically) realized on the previous syllable.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

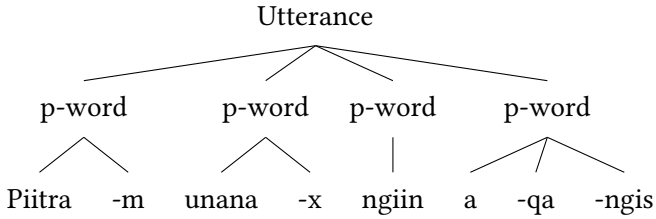
properties that Woodbury (2011) associates with p-words. A simplified depiction of the analysis of such forms presented in Woodbury (2011: 129) is presented below.

(16) p-word ⊂ g-word
Atkan Aleut:
Piitra-m *unana-x* *ngiin* *a-qa-ngis*
Peter-REL.SG cook-ABS.SG for.3.PL be-PST1-3PL.NS/3.SG.S
'Peter was a cook for them.'

(17) a.



b.



Woodbury argues that the element *ngiin*, while being a separate p-word is part of the g-word of the rest of the verb. It is an unclitic, because it inverts the standard relationship definition of clitics as prosodically dependent, but grammatically independent. Woodbury does not explain why the g-word which takes in two inflected elements could not be considered a phrasal or subphrasal constituent. But, the point is that his description is broadly in line with the assumptions of BLT, despite the misalignment.

These types of misalignments (g-word ⊂ p-word and p-word ⊂ g-word) exhaust what is statable in the BLT approach without modification. The researcher identifies grammatical and phonological principles, refers to the domains of structure where these principles hold as grammatical and phonological words respectively and describes how they align or do not.

There are at least two other types of misalignments that BLT does not have the vocabulary to express. These are cases where different candidate g-words (g-domains) or different candidate p-words (p-domains) misalign with each other. These were already discussed in Section 1, but they are worth mentioning again.

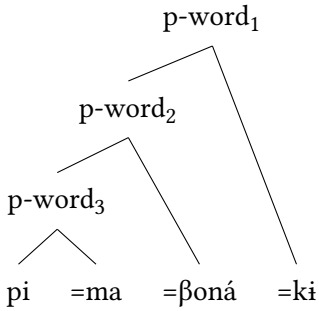
Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

For an example of cases where candidate p-words misalign with each other consider the example in 18.

(18) p-word$_1$ ≠ p-word$_2$ ≠ p-word$_3$ ...
Chácobo (Pano):
pi=má=βoná=kɨ
eat=CAUS=going=DECL:PST
'He made him eat on the go.'

In Chácobo the constituent *pi=ma* 'causative to eat' could be regarded as a p-word on the grounds that it is a domain of obligatory minimality, without *=ma* 'causative', the verb root can lengthen (*pii=βona...* 'eat while going.' However, it would not be accurate to simply state that *=βoná* 'going' does not phonologically interact with the rest of the verb complex as the identification of *pi=ma* 'make someone eat' as the phonological word implies. The clitic *=βoná* 'going' blocks the insertion of a default high tone by having a lexical tone itself. In cases where the rest of the verb complex has no underlying high tone, the presence of a high tone on *=βoná* blocks high tone insertion. For instance, without a high tone bearing suffix *hana* 'leave' is realized with a high tone on the first syllable, but otherwise this is blocked by morpheme like *=βoná*. Therefore we could also say that *pi=ma=βona* is the phonological word. This would ignore the fact that a different phonological principle identifies the whole string *pi=ma=βoná=kɨ* as a phonological word, however. All of the aforementioned elements are in a domain of obligatory tone reduction whereby adjacent lexical high tones delete Tallman 2018, 2021a. The ambiguity of is depicted in 19.

(19)



A number of issues arise in this case. One might claim that either p-word$_1$ or p-word$_2$ are phonological phrases (or 'composite groups'). the labelling issue (phonological word or phonological phrase) highlights a general problem

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

with the BLT framework.[10]. "Phonological principles" also apply at higher levels of structure. These data highlight the fact that an adequate typology of phonological and grammatical words cannot be decontextualized from issues of constituency in general.

Examples where grammatical principles misalign and thus provide competing notions of g-words are not hard to come by either.[11] Consider the following example from Teotitlán del Valle Zapotec.

(20) g-word$_1$ ≠ g-word$_2$ ≠ g-word$_3$ ...
Teotitlán del Valle Zapotec:
r- *œ-* *sut* *-nœ̄* *-i^{ʔ}ny* *=zá* *=an* *lą̆:n*
HAB- going- going:play -COM -DIM =also 3SG.INF 3SG.INF
'S/he goes to play with him/her (how nice!).' (Guttierez & Uchihara This Volume))

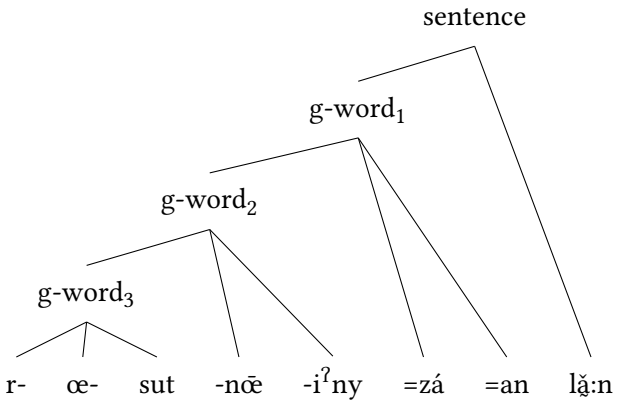
The syntagm *r-œ-sut* 'going to play' is a g-word under principles of selection, minimal free occurrence and sharing under conjunction. The syntagm *r-œ-sut-nœ̄-i^{ʔ}n* is a g-word under principles of nonpermutability and noninterruption by a free form. The syntagm *r-œ-sut-nœ̄-i^{ʔ}ny=za=an* is a g-word under principles of noninterruption by a noun phrase, repetition under conjunction and maximal free occurrence (Guttierez & Uchihara This Volume). The full picture is hard to depict in a tree diagram because a rigorous application of constituency tests gives us bracketing paradoxes in TV Zapotec. A simplified depiction of the results is provided below.

[10]Many current prosodic phonology analyses also posit that prosodic domains can 'recurse.' One might argue that 19 provides evidence that p-words are recursive in Chácobo. However, adopting recursion does not address the issue of ambiguity in label assignment, but rather exacerbates it, increasing the potential of arbitrariness of label assignment: in the case above, one could also claim that every single one of the candidate p-word domains are recursed phonological phrases, or perhaps any other layer of the prosodic hierarchy (see Section 4.3)

[11]Many morphosyntactic theories seem to be motivated by the fact that grammatical principles misalign, such as Baker 1988 movement analysis and Sadock 1991 autolexical approach to noun–incorporation. These authors do not appear to question the identification of 'words', however. They seem to rely heavily on orthographic practices to parse up the boundaries between the modules that their theories presuppose.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(21)



The same issues arise in this case. What is designated g-word$_1$ or g-word$_2$ could perhaps be reanalyzed as a 'phrase', but such an analysis does not fall out of the principles described in 13.

Given the fact that p-word and g-word domains misalign, a naive linguist might wonder what the purpose is in identifying 'words' at all in the description and comparison of individual languages. Entertaining such a possibility contradicts a central dogma underlying much contemporary descriptive and theoretical linguistics, however. I refer to this as the 'word bisection dogma.' Dixon articulates the dogma succinctly.

(22) The word bisection dogma:
"Units 'phonological word' and 'grammatical word' can without doubt, be recognized for all languages" (Dixon 2010a: 7).

I use the expression 'bisection', because the abstract notion of 'word' only need be split into two versions in this formulation. I refer to the claim as a 'dogma', because it is adopted uncritically in much language comparison and typology. If a descriptive linguist claims that the principle does not apply or work for a given language, they are generally treated as ignorant or insane.

On one reading the claim in 22 is simply a tautology, and, therefore, the expression 'without doubt' is warranted. I refer to this as the 'fiat-based word bisection dogma.' On another reading, Dixon is making an interesting empirical claim about the structure of all (or most?) languages. I refer to this as the 'empirical word bisection dogma.' On this reading the 'without doubt' expression is not warranted based on our current knowledge. The fiat-based and empirical interpretations of Dixon's claim should be kept distinct. However, many researchers

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

seem to assume that the more substantive empirical claim follows from the fiat-based one, which is fallacious. Below, I explain the issue in more detail.

The fiat-based word bisection dogma follows from the fact that (domain bound) grammatical and phonological principles exist at all. Once the linguist has found some domain of structure where a grammatical principle holds (e.g. 'fixedness of order') one can 'recognize' that domain as a g-word. If one finds another domain where a different grammatical principle holds (e.g. "noninterruption by a free form"), there is no problem at all if this does not line up with the domain that was already christened as a 'word.' When we have competing domains, the linguist simply *arbitrarily* designates one of the domains as a 'word', discarding the other grammatical principles as irrelevant or unreliable. Another linguist (or even the same linguist) could arbitrarily refer to the second domain as a g-word, even if they do not line up. The same holds for phonological domains. If a stress domain and a vowel harmony domain misalign, just christen one as *the* phonological word and be done with it. One need only insist that the other domain not-so-christened is not a reliable criterion in the language in question.[12] Since there is no justificatory logic behind fiat-based designations apart from appeals to authority such an explanation will suffice.

On the tautological interpretation Dixon is simply referring to the linguist's ability to label certain domains 'p-word' and/or 'g-word.' No claim is made about g-words or p-words having a unique interpretation from language to language or from description to description and the fact that grammatical and phonological principles might not line up to give the same results is not a problem. The linguist is free to discard certain grammatical and phonological principles as irrelevant to their identification of g-words and p-words according to the alignment of the stars, the flip of a coin, or the flippant suggestions of a more senior linguist. The misalignments described for Chácobo and Zapotec above pose no problem for the tautological fiat-based interpretation because the linguist is free to choose any of the competing p-words or g-words as constituting the 'real' instance of these categories according to how they feel, or perhaps according to precedence in their area of study ("Other Uto-Aztecanists/Zapotecanists/Arawakanists etc ... have defined it in this fashion and so I follow them".)

There is no problem, in principle, with the tautological word bisection dogma. It may even have expositional value in linguistic description and analysis. The expositional value of the fiat-based use of the notion of 'word' is expressed most clearly by Chao in his *Grammar of Spoken Chinese*.

[12]There could be a more empirically substantive notion of a test being poorly suited to a particular language. This could be defined as cases where a test is highly ambiguous providing a number of results.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> "Not every language has a kind of unit which behaves in most (not to speak of all) respects as does the unit called "word" when we talk or write *in* English about the subunits *of* English. It is therefore a matter of fiat and not a question of fact whether to apply the word "word" to a type of subunit in the Chinese sentence which has so many points in common with, and so few points divergent from, the English word "word" as to warrant the use of that term without danger of serious misunderstanding. As we shall see when we come to actual cases, we shall meet various types of word like units which can claim to be called the word, which overlap to a great extent, but which do not have quite the same scope. As usual, I shall prefer to use a familiar term, with a warning against making unwarranted inferences, in preference to using unfamiliar terms, which, though safe from being misunderstood, are often also safe from being understood." (Chao 2011: 159)

Thus, one can assign the label of 'word' as a matter of convenience since it could bootstrap understanding of an unfamiliar concept.[13]

But it follows as a matter of logic that one linguist's g-word and p-word will not necessarily be comparable to the next linguist's, even in the same language. There is also the danger that certain facts about the relevant language will remain poorly or imprecisely described. What would be the value in describing a potential diagnostic for g-words or p-words that does not line up with our preferred analysis (Haspelmath 2011) especially if authorities in the field *insist* that such constituents are manifested in all languages 'without doubt'?

On the empirical word bisection dogma, Dixon is making a substantive claim about regulative principles or constraints underlying the distribution of grammatical and phonological properties across the languages of the world. On this interpretation, Dixon is wrong to claim that grammatical and phonological words can be identified 'without doubt.' For this position to hold, Dixon would have to articulate how the grammatical and phonological principles he considers relevant would be patterned were the word bisection dogma false. All substantive empirical claims depend on a description language that allows them to articulate what it would mean for them to be falsified in order to show that they are not tautologies (Mayo 2018). This is what it means to have a substantive empirical claim.

[13] As a matter of descriptive convenience it is just as likely that the notion of 'word' obfuscates more than it clarifies and the purported understanding or agreement achieved is by and large an illusion. There is an important difference between a description *feeling* intelligible and having a detailed understanding of the case at hand as there is an important difference between agreement and the illusion of agreement (see Smaldino (2017), Kahneman et al. (2021) on the illusion of agreement).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

However, BLT has no vocabulary or descriptive language for even articulating the relevant counterfactual.

It is not always clear when a linguist is advocating a fiat-based or an empirically contentful conception of wordhood. Haspelmath (2023) is explicit in proposing a fiat based definition (not an empirically substantive theory) of 'word' for all languages. Certain passages in Dixon & Aikhenvald (2002), Dixon (2010b) and Aikhenvald et al. (2020) suggest that they are pushing an empirically contentful claim about the existence of "words" in languages:

> "It is not impossible that there would be a language that lacks phonological words and/or grammatical words, but we are not at present aware of one." (Dixon & Aikhenvald 2002: 32)

However, they do not articulate what such a hypothetical language would look like. It is hard to see from their discussion and their methodology how such a situation could arise, i.e. the claim appears to be tautological (Tallman 2020). As such all claims that insinuate that grammatical and phonological words are present in all languages in BLT as it is currently formulated are unfalsifiable and, therefore, ascientific. Insisting that all languages have grammatical and/or phonological words in the absence of any clear articulation of what the falseness of such a claim would entail empirically can only reflect a metaphysical prejudice rather than a scientifically valid position.

In any case it is interesting to consider what an empirical version of the word bisection dogma could amount to.

A strong version of the empirical word bisection dogma would claim that all phonological principles converge on a single domain and all grammatical principles converge on a single domain. However, this is clearly false and is well recognized as such by everyone who has discussed the topic to my knowledge (Carnie 2000, Hildebrandt 2007, Bickel et al. 2009, Bickel & Zúñiga 2017, Haspelmath 2011, Tallman & Auderset 2023). Such a claim would be implausible on diachronic grounds alone as we would expect grammaticalizing elements to gradually integrate into word domains over time (Bybee et al. 1998, Schiering 2006).

Another rendition of the empirical word bisection dogma is that it is probabilistic. This version of the claim seems to be presupposed in the following claim by Matthews:

> No [wordhood] criterion is either necessary or sufficient ... But they are relevant insofar as, in particular languages, they do tend to coincide (Matthews 2002: 274)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

One interpretation of this claim is that the g-domains and p-domains tend to converge around unique results more than one would expect if they were distributed according to chance alone. In this perspective the g-word and p-word are seen as regulatory principles that predict statistical clusterings of grammatical and phonological properties. We do not predict perfect coincidence between grammatical principles, nor between phonological ones, but enough to support the idea that grammar can be divided into word and phrase structure in the morphosyntactic and phonological domain.

In Bickel et al. (2009) this issue is engaged with, if not directly tested, in the phonological domain. Bickel et al. (2009) argue that the p-domains do not cluster around one abstract p-word domain cross-linguistically. Thus, on the interpretation that criteria should tend to cluster, it is not clear that Matthews' conjecture is correct. At least in the phonological domain the assumption seems to be falsified. As far as I know, Matthews claim about the tendency of wordhood criteria to cluster has not been tested systematically in the morphosyntactic domain. In Tallman 2021a I argue that it is not obviously true based on the application of wordhood tests in Chácobo.

This does not mean some version of the word bisection dogma as a regulative principle cannot be established when we look at the relevant phenomena cross-linguistically. This question is partly what motivated the collaborative project which resulted in this volume: is there an empirically contentful, but perhaps statistically justifiable version of the word bisection dogma that can be defended? Addressing this question requires a typological project that codes and measures the degree to which the relevant criteria align.

Insisting on a definition of the concepts by fiat may have some value in another research context (Haspelmath 2023), but it is not the concern of this volume.[14] We are concerned with describing and theorizing about patterns relevant to understanding identifiable empirical phenomena of the languages of the world, not with ratifying or rejecting some fetish in linguistics for "traditional" terminology.

[14]Note that Haspelmath 2023 provides a definition of a "word", which is not based on any phonological criteria. Those wishing to maintain a distinction between g-words and p-words might choose another phonological criterion to define the p-word. For instance, one could claim that the phonological word is always a minimality domain or *always* the stress domain. There may be some research contexts where such a universal definition is useful or even necessary. But it remains unclear why such a designation would invalidate a research program that seeks to investigate how different notions of the word, or different domains, cluster with one another cross-linguistically.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4 The prosodic hierarchy (hypothesis)

The Prosodic Hierarchy Hypothesis (PHH) is perhaps the most prominent hypothesis that is concerned with the relationship between morphosyntactic and phonological domains. The more orthodox articulations of the theory state that all languages come with a fixed number of (post-lexical) phonological layers (prosodic word, phonological phrase, utterance phrase etc.), which are projected (or mapped) from morphosyntactic constituency in a constrained fashion (Vogel 2023: 111). A corollary of this idea is that the relationship between morphosyntax and phonology is 'indirect': Morphosyntactic objects are translated into phonological ones where they can be interpreted by a phonological and/or phonetic component of grammar. The mapping process eliminates details from the morphosyntax from phonology's vantage point. This information reduction constrains the types of relationships that phonology can bear with morphosyntax. That is the idea anyway. In practice, the diversity in projection and parsing rules and the flexibility with which morphosyntactic and phonological domains can be constructed by the analyst makes the PHH (and associated auxiliary hypotheses) hard (or impossible) to test.

This section provides a brief overview of the PHH and the typological studies which have sought to test it. The methodology employed in this volume was inspired by the latter studies but sought to advance from them and overcome some of their shortcomings.

To illustrate the basic idea of the PHH and indirect reference consider the following sentences from Chácobo in 23 and 24.

(23) *kamano=́ ina pi=kɨ*
jaguar=ERG dog eat=DECL:PST
'The jaguar ate the dog.'

(24) *ina pii kamano=́ =wa=kɨ*
dog eat jaguar=ERG =TR=DECL:PST
'The jaguar ate the dog.'

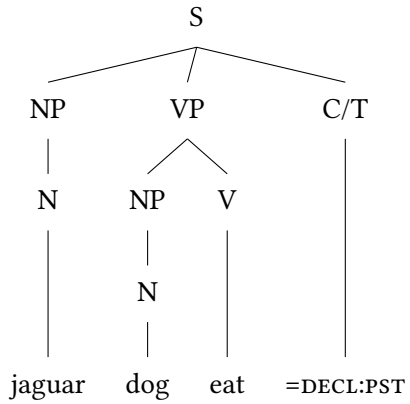
The sentences above serve to illustrate two facts about Chácobo. The displacement of the syntagm *ina pi* 'dog eat' from its position in 11 to its position in 24 suggests that the object and the verb root in Chácobo form a constituent excluding the clause-type and tense clitic *=kɨ* 'declarative past'. On the other hand comparison of the two examples shows that when *=kɨ* 'declarative past' is right-adjacent with the verb root as it is in 11 it behaves as a phonological constituent with the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

root, blocking the vowel lengthening manifest in 24. Assuming that the blocking of the vowel lengthening signals that *pi* and *=kɨ* are a phonological constituent in 11, we thereby arrive at an analysis where the an abstract syntactic structure motivated through constituency tests does not line up with phonological groupings based on minimality-induced processes (blocking, permitted or obliging the insertion of phonological material to meet a bimoraicity requirement).

We could posit the morphosyntactic structure for the Chácobo sentence in 11 with the translation rules in 26, resulting in the prosodic tree in 27. The structures below are simplified, only presenting constituency structures I discussed evidence for in the preceding paragraph (I assume that nouns and verbs are distinct and that Chácobo has noun and verb phrases, C/T stands for clause-type and/or tense, S stands for sentence).

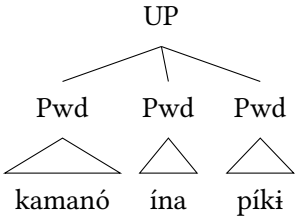
(25)



(26) Morphosyntactic to phonological constituency translation

a. Lexical X^0 elements project a phonological word
b. A lexical X^0 root parses nonlexical (clitic?) elements to its right into a phonological word (Pwd) if they are not already in a Pwd of their own (or clitic elements integrate into the prosodic word to their left)
c. Translate the highest projection into an intonational phrase (UP).

(27)

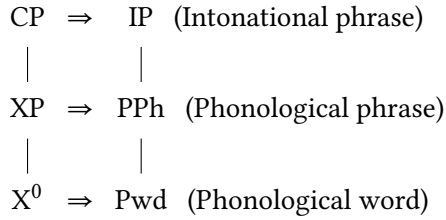


Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

We stipulate that if a Pwd is not minimally bimoraic, a root will undergo vowel lengthening. This captures the obligatory lengthening of *pi* 'eat' to *pii* 'eat' in the example in 24.

The analysis sketched above illustrates non-isomorphy between morphosyntactic and phonological domains: in the morphosyntactic analysis *pi=kɨ* is not a constituent, but in the phonological analysis it is. Or, put another way. Phonological rules of Chácobo rely on a (surface) constituent structure which is different from that which is motivated from morphosyntactic constituency tests. The analysis illustrates what is meant by 'indirect' reference: minimality is built out of *Pwd*, which is in turn parsed from abstract notions like lexical X^0.[15] Note that the translation from morphosyntax to phonology does not make reference to part of speech categories like 'noun' or morphosemantic content like PST. It only makes references to different layers of X and the distinction between lexical and functional categories. Typically lexical categories will project a phonological word but non-lexical categories will not (Selkirk 1996, Selkirk 2011, Werle 2009). The mapping rule also requires a morphosyntactic analysis with some type of division into levels for a correct formulation. If we gave Chácobo a different morphosyntactic structure by, for instance, assuming that *ina pi* 'dog eat' was under X^0 our parsing rules would no longer make the correct predictions. Thus, articulating one's morphosyntactic analysis is crucial for meaningful assessment of the predictions of any prosodic phonology analysis. If one does not present the evidence for X^0 or any of the presupposed constituency structures, the prosodic analysis will not make meaningful cross-linguistic predictions, nor be comparable to other prosodic analyses.

The Prosodic Hierarchy Hypothesis assumes that all languages manifest a universal prosodic hierarchy which is mapped from morphosyntactic constituency in a constrained fashion, depicted in 28.

(28)


Phonological processes make reference to phonological domains, not morphosyntax directly. A phonological rule that refers to morphosyntactic words or phrases

[15]The parsing rule provided more or less gives a "relational rule"; an edge-based formulation might say that Pwd is parsed from the left edge of lexical X^0.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

is banned. This requirement will not make an empirical difference unless the mapping rules result in non-isomorphy. Phonological domains are constructed out of structures such as X^0 and XP. They do not make reference to noun phrases or verb phrases as such. X'-theory or one of its descendants, which presumes that there is meaningful structural homogeneity across verbal, nominal and adjectival domains is presupposed. Indeed it is necessary for the translation process to occur. This prevents a phonological domain from being specific to a part of speech class or specific construction.

It is important to highlight what this perspective shares and what it does not share with the BLT formulation of morphosyntactic and phonological wordhood. Both the PHH and BLT assume that there is a hierarchy of constituents. Discussions of such issues generally presuppose that the identification of distinct and comparable levels cross-linguistically is somehow obvious: not much attention is given to the possibility that there might be some ambiguity in distinguishing between 'word' and 'phrase.' The PHH also assumes the word bisection dogma: that a distinction between morphosyntactic and phonological words is sufficient for describing misalignments between potential wordhood diagnostics. The PHH often comes coupled with a few other auxiliary positions, not explicitly articulated by BLT. For instance, BLT does not make explicit a distinction between lexical and postlexical phonology, but this is assumed in much prosodic phonology research (Scheer 2010). Relatedly, in most formulations of prosodic phonology, mapping rules do not make direct reference to information like part of speech classes. But this assumption is not made explicit in BLT. One wonders, however, whether such assumptions are *implicit* in the word bisection dogma. Does the notion of a phonological word really make sense if its content and/or relationship to morphosyntax varies from construction to construction, or part of speech category to part of speech category?

The PHH shares with BLT the adoption the word bisection dogma and presents us with a sophisticated set of labeling conventions for dealing with misalignments of the types $p_1 \neq p_2 \neq p_n$ and $g_1 \neq g_2 \neq g_n$. Misalignments in the morphosyntax can be handled by positing that the relevant g-domain is a phrasal, subphrasal or even a subword constituent. Despite the fact that misaligning domains can be dealt with a more elaborate set of labels, there is still unresolved ambiguity with respect to which domain receives which label, a point I elaborate on below.

The PHH purports to make substantive predictions about the relationship between morphosyntactic and phonological domains. It is often implied that there is wide scale empirical support for the hypothesis and that it makes substantive predictions about language structure (Bennett & Elfner 2019), i.e. it is not just a

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

set of arbitrary labeling conventions. Despite such triumphalist claims, it is not really accurate to discuss a single PHH. The empirical content of the hypothesis will vary drastically depending on what supporting auxiliary hypotheses are adopted and how one maps the metalanguage of the theory to language specific facts. Furthermore, the auxiliary hypotheses often weaken the predictions of the theory substantially. Below I take stock of these auxiliary hypotheses and assess their importance for the testability of (some version of) the PHH and the general usefulness of the PHH for language comparison. The first three points are well known and widely discussed and debated in the prosodic phonology literature: (i) adding more layers (section 4.1); (ii) skipping layers (Section 4.2); (iii) recursion (Section 4.3). The next two points concern issues which are less discussed, but further weaken the claims of the PHH (Sections 4.4 and 4.5). The final point concerns the most obvious empirical prediction of the PHH about domain clustering, which has shown to be false. More generally though, I argue that PHH is not strictly testable and therefore the idea that the PHH is well supported empirically is nonsensical. The best we can do is say that there are certain versions of the PHH that have been shown to be false. Furthermore, I argue that as a typological metalanguage for linguistic comparison, the PHH is problematic due to the ambiguity in mapping its categories and structures to actual languages. Linguists should move with caution when using concepts from the PHH for language comparison as the concepts are abstract and their mapping to language particulars indeterminate. I suggest that the planar-fractal method offers a better alternative for language comparison (for now).

4.1 More layers

While the three layers displayed in 28 are assumed by most researchers, the literature attests to a wide variety of positions regarding which other layers might be relevant. In Nespor 2007 a domain called the "clitic group" is posited to account for the behavior of combinations containing clitics between the prosodic word and the phonological phrase. The clitic group was abandoned when more sophisticated theories of clitic integration were developed in the 1990s (Booij 1996, Selkirk 1996, Peperkamp 1996, 1997). However, Vogel 2008 argues that such a constituent is still necessary, renaming it the "composite group". Downing & Kadenge 2020 adopts the "prosodic stem", a constituent lower than the prosodic word. Hildebrandt 2007 has shown that Limbu has too many domains to be able to be easily accounted for with the PHH. The possibility of adding (or removing) domains ad-hoc weakens the predictions of the PHH. At no point (except in the case of Schiering, Hildebrandt and Bickel) was the necessity of adding new

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

domains seen as evidence against the PHH, but the possibility immunizes the theory against a specific type of counter-evidence. Actual practice in the field suggests, therefore, that the PHH does not place any constraints on the *number* of phonological constituents a language might have. It is perhaps true that the PHH could make some claim to the number of phonological layers that any given language *tends* to have, but this has not been shown.

From the perspective of language comparison the possibility of adding new domains adds more ambiguity. Consider the case of adding the composite group or "kappa" to our vocabulary (Miller 2018, Vogel 2019). Now for a given p-domain in a language where the kappa was not originally introduced, we are not just faced with potential ambiguity between p-word and p-phrase, but also between p-word, p-phrase and kappa. This problem could only in principle be resolved with attention to cross-linguistically identifiable morphosyntactic domains: kappa or whatever should relate to a kappa specific morphosyntactic domain in a specific way. Otherwise the extra domain has no mnemonic value for language comparison and introduces noise in language comparison. How are we to know that one linguists' kappa is not another linguist's phonological word or phonological phrase?

4.2 Layer skipping

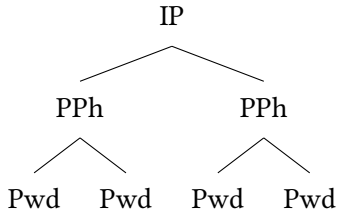
The original version of the PHH posited 'strict layering.' An analysis that follows Strict layering is one where in the parsing of elements into the prosodic hierarchy none of the layers can be skipped (Hayes 1989, Selkirk 1996, Nespor 2007). I quote Selkirk for a more precise definition.

(29) The strict layer hypothesis
A constituent of category-level *n* in the prosodic hierarchy immediately dominates only a (sequence of) constituents at category-level *n*-1 in the hierarchy (Selkirk 1984: 437)

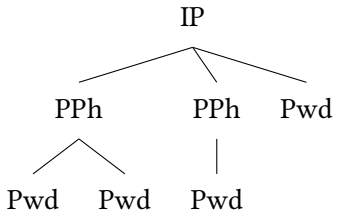
A prosodic word can only be composed of feet. A phonological phrase can only be composed of prosodic words. It cannot contain prosodic words and syllables. This hypothesis constrains the structure of phonological constituency. There are two ways of violating the strict layer hypothesis. One is through layer-skipping and the other is through recursion. I start with layer skipping. A structure without layer skipping would be as in 30a and one with layer skipping would be as in 30b. The right-most prosodic word in this tree 'skips' the phonological phrase.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(30) a.



b.



To illustrate the basic idea of layer skipping, consider the example from Chácobo below. There is a phonological domain in Chácobo where high tones are inserted if there is no underlying lexical L(ow)-H(igh) tone present.

(31) [nŏjaki ↓]
nŏya =kɨ́
fly =DECL:PST
'S/he flew'

When an underlying LH tone is present as in the example below, the high tone insertion is blocked. An H is not inserted on the first syllable as in the previous example.

(32) [nòjàjóki ↓]
noya =yŏ *=ki*
fly =COMPL =DECL:PST
'They all flew'

The domain of initial H tone insertion/blocking is larger than the minimality domain that I identified as the PWd above (Tallman 2018). I thus assume it is the PPh, following the assumptions of the prosodic hierarchy.[16]

In the example in 33, the clitic *=kɨ̌* 'prior event, different subject' (not to be confused with toneless *=kɨ* 'declarative, past') blocks the insertion of the H tone as expected on *noya* 'fly' as expected.

[16]But I could call it the 'composite group.'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

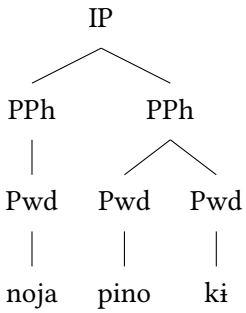
(33) [píno nojàkɨ́ tsí honi tsájakɨ ↓]
pino *noya =kɨ̌* *tsi* *honi =* ´ *tsaya =kɨ*
humming.bird fly =PRIOR:DS LNK man =ERG see=DECL:PST
'When the humming bird flew the man saw it.'

In different subject dependent clauses, verb phrases can front as in the example below, where *noya* 'fly' appears before the subject *pino* 'hummingbird.' Note that in this example, the H tone is inserted on *noya* 'fly.'

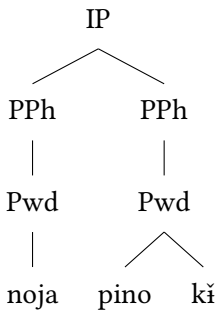
(34) [nója píno kɨ̌ tsí ɨ tsájakɨ ↓]
noya pino *=kɨ* *tsi* *honi =* ´ *tsaja* *=kɨ* '
fly humming.bird =PRIOR:DS LNK man =ERG see =DECL:PST
What the humming bird did was fly when the man saw it.'

In the string *noya pino kɨ̌* two possibilities are warranted under strict layering. Either, the *=kɨ* must integrate into the Pwd projected from *pino* or it must itself project its own Pwd. The two possibilities are depicted below (excluding an analysis whereby the clitic projects its own PPh, which wouldn't solve anything in any case).

(35)



(36)

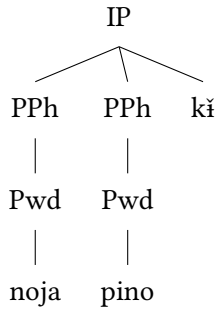


Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Neither analysis makes correct predictions. If we assume that *=kǐ* 'prior, different subject' projects its own Pwd, then it should lengthen to meet minimality requirements. Even if we allow it to integrate into an adjacent Pwd (for which there is no evidence based on vowel lengthening), its presence should block the insertion of an H tone on *pino* 'hummingbird', contrary to fact.

Our prosodic analysis can be saved from quick falsification, if we allow *=kǐ* to integrate post-lexically with a higher prosodic domain, say IP, depicted in the tree below. This involves 'skipping' both the Pwd and the PPh layer.

(37)



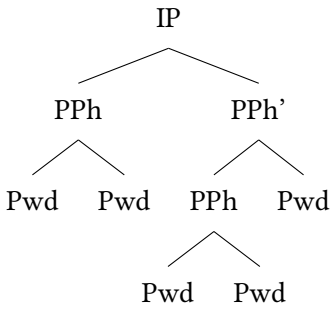
Violating strict layering makes the PHH weaker as it immunizes the theory further against potentially falsifying evidence, bringing it closer to the status of a tautology, i.e. a set of labels for annotating phonological domains and nothing else. As far as I have been able to discern the ability for clitics to integrate at various levels of the prosodic hierarchy does nothing except redescribe their phonological behavior in a stipulative fashion. Insofar as this interpretation is correct, layer skipping exonerates the PHH from making any predictions about clitic phenomena cross-linguistically. While it may be an elegant expositional device for representing language-internal and cross-linguistic differences in the behavior of clitics (Peperkamp 1996), it should be recognized as just that, not a theory that constrains language variation.

4.3 Recursion

As stated above, in the original PHH, *strict-layering* prevented individual prosodic domains from recursing. An example of a recursive structure in prosodic phonology would be as follows. In the structure below PPh' is a recursed PPh of the lower domain.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(38)



The issue of whether recursive structures exist in phonology is somewhat controversial (Féry 2017, Tallman 2021c, Ishihara & Myrberg 2023, Kügler 2023, Bögel 2021, Cheng 2021, Ito 2021, Miller & Sande 2021). The reason seems to be related to the fact that different authors adopt different criteria for identifying recursion. Here I will limit the discussion to how the issue of recursive phonological domains is relevant for language comparison (see Miller & Sande 2021 for an important discussion about how recursion might be constrained cross-linguistically).

An important first cut in understanding recursion in phonology would be to recognize a distinction between notational and empirical recursion. The distinction is inspired by the discussions in Schiering et al. 2010 and Miller & Sande 2021.

(39) a. **Notational recursion**: A category is embedded under another category with the same label. The different instances of the label need not have the same empirical signal (i.e. they do not refer to identical empirical phenomena).

b. **Empirical recursion**: A category is embedded under another category. Each layer signals the same empirical phenomenon.

In notational recursion one label is just formally represented as embedded under another one. I can illustrate notational recursion with an example from Chácobo. In Chácobo, I associated minimality with Pwd, default H tone insertion with PPh and intonational phrasing with IP. There is an important prosodic phenomenon in Chácobo whose span of structure is in between that of the PPh and the IP. Trisyllabic nouns truncate their final syllable if they occur before the clause-type morpheme. Otherwise they occur in their 'long forms'. The long form of the morpheme *kàmáno* 'jaguar' is illustrated in 40 and the short form in 41. Likewise the short and long forms of *tsăkaka tsăka* are provided in these examples.

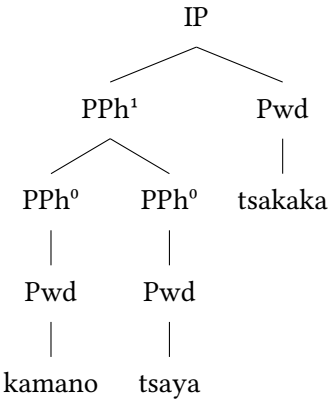
Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(40) [tsǎkà tsájaki̵ kàmáno ↓]
tsǎkaka tsǎya =ki̵ kamǎno
agouti see =ANT jaguar
'The jaguar has seen the agouti.'

(41) [kàmá tsájaki̵ tsákaka ↓]
kamáno tsaya =ki̵ tsákaka
jaguar see =ANT agouti
'The agouti has seen the jaguar.'

Rather than positing a new domain for noun and adjective apocope, I can assume that the PPh recurses. The lower PPh^0 is relevant to H tone insertion and blocking and the higher PPh^1 is the domain where trisyllabic or larger nouns and adjectives truncate their final syllable.[17]

(42)



Thus truncation only occurs in PPh^1. It should be obvious, however, that this is no different empirically from just positing an extra layer. The only difference is that using notational recursion introduces labeling ambiguity (a point made *en passant* by Féry (2017: 62) and Richards (2016: 97) without discussion of the resulting epistemic problems this ambiguity entails): once recursion is admitted there isn't a clear reason why we should not label PPh^1 as UP^0, and the original UP as UP^1, shifting the burden of recursion to another domain. For the purposes of language comparison I, therefore, cannot see any advantage to using notational recursion.

[17] I could even justify this decision based on a syntactic analysis whereby both PPh^0 and PPh^1 are mapped from XPs as in Match Theory (Selkirk 2011).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

On theoretical grounds, the adoption of notational recursion weakens the prosodic hierarchy for the same reason that adding new domains does. Without further constraints, the effect of adding recursion into the categories of the prosodic hierarchy seems to mean that this theory now puts no upper bound on the number of prosodic domains it allows (Tallman 2021c). It is not clear to me what the purpose is of advocating notational recursion over just adding extra domains.

There might be examples of real empirical recursion (see Féry 2017 for a review). For instance, let's say that inside PPh^1, a super H tone with twice the distance in semitones from L tones was inserted on the first syllable of the domain. One could argue that the relevant phonetic effects have now been stacked in proportion to how embedded the domain is, but that the phonetic properties of the domain have remain unchanged. Something like this might be true for the prosodic behavior of some embedded clauses as they can display similar prosodic properties but with phonetic differences shrunk down (Vigário 2010).

4.4 Empirically contentless layers

Nespor (2007: 11) argues that if one does not find evidence for a given layer of the prosodic hierarchy one is not necessarily warranted in assuming that the layer is not present. While the layer may not be causally related to a specific phonological process or phonetic effect, stipulating its presence may help formulate rules for other prosodic domains. An example might be positing CVV syllables in Araona. While it is not strictly necessary to state the stress rule/pitch accent rule of the language (for this all you need are vowels and consonants), positing a syllabification rule makes the statement of the stress rule simpler (Tallman & Gallinate accepted). In this case perhaps one is warranted in positing the syllabification rule and syllables as a prosodic layer in the language. We might also find a pitch accent rule in a language which applies as PPh, which inserts a pitch accent on the leftmost prosodic word in the PPh. The prosodic word itself might not have any independent phonological processes, but assuming prosodic words are present helps in articulating the phonological phrase.

Nespor (2007: 11-12) seem to take the idea of empirically empty layers even further, however, suggesting a strong burden of proof for positing the absence of one of the domains of the prosodic hierarchy in a given language, thus letting the PHH off the hook again, this time in terms of making any predictions regarding the minimum number of phonological layers one needs to ratify the theory.

> "If ... it turned out to be the case that all of the languages that appeared not to have phonological rules that refer to X^i shared some other feature as well,

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> this would be a more convincing type of evidence that X^i may be absent in a particular category of languages characterized by this feature."

The interesting empirical question raised by this point notwithstanding, it should be noted that the epistemic consequences of allowing categories of the prosodic hierarchy to be empirically invisible makes the PHH even weaker as a theory: cases where no phonological rule can be found for a Pwd domain in a given language predicted to exist because of the prosodic hierarchy cannot be regarded as counter-evidence.

The suggestion that some languages may have little or no [!?] empirical phenomena which are causally related to their prosodic words has been taken up by Féry (2017) Such languages are referred to as 'phrase languages' (they include Hindi, Georgian, Turkish among others).

> "... tonal specifications are mostly assigned at the level of Φ - phrases and ι-phrases. But contrary to intonation languages, specifications at the level of the word are sparse, absent or only weakly implemented. Phrase languages do not automatically associate pitch accents with stressed syllables, most tones are nonlexical (or 'postlexical')." (Féry 2017: 270)

This position begs the question as to when one is ever justified in questioning the universality of a specific domain according to prosodic phonologists, since the criterion of finding something in common in such languages is at least suggested by Féry. Note that the position seems to differ from Nespor 2007. Féry finds evidence that the languages where no p-word is present have something in common, but assumes that the p-word is there anyways.

The analytic possibility of positing empirically contentless layers potentially adds more indeterminacy for language comparison. Instead of positing that a given Pwd has little or no empirical signal the question arises as to whether the PPh should be relabeled as the Pwd. This is a general problem when the number of phonological domains is smaller than the set predicted to exist from the prosodic hierarchy (Tallman 2020). One linguist's Pwd might be another's PPh, for instance (see Michaud 2017: 321-322 for relevant commentary).

Domain labeling ambiguity arises as a consequence of a lower number of prosodic domains when we only consider the prosodic tree geometry without considering the structural relations between the prosodic tree and analogous morphosyntactic domains. A prosodic word is not (just) the domain between the foot and the phonological phrase, but also the domain which is structurally closest in some sense to the morphosyntactic word.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

4.5 But what morphosyntactic structure?

The validity of using structural closeness to morphosyntactic domains to label prosodic ones, depends on those morphosyntactic domains also being consistently definable from language to language (Miller 2018). That is, in case after case, the identification of and the distinction between X^0, XP and other constituents has to be made consistently. However, in general, the prosodic phonology literature rarely discusses morphosyntactic criteria. For instance, in Féry 2017 only a single criterion is provided for morphosyntactic wordhood (coordination), and as far as I could discern no literature is cited that helps the reader discern how to parse up morphosyntactic constituents in a way that makes predictions about the morphosyntax-phonology interface operationalizable cross-linguistically.

Exacerbating the problem, the morphosyntactic literature is not obviously unified in its prescriptions for how one should go about identifying the relevant constituents or if the morphosyntactic constituents presupposed by the PHH are even valid at all. Carnie (2000), for instance, argues that the there really is no discrete distinction between X^0 and XP. The reasoning behind this is that the properties associated with (head moving) X^0 constituents and (A/A'-moving) XPs do not perfectly cluster. Similar problems have been discussed outside of the generative literature (Russell 1999, Haspelmath 2011, Bickel & Zúñiga 2017, Tallman 2021a). There are constituents that behave like X^0s according to some criteria and like XPs according to others.

We can add that part of controversy about direct versus indirect reference theories relates to what the correct morphosyntactic analysis is, as morphosyntax-phonology non-isomorphisms could be the result of an incorrect analysis of the morphosyntax (Seidl 2001). The possibility that non-isomorphisms might be the result of unmotivated analyses in the morphosyntax, was also highlighted in the usage-based literature (Bybee 1999). Nonchalance about the labeling of *morphosyntactic* domains, not to mention how to motivate the correct constituency structure, is, therefore, not justified for linguists interested in testing and/or developing theories about the relationship between morphosyntax and phonology.

4.6 Clustering hypothesis

One prediction of the PHH is domain clustering or bundling (Bennett & Elfner 2019). This is the only claim of the prosodic hierarchy that has been tested in a typological study.

Bickel et al. (2009), Schiering et al. (2012) developed a word-domain database. This database coded phonological processes in 70 typologically diverse languages.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

It taxonomized the phonological processes that define p-domains into a number of types (e.g. metrical based, harmony, segmental). Each domain could be coded as being mapped over a set of structural categories (e.g. prefix-root vs. prefix-root-suffix). The relative clustering of domains could then be assessed cross-linguistically. The structure of the database allowed the researchers to assess a number of statistical relationships between phonological domains: (i) which phonological processes tend to occur in "higher" or "lower" domains than others; (ii) which phonological processes tend to cluster together in terms of span of application; (iii) whether there an overall tendency for domains to cluster or bundle together better than one might expect.

An answer to the last question is most relevant to the claims about the prosodic hierarchy. Using multidimensional scaling Bickel et al. (2009) argue that there is no tendency for the phonological domains of their study to cluster. They argue that this result refutes the claims of the Prosodic Hierarchy Hypothesis. The idea is that if the PHH were correct, we would expect prosodic domains to cluster around a single formal category, but they do not evince any tendency to do so. In another publication taking a close look at Thai and Limbu (Schiering et al. 2010), two languages which present challenges to the PHH in that they do not have the right number of layers, the authors suggest that the reason for the observed non-clustering is that prosodic domains are 'emergent.' There is no set of innate formal categories constraining the distribution of phonological domains, these emerge from language history. The studies by Réné Schiering, Kristine Hildebrandt and Balthasar Bickel were the first to systematically investigate the issue of domain clustering. In certain aspects the methodology employed by these authors overcomes many of the epistemic difficulties associated with the prosodic hierarchy I discussed above. The methodology employed in the current study builds on Schiering and company's methodology in important respects. We try to overcome some of the shortcomings of their approach, and so these shortcomings are worth commenting on.

The first shortcoming is that the project focused only on 'word-domains', rather than assessing the relationship of phonological domains from morph to utterance (or at least prosodic word to utterance phrase). This opens the research up to criticisms that *perhaps* some of the domain misalignments could be related to the fact that some of these domains are 'phonological phrases' (or higher domains). If there is no consistent way of distinguishing between prosodic words and phonological phrases based on phonological criteria, then it becomes unclear why some of these p-word domains are not actually indicating a higher level of structure.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

As far as I understand, the identification of p-domains in Schiering et al. 2012, Bickel et al. 2009 were limited to 'lexical' phonological processes. This issue was not explicitly discussed in the published materials to my knowledge and thus my comments here should be taken with a grain of salt (see Hildebrandt This Volume for clarification). Lexical phonological processes are supposed to be different from post-lexical processes based on a number of properties: structure preservation, optionality, reference to morphosemantic information, categoriality, among other properties (Zsiga 2020: 201). Lexical phonological processes are also supposed to be word-internal. A phonological process is structure preserving if it involves changing one contrastive phonological unit to another. For instance, vowel tensing in English is structure preserving: the change of *grain* /gren/ to /græn/ in the context of the form /grænular/ is structure preserving because /e/ and /æ/ contrast in English. Such a process would be considered 'lexical.' Flapping in English, which only results in the introduction of a noncontrastive allomorphy [ɾ] is considered post-lexical. Schiering et al. (2010) only focused on lexical phonology.

A problem arises when we consider the fact that the criteria for distinguishing between lexical and post-lexical processes do not themselves cluster together. For instance, the morphophonetics literature has shown that there are many word-internal processes which are not structure preserving (Plag 2014). Some research has also uncovered structure preserving processes that are 'post-lexical' in the sense that they occur at phrase level domains (Hyman 1993). Bybee (2001: 214) points out that the distinction between lexical and post-lexical is probably graded, rather than discrete. It is not clear, therefore, that a distinction between lexical and post-lexical phonology can serve as a basis for a 'word domains project.' We seem to be forced by the empirical phenomena to look at the whole picture without presupposing that phonological processes can be divided neatly into lexical and post-lexical categories.

On theoretical grounds, focusing only on the word domain means that the Prosodic Hierarchy Hypothesis cannot be systematically engaged with. In many current formulations the PHH, it is *only* concerned with post-lexical processes. Lexical phonological processes are handled by lexical phonology, where there is no expectation of domain convergence. Rather, layering and cyclicity is all that is expected in the word. Contrary to the assumptions that are made in Bickel et al. (2009), clustering of phonological processes around a single domain is not.

Another criticism of the word domains project is that it did not explicitly engage with morphosyntactic information. This criticism is present in Miller & Sande (2021), for example. The PHH is not just a theory about the clustering of phonological domains. It also purports to be a theory which constrains the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

relationship between morphosyntax and phonology. Miller's criticisms suggest that one should not conduct a typological project of phonological word domains without also including morphosyntactic information. Insofar as Miller's criticisms are meant as a defense of PHH (rather than simply a critique of the word-domains project) they are somewhat weak, however, because prosodic phonology literature suffers from a general dearth of argumentation for its presupposed morphosyntactic analyses even where it posits abstract morphosyntactic structures.[18]

Another critique of the word-domains project is that it did not present an alternative theory which meaningfully constrains the distribution of p-domains cross-linguistically. The force of Bickel (2015), Schiering et al. (2012, 2010) is largely methodological. They argue that typological research should start from language specific processes rather than positing *a priori* structures. Such an approach seems necessary if we are going to hope to test competing claims about prosodic phonology. Schiering et al. (2010) also suggest that their results support an emergentist approach to phonological domains: "This leads us to conclude that the prosodic word is a language-particular category which emerges through frequent reference of phonological patterns to a given morphological construction type." (Schiering et al. 2010: 705). The argument seems to be based largely on the failure of formal theories to account for linguistic variation, rather than the development of a testable emergentist theory of prosodic domains (see Mielke (2008) for discussion). Future research should be dedicated to fleshing out an empirically contentful emergentist alternative. If this is done we will be able to actually assess how much formal innate structure is really necessary, if any (Schmidtke-Bode & Grossman 2019).

5 Typological description languages, falsifiable theories and selection bias

The previous discussion has suggested that prosodic phonology suffers from two serious problems. First, insofar as it professes to be a theory about language structure it suffers from a lack of falsifiability.

Second, insofar as it might serve a function for language comparison it is also problematic: the theory posits a repertoire of formal categories and structures,

[18]It is my understanding that morphosyntactic information (g-domains) were included in the original autotype database. However, it is not obvious to me how the morphosyntactic domains related to the findings reported in published materials.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

but the mapping between these and language specific facts is highly underdetermined, resulting in a lack of commensurability from description to description.

The planar-fractal method seeks to be a typological description language in the spirit of Schiering et al. (2010) and Good (2016a) that addresses these issues. This means that it is a method for comparing structures from language to language. It does not seek to be a theory which constrains typological variation. But it can serve as a methodology for testing or developing such theories. It is developed in such a way that it can be used to create machine-readable databases. This will allow researchers to discover statistical trends in the relationship between morphosyntactic and phonological domains.

Some researchers find this strange because they assume all formal frameworks for describing linguistic facts should necessarily be theories about typological variation or the nature of language, or language universals or whatever. However, developing a description language for stating facts independently of a theory is necessary to assess the relative merits of competing theories and to avoid lapsing into self-sealing tautologies in theory construction. Relatedly, some have criticized generative linguistics specifically for conflating "theory" with "notation" or "metalanguage" (Dryer 2006). While it is true that the planar-fractal method makes certain assumptions about language structure and assumptions about what data are important via its notation, distinguishing between data structures and theoretical models is crucial in all the sciences. Data structures are useful because they allow us to state or even simulate explicitly what data patterns we would observe if our theories were false or true. Distinguishing data structures from our theories allows us to actually assess whether a theory is testable (Mayo 2018).

The planar-fractal method does not compete with the PHH or any other prosodic theory for status as a theory. However, as a description language for coding, testing and developing theories concerning the relationship between morphosyntax and phonology it is superior. It attempts to eliminate mapping ambiguity between language specific facts and language structures (e.g. is phonological domain *x* a p-word or a p-phrase?) and code cross-linguistic data in a machine-readable database (See Auderset et al. This Volume for the database structure). It does not posit *a priori* structures presupposed by certain theories. Rather it is designed in such a way that it could be used to test such theories and/or their auxiliary hypotheses. In this way it functions as a 'comparative concept' (Haspelmath 2010, Good 2016b) allowing constituency facts to be coded in a commensurate fashion from language to language.

One of the motivations for conducting a cross-linguistic study and developing a methodology such as the one used in this volume is to overcome certain

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

methodological flaws in traditional linguistic analysis. In particular some recent literature has pointed out such a flaw in syntactic, if not just general linguistic, argumentation. The methodological flaw is referred to as 'methodological opportunism' or 'diagnostic fishing' (Croft 2001b, 2010, Haspelmath 2011). In essence, the idea behind this criticism is that in certain cases, linguistic frameworks, theories or hypotheses are coupled with a methodology that allows (or perhaps impels) researchers to discard or ignore data that might contradict a preferred hypotheses. Croft 2001b has argued that one of the reasons that there are so many competing syntactic theories is because researchers are simply using different data to construct their analyses throwing out or dismissing as irrelevant the data used by their competitors. Haspelmath 2011 applied a similar criticism to the literature that makes use of some notion of 'word.' Because there is no jointly agreed upon set of wordhood criteria, criteria can be used to because they fit a preferred analysis or discarded if they do not.[19]

In general terms, biases of this kind are well-known outside of linguistics, especially in discussions about replicability and hypothesis testing (Risen & Gilovich 2006, Nosek et al. 2018, Mayo 2018). More closely inspired by the latter literature, I refer to the problem as 'selection bias' (Tallman 2021b). The solution to selection bias that I propose below is called 'full reporting.' Rather than every linguist pulling criteria in an 'opportunistic' fashion from the literature and interpreting the criteria *just so* they fit with their preferred analysis, **full reporting** means applying constituency diagnostics according to a protocol, developed by a team of researchers working on different languages. The idea is that full reporting forces the linguist to be held accountable to constituency diagnostics they might not have used otherwise. With this methodology we hope to assess claims about domain clustering in a less biased fashion, because we are not as beholden to the implicit biases of individual linguists working in isolation.

6 Planar structures

A planar structure is a hypothesis space for coding the results of constituency tests or domains, phonological and morphosyntactic alike. It is a 'comparative concept' in the sense that Good 2016b uses the term in his discussion of templates. A planar structure is a maximally flat structure that contains **positions**

[19]While the criticism seems to be directed at generative linguistics, it is not clear why the same criticism does not apply to other theory-driven endeavors in linguistics. BLT, for instance, presupposes a distinction between phonological and morphosyntactic words: why does the methodological critique of generative linguistics not extend to this approach as well (see Section 3)?

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

which are **fit out** by **elements**. The positions are ordered into a template. The planar structure is an extension of the coding methods developed by Bickel et al. (2009) and Bickel & Zúñiga (2017). Unlike the structures of the latter sources, however, it is not delimited by orthographic word boundaries. Rather it contains syntagmatically distinct positions where presupposed morphological and syntactic elements are positioned on the same 'plane' with a caveat: languages have a planar structure for each part of speech distinction they contain. A verbal planar structure contains positions within a presupposed verbal word, 'free' adverbials, and other syntactic elements and **noun phrases** (nominal planar structures) all in the same template. A nominal template will have the noun root, all affixes which can combine with a noun and any syntactic noun modifiers.

In this section, I describe the planar structure by comparing it to phrase structure analyses. First, I provide a conceptual introduction to planar structures by articulating them as 'flattened out' phrase structure grammars in Section 6.1. Then I provide a more precise formal sketch of planar structure grammars in Section 6.2 describing them as a species of phrase structure grammar with more rigid conditions on what constitutes an admissible non-terminal node. Section 6.3 discusses the 'tangling' of different planar structures, referring to cases where modifiers of one domain (predication, reference) appear in another and how this is handled. Another constraint on planar structures is that they contain a 'base' element which is fixed in place in the template, a condition not put on non-base elements (Section 6.4). Planar-structures analyze elements into positions and elements are analyzed into minimal morphs where possible and larger structures where necessary. The minimal morph condition is discussed in Section 6.5. Finally I briefly comment on a criticism of the methodology that has arisen through its presentation in Section 6.6.

6.1 Flattening phrase structure grammar

In order to explain the planar-structure I will compare it with a typical phrase structure grammar. To start off I point out that a planar structure could be viewed as a phrase structure grammar which is 'flattened out' until issues of recursion make the flattening unworkable. This is not done because of a commitment to the idea that sentence structure is non-hierarchical. *Rather* it is done in order to construct a template over which constituency test results can be coded in a commensurate fashion across languages. Furthermore, the formalism gives us the possibility of coding bracketing paradoxes in a given language, which are not

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

straightforwardly supported in phrase structure grammars.[20] I will emphasize throughout that the planar-structure is not meant to compete with any given phrase structure grammar as a tool for the development of testing of linguistic hypotheses. My view is that they should complement them. The planar structure is a *cross-linguistic comparison tool* and *constituency test or domain measuring device*, not a hypothetico-deductive model.

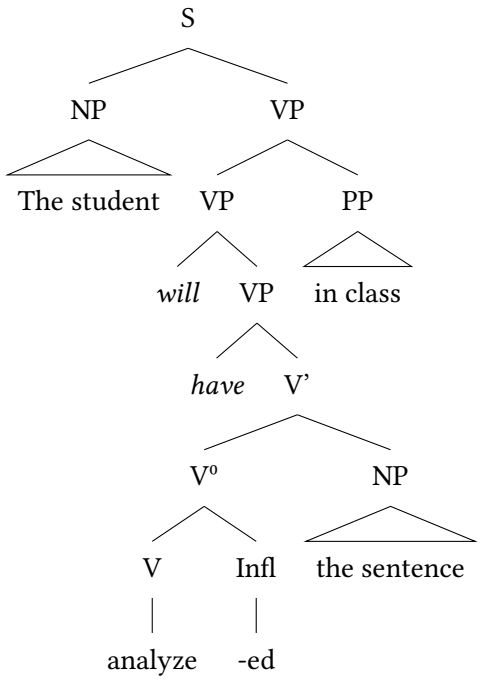
The idea of 'flattening out' a constituency structure should be intuitive for linguists who are familiar with competing syntactic theories where more or less hierarchical analyses can be contrasted with more or less flatter constituency ones (e.g. Culicover & Jackendoff 2005, Sobin 2008 for discussion). Consider the English sentence *The student will have analyzed the sentence in class.* A fairly standard constituency analysis might posit the phrase structure rules in 43a, with the corresponding constituency analysis in 43b (see McCawley 1988: 207-261 Baker 1995 for rough equivalents in terms of the degree of hierarchical structure).

(43) a. S ⟶ NP VP
VP ⟶ VP PP
VP ⟶ *will* VP/V'
VP ⟶ *have* VP/V'
V' ⟶ V^{0} NP
V^{0} ⟶ V Infl

[20] One has to posit multiple phrase structure grammars (Sadock 1991) or toss out certain test results.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b.



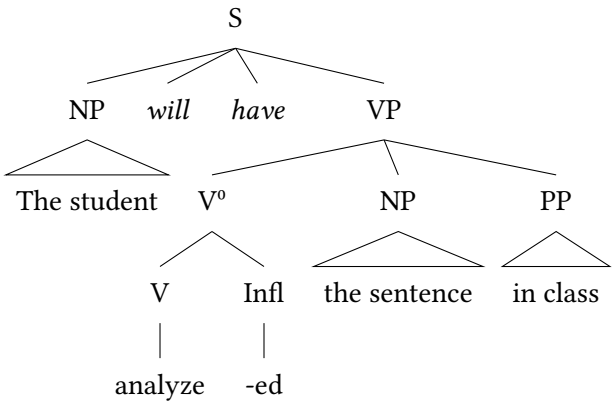
A few arguments can be rallied in favor of the layered VP structure above (e.g. V'-deletion (McCawley 1988: 210), affix-hopping combined with X'-theory (Ouhalla 1999: 95–99). *Do-so* proform replacement or perhaps considerations of scope can be rallied to support the idea that the prepositional phrase *in class* requires an additional VP-layer (Sobin 2008).

Another analysis might 'flatten out' the structure on the grounds that the evidence for the layered VP above is weak and/or problematic for a variety of reasons (Culicover & Jackendoff 2005). We might posit a flatter structure as in 44a.

(44) a. S ⟶ NP *will have* VP
VP ⟶V⁰ NP PP
V⁰ ⟶ V Infl

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b.

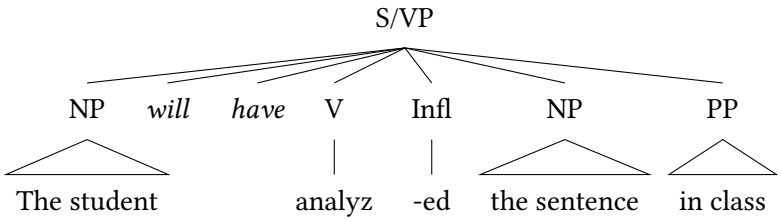


And this is as far as any linguist would go with English in terms of 'flattness' (to my knowledge) if English was our primary consideration. For typological investigation though we want a representation that allows us to code constituency tests regardless of whether these support a specific constituency analysis. It is at this point that an important conceptual difference between phrase structure grammars and planar structures arises. We are interested in phrase structure grammars only insofar as they give as position classes over which we can state test results. We are not interested in an elegant account of English grammar but one which allows unbiased comparison of constituency tests with other languages unmediated by the abstract ('chimerical', see below) constituents posited in phrase structure grammars. I fact, our goal is to represent all languages *as if* they had the same degree of structural flatness so that we can assess how constituency tests support hierarchical structures to different degrees across languages.

In order to do this we flatten the structure further as illustrated in 45 below.

(45) a. S/VP ⟶ NP *will have* V Infl NP PP

b.



Constituency tests and constituency test fracturing are discussed in Section 7, but the relationship between a constituency test and a planar structure needs to

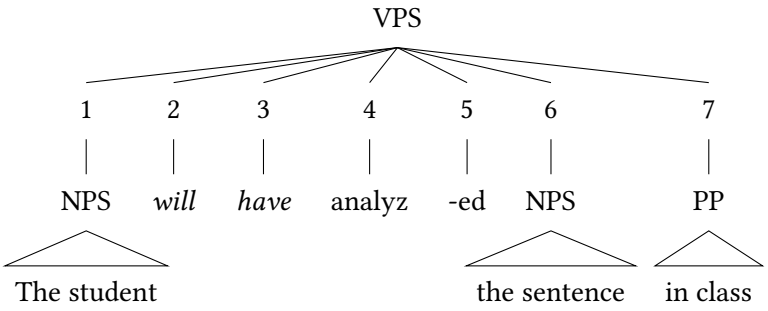
Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

be introduced to understand the next step in explaining the motivation for our representational device. Putting aside the noun phrase and prepositional phrase, the structure posited above represents nothing except the relative ordering of elements in the verbal word and/or the verb phrase with its functional projects or modifiers. To discuss constituency test results, we will refer to *spans of structure* identified by these tests and attempt to define them in a consistent way cross-linguistically.

Let us say we want to code the result of a *do-so* proform test in English. We could say that the test identifies a span of structure [V...NP][21] and a span of structure [V...PP][22] over the template defined by the phrase structure rule in 45. Such a notation will quickly get out of hand and become ambiguous with more complex structures, however.

We, therefore, take our flattened out representation and add consecutive numbers over the positions classes. As in the example in 46, where *VPS* stands for 'verbal planar structure' and NPS stands for 'nominal planar structure.'

(46)



The relevant phrase structure rules would be as follows. The first rule giving the verbal planar structure and the other rule giving the structural positions and the elements that can fit them out.

(47) VPS ⟶ 1 2 3 4 5 6 7
1 ⟶ NPS
2 ⟶ *will*
3 ⟶ *have*
4 ⟶ analyze

[21] As in *The student will have* [*analyzed the sentence*]$_i$ *in class and his teacher will have* [*done so*]$_i$ *too in his office.*

[22] As in *The student will have* [*analyzed the sentence in class*]$_i$ *and his teacher will have* [*done so*]$_i$ *too.*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

5 ⟶ *ed*
6 ⟶ NPS
7 ⟶ [P NPS]

We can now say that the *do-so test* in English identifies a 4-6 and a 4-7 span. We use a flat template in order to state the constituency tests that motivate our constituency analysis. If we report only those tests that allow us to motivate the constituency analysis we consider valid, a phrase structure grammar and the planar structure with constituency tests would be notational variants of one another. But the planar structure allows us to approach the question with more agnosticism. We can state and code the results of tests which we are unsure about (i.e. unsure if they are meaningful) and we can more easily state which groupings (spans) have more or less support.

At this point the reader might wonder whether the planar-fractal method provides nothing except an awkward notational variant of constituency analyses which allow overlapping constituency structures (Sadock 1980, 1991). This impression would be legitimate if we stopped short of developing the method for cross-linguistic comparison.

For typological comparison there is an important difference between providing a phrase structure grammar which manifests a particular constituency analysis that implicitly codes some set of constituency tests and a planar structure which allows for explicit coding of those constituency test results. The former is mediated by abstract constituent categories such as V^0, VP, word, phrase etc., the latter is only mediated by a notion of verb/predicate, a notion of noun/referential expression, and (perhaps) a notion of adjective/modifier. Apart from this, the planar structure coupled with reported spans is only mediated by structural positions (which is also true of phrase structure grammars in any case).

The problem with abstract constituents for typological comparison is that they can stand in for groupings that are based on an open-ended set of constituency diagnostics and linguists can differ in terms of which of these constituency tests they think ought to be captured by the phrase structure representation. This can lead to obfuscation of empirical differences and similarities in constituents or domains across languages. A VP in one case might not mean the same thing as a VP in another. We can consider the English case and compare it with Chácobo. I stated that the VP containing an object NP and a V in Chácobo could be motivated by displacement in Section 4. There are no verbal proform tests that provides evidence for a verb and object constituent which excludes the subject.[23].

[23]The translation equivalent *toka ... a...* 'do so' can replace a verb without an object noun phrase

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

However, in English there are a host of tests that provide evidence for the verb phrase (see the sources cited in Osborne (2018) for example). The set of tests that motivate a constituent in one case are different in kind and quantity than they are in the next. At this more granular perspective, debates about whether some language 'has' or does not 'have' an NP or VP (Austin & Bresnan 1996, Louagie 2021) miss the point that languages might still vary in terms of the degree to which the latter structures are supported and what types of constituency tests support them.

That it might be theoretically legitimate to treat the Chácobo VP and the English VP as the same in some sense would be beside the point. If we are interested in comparing language constituent structures to the finest degree of detail, we need to start out by dissecting abstract constituents down to the tests that are used to justify them. The planar-structure is designed to do just that.

To further develop the English planar structure we would continue adding positions until any and all predicative sentences of the language could be 'fit out' with planar structure positions. Thus we would add positions, for negative marking, adverbs, verbal particles, all of the auxiliaries, fronted constituents etc. This should be kept in mind in the following structure. A complete planar-structure analysis of English would require a paper of its own.

Given that the structure is built specifically to represent linear ordering among elements, a question arises as to how variably ordered elements can be represented in the structure. As with typical phrase structure grammars, we can add structural positions that allow elements to 'base generate' in alternative positions. For instance, to represent the variable ordering of *quickly* with the verb phrase in English as in 48.

(48) a. *The student analyzed the sentence* ***quickly***
b. *The student* ***quickly*** *analyzed the sentence*

We add the requisite positions for *quickly* in the planar-structure in order to account for its ordering in relation to the elements we already have as in 49.

(49) S/VP ⟶ 1 2 3 4 5 6 7 8 9 10 11
1 ⟶ NP
2 ⟶ *quickly*
3 ⟶ *will*
4 ⟶ *quickly*
5 ⟶ *have*
6 ⟶ *quickly*

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

7 ⟶ *analyze*
8 ⟶ *ed*
9 ⟶ NP
10 ⟶ *quickly*
11 ⟶ [P NP]

Another issue arises when we consider the fact that certain modifiers of the verb can combine with the verb complex iteratively (Vater 1978, Forker 2014). Prepositional phrases in English display this property.

(50) *The student analyzed the sentence* [*at his desk*]$_{PP}$ [*in class*]$_{PP}$ [*without thinking*]$_{PP}$ ...

To accommodate iterably combining modifiers we introduce a distinction in positions between slots and zones (Tallman 2018, 2021d).

(51) a. **slot**: Can only fit out a single element at a time.
b. **zone**: Can fit out multiple elements which can surface in any order.

The last planar-structure rule only has to be modified by making position 11 a zone (11_{zone} ⟶ PP) , which means that the category PP can repeat itself in that position.

Planar structures do not flatten out word and phrase structure without limits. We can only flatten out the templates insofar as we do not run into self-similar embedding or recursion. A relative clause in a nominal template will be represented as a single element, rather than flattening out a whole sentential template along with the nominal elaborators. A noun phrase (or more appropriate a nominal planar structure) in a verbal template will typically just be represented as a single element as well. Thus, we will have multiple planar structures for each functional domain (predicate, referential expression) or part of speech. This is why in the example above NP and PP are represented as elements of the verbal plane. A noun will receive its own planar structure. The prepositional phrase will be coded as a nominal planar structure plus an element that codes the relationship between the verb and the noun, i.e. a case or P in rule 11.

6.2 A formal sketch of planar structure grammars

A planar structure grammar is a coding device outfitted with the following elements:

(52) a. Planar structures (V, N, Adj, Adv ...)

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

b. Non-terminal elements / "Positions"
c. Terminal elements that occur inside positions
d. Planar structure rules/templates
e. Two types of rules for positions ("slots" versus "zones")

Each terminal planar structure has a fixed number of non-terminal elements we call **positions** (see Partee et al. 1990 for discussion of terminal versus non-terminal elements). Apart the 'initial symbol' introducing the planar structure and positions of the planar structures, nonterminal elements are not allowed. All other elements associated with planar structures are terminal nodes. We call these terminal elements just 'elements' for short.

The **positions** are of two types: slots and zones defined below. The slash / represents 'or'. The curly brackets are used for an unordered set of elements which do not have a precedence relationship with each other.

(53) a. $P_{slot} \longrightarrow$ a/b/c ....
b. $P_{zone} \longrightarrow$ {a,b,c ...}

Only one element can fit out a slot. The rule above outputs the following.

(54) a
b
c
∅

Inside a zone multiple elements can occur and these can occur in any order. Thus the rule $P_{zone} \longrightarrow$ {a,b,c ...}, produces the following possibilities.

(55) a b c
a c b
b a c
b c a
c a b
c b a
a b
b a
a c
c a
b c
c b

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

a
b
c
∅

If a planar structure is embedded in a zone it is understood that this planar structure can iterate (like the prepositional phrase in the example above). Thus if we have a rule as in the following

(56) $P_{zone} \longrightarrow \Pi$

where Π is or contains a planar structure, the output is as follows.

(57) Π_1
$\Pi_1\ \Pi_2$
$\Pi_1\ \Pi_2\ \Pi_3$
...

As stated above, planar structure rules consist only of non-terminal nodes called positions with precedence relations between them.

(58) vps ⟶ 1 2 3 4 5 6 ....
nps ⟶ 1 2 3 4 5 ....
...

Languages can vary in terms of the number of positions each planar structure has. Some languages might have a verbal planar structure with only around 20 positions (e.g. Araona), while others can have around 40 (e.g. Chorote). Languages can further vary in terms of how many and which positions are slots or zones.

A language with more fixed orderings will typically be represented with more slots overall. A language without any fixed ordering at all would have a single zone. So called 'free word-order' languages are are not represented with only zones. The reason is that they typically display some degree of fixed ordering inside their verbal or nominal 'words', which is represented on the same plane.

6.3 Tangling of planar structures

Planar structures can be 'tangled' with one another. This aspect of planar structures has not been systematically discussed across the studies, because most studies have focused only on verbal planar structures. Nevertheless, it is an important

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

aspect of some planar structures that needs to be described to adequately compare these with phrase structure grammars. Furthermore, if the study of constituency using planar-structures advances beyond comparing verbal structures, tangling will be dealt with in more detail.

Normal phrase structure grammars allow different types of non-terminal elements. However, in a planar structure grammar, the only types of non-terminal node are 'position' and the 'initial' symbols of the planar structures themselves. In this sense, planar structure grammars are more rigid than normal phrase structure grammars. Once again: this rigid flatness of planar structure grammar is imposed to for cross-linguistic commensurabilty so that planar structures can be constructed as constituency test coding and measurement devices, not because a linguist who uses a planar-fractal method believes that linguistic structure is flat.

More is needed to describe structural relations in a sentence apart from the formal properties described above. The reason for this is the well-known fact that verbal and nominal categories and modifiers can intermingle syntagmatically. When developing a planar-structure we allow 'tangling' between nodes if necessary in order to capture such cases (Partee et al. 1990: 442).[24] An example from English comes from the quantifier *all*, which displays a well-known property of 'stranding.'

(59) *All the students will analyze the sentences*
The students will all analyze the sentences

The problem with such sentences is that there is a nominal modifier interspersed with a verbal modifiers, yet nominal and verbal modifiers should be on distinct planar structures according to the planar-structure formalism.

To accommodate cases of part-of-speech modifier intermingling, we add a position in the verbal planar structure for the quantifier *all*. We only allow such intermingling if it is necessary, otherwise elements should be placed uniquely in their own planar structure. Preliminary verbal and nominal planar structures are provided in 60 and 61. Once again, these are only partial planar structure grammars of English developed for expository purposes.

(60) vps ⟶ 1 2 3 4 5 6 7 8 9 10 11 12

[24] Partee et al. (1990: 442) defines the 'Nontangling Condition' for a typical constituent structure grammar as follows: "In any well-formed constituent structure tree, for any nodes x and y, if x precedes y, then all nodes dominated by x precede all nodes dominated by y." Trees for planar structure grammars can violate the non-tangling condition, whenever positions of distinct planar structures are intermingled.

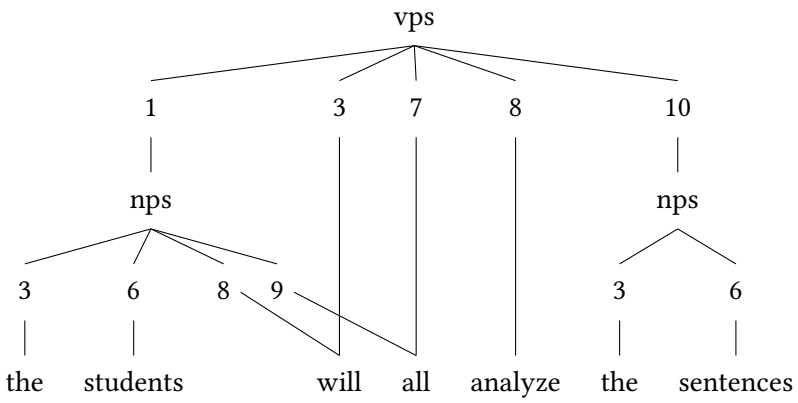
Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

1_{slot} ⟶ nps
2_{slot} ⟶ *quickly*
3_{slot} ⟶ *will*
4_{slot} ⟶ *quickly*
5_{slot} ⟶ *have*
6_{slot} ⟶ *quickly*
7_{slot} ⟶ *all*
8_{slot} ⟶ V-ROOT
9_{slot} ⟶ *ed*
10_{slot} ⟶ nps
11_{slot} ⟶ *quickly*
12_{zone} ⟶ [P nps]

Note that the element *all* would be represented in the nominal planar structure as

(61) nps ⟶ 1 2 3 4 5 6 7 8 9 10
1_{slot} ⟶ QUANTIFIER
2_{slot} ⟶ *of*
3_{slot} ⟶ *the, a, all,* nps*'s*
4_{slot} ⟶ *one, two, three ...*
5_{zone} ⟶ APS
6_{zone} ⟶ N-ROOT
7_{zone} ⟶ [*who/which...* vps]
8_{slot} ⟶ vsp[2–6]
9_{slot} ⟶ *all*

(62)



Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

In a sense admitting tangled planar structures violates the constraint I placed earlier on flattening out planar structure, since in the representation above, *all* is a modifier of the noun but also in the verbal planar structure. If we are allowed to tangle planar structures in this fashion, why not completely collapse them? The reason is because this would make planar structures infinitely long, thus becoming impractical for database construction. In order to accommodate intermingled structures while also allowing planar structures to have some practical use, we adopt the following protocol in the development of planar structures.

(63) **Tangle-only-if-necessary protocol**: Do not tangle planar structures unless it is necessary to account for the relative ordering of elements. Then, introduce the least amount of positions possible in order to capture the relevant precedence relations.

The protocol is followed by all descriptions in this volume. The restriction is imposed to guarantee commensurability across descriptions and to capture the relative ordering of elements, while enforcing finiteness on planar structures.[25]

6.4 Base elements and positions in planar structures

Another restriction on planar structures relates to their base elements. Base elements can be regarded as the phrase structure equivalents of 'heads.' But actually defining a base element as a comparative concept turns out to not be entire trivial. I will introduce these restrictions and then explain why they are adopted. The first restriction is stated below.

(64) **Base position restriction**: All planar structures have a base element. The base element is the semantic head of the planar structure (Croft 2001a, 2022). The part of speech of the base element defines the type of planar structure.

For instance, a verbal planar structure must contain a verb root, and a nominal planar structure must contain a noun root. Another restriction is imposed on how base elements are fit out in a given planar structure. Of course an immediate problem arises as to whether it is really obvious which element is the semantic head in any given case. I discuss this issue below.

Before delving into this issue a second restriction has to be imposed on the distribution of base elements within planar structures.

[25] It is not yet clear though that all tangled elements have been appropriately represented in the nominal planar structures that are presented in this volume. Such NPSs will perhaps require revision at a later stage.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(65) **Only-one-base-position restriction**: There can be no more than one position for a base element or formative that is part of a base element per planar structure.

First, note that this restriction *does not* mean that a base element cannot occupy more than one position at the same time. A base element can display multiple exponence allowing formatives split across more than one position of the planar structure if necessary. What it means is that we do not allow the same base element formative to be generated in different positions of a planar structure. Such a condition seems to be implicit in the construction of morphological templates, but in syntax it is common to think of a verb 'moving' or 'dislocating' to different positions of the clause, so the restriction requires more commentary. A similar interpretative warning is in order: we are not imposing this condition because we think 'verb root/stems never move' or 'verb roots/stems never base generate in more than one position.' Rather it is a restriction imposed to adequately code the results of constituency tests in a practical fashion.[26]

I will illustrate what this means in practice with an example from Chácobo. In Chácobo a subject NP and the verb stem (verb plus affixes) can variably order. That is S-V and V-S orders are both permissible Tallman (2018) describes cases where the V occurs before the NP S/A argument as 'verb-fronting.' An example of verb fronting is provided in 66. The first example displays S V order and the second displays V S order, where the verb and an associated motion clitic 'move' to the front of the sentence.

(66) a. *βakɨ́ tsi oʂa =kana =kɨ* '
child LNK sleep=going.ITR LNK =DECL:PST
'The child slept while going (e.g. in a truck)'

b. *oʂa=kana tsi βakɨ́ =kɨ*
sleep=going.ITR LNK child =DECL:PST
'The child slept while going (e.g. in a truck).'

Using our planar fractal notation, (at least) two competing grammars emerge for the distributional facts above. The first, allows the verb to be generated in different positions in the planar structure depicted in 67 where V^c represents the verb base and NPS represents a nominal planar structure.

[26] Furthermore, it is perfectly possible that a methodology could exist where the **Only-one-base-position restriction** is rejected. It is adopted here because when it was not imposed the reporting of constituency tests became unwieldy as one would have to fracture tests according to the position of the base element. Relaxing this condition also very naturally results in competing planar-structure analyses for the same language.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(67) a. vps ⟶ 1 2 3 4 5 6 7 8 9
b. 1 ⟶ V^{c}
c. 2 ⟶ *=kana*
d. 4 ⟶ *tsi*
e. 5 ⟶ nps
f. 6 ⟶ *tsi*
g. 7 ⟶ V^{c}
h. 8 ⟶ *=kana*
i. 9 ⟶ *=kɨ*

This planar structure requires some extra restrictions to get the distributional facts right.[27]

Another grammar might let the NPS move around in different positions and force the V□ to stay in place as in 68.

(68) a. vps ⟶ 1 2 3 4 5 6 7
b. 1 ⟶ nps
c. 2 ⟶ *tsi*
d. 3 ⟶ V^{c}
e. 4 ⟶ *=kana*
f. 5 ⟶ *tsi*
g. 6 ⟶ nps
h. 7 ⟶ *kɨ*

In the context of this project we would always choose the second grammar. The reason is that when we construct a verbal planar structure we do it with the goal of reporting constituency test results that include the verb. This restriction sometimes results in proliferation of positions around the verb in a way that many linguists might consider counter-intuitive. For instance, in South Bolivian Quechua there are a relatively large number of clitics which occur in a fixed order with respect to one another. Since they modify the predicate they are all in the verbal planar structure but they can occur before or after the verb with the

[27] Certain positions would be open or closed depending on which position the verbal base fit out. Position 6 would be open if position 7 was filled by V^{c} and otherwise closed.. Position 4 would be open if position 1 was filled by V^{c} and otherwise closed.. Position 8 would be open if position 7 was filled by V^{c} and otherwise closed. Position 2 would be open if position 1 is filled with V^{c} and otherwise closed.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

same restrictions of linear order with respect to each other. As a consequence of **Only-one-base-position restriction** we have dedicated positions for the clitics before and after verb which recode their the linear constraints these elements have with one another.

There are two reasons for imposing the only-one-base-position restriction. The most important is practical and involves limiting the scope of constituency test application to make is manageable and also (I think) more in line with how constituency tests are actually used.

For each constituency test we assume that it must overlap with the base element of a planar structure. What this basically means is that a verbal planar structure is a device for coding constituency tests which group the verb inside them. This reduces the number of constituency tests that have to be reported, but also makes the planar structure a more coherent tool for research. Defining constituency tests such that they must overlap with a specific position makes them easier to define and apply consistently.

Allowing a base element to potentially occupy more than one position complicates constituency test reporting. We would have to report different constituency test for every position we allow the verb to occupy as the spans of structure would change accordingly.

The second reason this restriction is imposed is because it restricts the number of possible planar structures that are compatible with the data. This increases comparability between the descriptions, because it reduces the number of competing planar structure analyses that a research could construction. The ideal is actually to have the construction of the planar-structure to be completely unambiguous insofar as the relevant facts are known (see Section 6.6 for discussion). This is achieved through imposing protocols and constraints on the construction of planar structures.

I now return to the notion of a semantic head which the original definition makes reference to. Simplifying Croft's discussion somewhat, a semantic head combines the notions of profile equivalent with the highest paradigmatic contrast. In a combination X+Y the profile equivalent is X if X+Y is a type of X (Croft 2001a: 257). In a combination X+Y, X is the element with the highest paradigmatic contrast if it is in paradigmatic contrast with more elements than Y (Croft 2001a: 270). In the context of the planar structure, I assume that elements that can occur in the same position are in paradigmatic contrast with each other in that position. Croft conjectures that while both profile equivalence and relative paradigmatic contrast tend to align in defining headedness at the syntactic level, in morphology, these criteria tend to misalign such that the root displays the highest paradigmatic contrast while the affix is the profile equivalent.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

As the planar-fracture method starts from the premise that we should homogenize morphological/word and syntactic/phrase structure representations as much as possible in order to investigate the actual empirical motivation for the division, Croft's notion of 'relative paradigmatic contrast' would appear to be more appropriate in defining the semantic head since it generalizes across syntax and morphology.

The main problem with systematically associating our verb base with a profile equivalent is because it is frequently the case in many languages that there is more than one element that can be considered the profile equivalent. This would seem to be especially true of languages that are traditionally labelled as polysynthetic as they contain many 'lexically heavy' elements that are neither roots in an obvious sense nor do they necessarily project their own planar structure.

To take one example, if we consider, for instance, *do-bea-tsoa* [carry-come-go.up] 'bring something up a hill' from Araona it is not clear which of the morphemes (all classified as 'roots' by (Pitman 1980)) is the profile equivalent of the whole (the action is a type of carrying, a type of coming and a type of upwards motion). Nor is this issue particularly uncommon (see Woodbury This Volume).

The paradox dissolves if we move away from identifying the verb base based on the properties of elements and define the notion based on the more abstract notion of **position**. If we associate relative paradigmatic complexity with positions, then we ask whether, when aggregating over the elements that can occur in each position, we find one position which simultaneously can function as a profile equivalent and displays a high degree of paradigmatic complexity. The verb base position is the position whose elements in the aggregate display the highest degree of paradigmatic complexity compared to other positions. The issue clearly requires more discussion, but based on the data I have observed thus far, it appears that conceptualizing the base in terms of a single position in the planar structure seems to resolve the issue of semantic head ambiguity. Another possibility would be, of course, to drop the condition that the can only be a single base, or that a base is necessary at all to define the planar structure. We have not adopted this strategy in this volume for practical reasons, but it does not mean that it is not an avenue that ought to be explored.

Developing a coding device with different formal properties and constraints might highlight different aspects of constituency structure and allow different generalizations to come to light. The main point for typological comparison though is that whatever measurement instrument is developed and used that it be applied as consistently as possible across languages.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

6.5 Minimal morphs

I stated above that the planar structure breaks down elements into positions and those positions can be composed of morphs. However, the identification of morphs is known to lead ambiguities. In a recent review of the notion of 'morph' in morphosyntactic analysis, Haspelmath (2020: 124) states 'whether a form is minimal or can be further divided into smaller forms with their own content is not always clear." (see Blevins (2016) for important discussion).

In the planar-fractal approach, we always divide forms into their smallest parts ('minimal morphs'). This means that many of the morphs will not necessarily have semantic content, rather they could just be 'recurrent partials' in the sense of Crysmann & Bonami (2016: 314). The condition is stated below.

(69) **Minimal-morph condition**: Analyze elements into morphs. Where ambiguity arises in terms of the number of morphs into which a form can be broken down, always chose the smallest element (or the analytic result that gives the most morphs).

There are two reasons for this condition. One is to impose consistency across the descriptions. The other relates to what the planar-structure is for. It is a device for measuring (mis)alignments between constituency test results. Conflation of elements could result in conflation of positions, which could result in spurious convergences between constituency tests (i.e. a loss of precision and a loss of potentially important information). In contrast, it is hard to see how any sort of spurious misalignment between tests could arise because of overly splitting morphs. If it is truly correct that some purported combination of two or more morphs should really be regarded as one, there should be no reason to expect that a constituency test would break into pieces.

6.6 Competing planar structures

One of the reasons for not using constituency structure or phrase structure analyses to compare languages is that, for a given language, even for the same set of facts considered, there are competing constituency structure analyses. This point should be obvious enough to anyone who has read debates in the syntax literature (Croft (2001a), Culicover & Jackendoff (2005) among others). Constituency tests do not apparently point to one and only one analysis. Self-described descriptive linguists might imagine they are sheltered from this problem when they claim to be following Basic Linguistic Theory, but this is an illusion, for there

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

can be competing analyses of what constitutes the grammatical and phonological word in this approach.

One criticism (or worry) that has arisen in the presentation of the methodology is the possibility that, even given the principles specified above, it might be possible that competing planar structure analyses are possible for a given language. That is, just as there are competing phrase structure analyses, there could be competing planar structure analyses.

This critique has some validity in principle. But there is an important difference between our critique of Basic Linguistic Theory, the Prosodic Hierarchy Hypothesis and traditional constituency analysis as tools for comparison and the such a critique of the planar-fractal method. In the latter cases, the ways in which ambiguities arise are easy to state (e.g. different 'wordhood tests' identify different domains of structure; different phonological domains could be mapped to different levels in the prosodic hierarchy; different constituency tests could be used or discarded in the development of a constituency analysis) and there are known empirical facts lead to such ambiguities. For the planar-fractal method, the criticism amounts to a speculation that if different researchers looking at the same set of facts from a given language *somehow* develop distinct planar structures these same researchers might *somehow* arrive at different results for the relative convergence and non-convergence of constituency tests.

But this criticism (or perhaps worry) could be applied to *all* comparative concepts. Anytime a comparative concept is proposed we might upon closer empirical scrutiny find that the concept is more ambiguous than intended.[28] In fact one of the goals of empirical research is to make sure that the comparative concept allows for consistent comparison. The solution to finding that our comparative concept is more ambiguous than intended is either to impose further restrictions on the concept or to split the concept into more variables. In the context of planar structures this would entail further tightening the protocol for building them or reporting competing analyses according to different principles. But if we simply start off with the premise that we need to develop a methodology that ensures no ambiguity could ever arise before engaging in any empirical studies, we will never engage in any empirical studies.

At a minimum someone who has such a worry about planar-structures should explain how the relevant ambiguity might arise and actually provide a case study demonstrating that it exists, in fact, and matters for the comparison of constituency tests and domains. Otherwise, it seems that the criticism amounts to prejudice

[28] A sure-fire way of never having a comparative concept scrutinized is for it to never actually be used in any typological study.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

against the idea of investigating constituency from a typological perspective in general.

7 Fracturing constituency tests

The constituency tests that one finds in the literature are ambiguous. For a given 'constituency test' or 'wordhood test' you will generally find (although not always) more than one interpretation when they are specified more precisely.

An obvious example of ambiguity in a constituency test comes from noninterruption or contiguity. The elements of words or constituents are noninterruptable or contiguous. The problem with this claim is that it is contingent on identifying an appropriate 'interrupting element.' Take a word like *post-dependence* in English. This prefix *post-* can be interrupted from *dependence* by the morph *in-* as in *post-independence*. We do not regard this as evidence that *post* does not form a word with *dependence* in the first example because of the status of *in-* as a prefix. To make the criterion more precise we might say that this is because *in-* is bound (cannot be a free form) and is highly selective of its particular base: *in-* cannot be a full utterance by itself and selects noun roots. A combination of elements that can be interrupted by a non-selective free form would be regarded as more than one word. Importantly, the criterion cannot be used unless we have stated something about the interrupting element.

When researchers assume the existence of endoclitics, the criterion for noninterruptability is implicitly relaxed. For instance, in European Portuguese the form *mostrar-emos* 'we will show' can be interrupted by a bound pronoun *-lho* as in *mostrar-lho-emos* 'we will show it to him' (Luís & Spencer 2004). The question arises as to why such constructions are not simply seen as a violation noninterruption: why are *mostrar* and *-hemos* not distinct words? Here the interrupting element is bound and one could claim that on these grounds it does not constitute a genuine instance of interruption (Bauer 2017 for the contrary position). In certain types of incorporating or compounding structures the criterion of noninterruption is further weakened if not dropped all together.

We can go even further though. In Chácobo, what the domain of noninterruption is, will depend on whether our interrupting element is a free form or a combination of free forms (e.g. a noun phrase). If we use a combination of free forms (e.g. *honi* 'man' and *siri* 'old' in a noun phrase) as the interrupting element, then the causative is part of the verbal word. If we say the interrupting element ought to be fixed as a single free form, then the causative is not part of the verbal word. This is illustrated in 70 (Tallman 2021a).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(70) *tsaya =yáma =má honi siri=́ =wa =kɨ*
see =NEG =CAUS man old=ERG =TR =DECL:PST
'The old man didn't show it to him.'

Thus, whether the constituent identified by noninterruption will depend on what we choose as an appropriate interrupting element.

One way of dealing with this issue is to choose a "correct" noninterruption test by fiat, as suggested in (Haspelmath 2011, 2023). The problem with this solution is that the result is bound to be arbitrary. Such a solution also pointlessly limits the amount of variation we are can cover in our typological study of constituency. We do not know which one of these versions of the test will be the most revealing a priori – why should we engage in a research program that pretends that we do?[29] Rather we 'fracture' the test into its different interpretations and apply all of these, coding the relevant details in the database. We define domains for interruption by a free form, by a combination of free forms, or by some promiscuous element insofar as the fractures give distinct results.

Similar considerations about ambiguity apply to phonological domains as well. The most obvious problem with identifying the span of application of a phonological process arises with 'vacuous application' of a phonological rule. Vacuous application occurs when the phonological conditions for a specific phonological rule are never met in a certain environment. If the relevant conditions are never met, one cannot tell whether the relevant phonological process and domain spans over such structural positions or not. The solution, as with morphosyntactic domains, is to fracture. I will illustrate the issue with glottal stop insertion from Chácobo below.

In Chácobo there is one environment where glottal stop insertion is obligatory: this is between two vowels at the boundary between a prefix and a root. The process does not occur if the root begins with a consonant, however. The glottal

[29] Haspelmath notes that comparative concepts should be 'useful' (Haspelmath 2010) - they are not true or false. However, in the case of his word 'retro–definition', which amounts to a domain that cannot be interrupted by any free form, he does not show how it might be useful for any conceivable typological study. In order for Haspelmath's recent intervention of the question of wordhood to be of value for empirical studies, he needs to show why christening one the many domains coded in our study as*the* 'word', as opposed to any of the other domains, is revealing. The perspective taken in this volume is different. We assume that languages might be organized in such a way that a 'word' might be definable based on a different set of diagnostics from case to case. The organization of constituency tests might show some sort of dichotomous patterning regardless of whether there is a single defining criterion across all languages. Note that this perspective is ostensibly empirical since it is not a foregone conclusion that we should find such a pattern. On the other hand, no empirical questions arise from Haspelmath's retro-definition.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

stop insertion is shown at the prefix-root boundary in 71. The non-application of the rule is found in 72.

(71) [βáʔàtʃíkɨ]
βă- atʃ -ǐ =kɨ
arm- grab -ITR =DECL:PST
'S/he grabbed her/his own arm.'

(72) [βánɨ̀ʂíkɨ]
βă- nɨʂ -ǐ =kɨ
arm- tie -ITR =DECL:PST
'S/he tied his/her arm.'

We have evidence for the existence of the process of glottal stop insertion at the boundary between prefix and root. However, at the juncture between the root and suffixes or enclitics in Chácobo no evidence for or against the application of the glottal stop insertion rule ever arises. The reason is that vowel initial transitivity markers such as *-ɨ* only ever combine with consonant final roots. Otherwise all suffixes and enclitics in Chácobo are consonant initial.

How are we to characterize the domain of application of glottal stop insertion? Does the glottal stop insertion domain span over suffixes or not? In principle there appear to be two options. One of these is to assume only positive evidence counts. This would define the prefix-root constituent as the domain for glottal stop insertion. The other is to assume that the rule applies 'vacuously' in all cases where there is no evidence against the application of the rule, i.e. where there are adjacent vowels spanning morph boundaries, but where no glottal stop insertion applies. I refer to the smaller (positive evidence only domain) as the 'minimal' domain. And the larger (negative evidence only domain) as the 'maximal domain.'

The problem with leaving the issue open to interpretation is that it allows researchers to identify spurious convergences between domains. Since the maximal domain is substantially larger than the minimal domain in Chácobo, one could claim that it convergences with any other domain of intermediate size between the minimal and maximal domains of glottal stop insertion. To be somewhat more formal, imagine the minimal domain spans 3-4 and the maximal domain 1-6 for glottal stop insertion. If we have a stress domain that spans 2-5, we can claim that the glottal stop and stress domains line up with one another if we leave the space between minimal and maximal domains open to interpretation rather than being more specific (see Tallman 2021a for the actual details in Chácobo). Not providing a formalization of the degrees of freedom in domain

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

interpretation will naturally result in theories of phonological parsing being confirmationally lax: if there is ambiguity chose the interpretation that makes your theory work.

Test fractures can be divided into different types. The first type, which reoccurs throughout the database, is the *minimal–maximal* fracture. I assume that a minimal–maximal fracture arises anytime the minimal domain is by definition a subspan of the maximal. An example of this is provided with the glottal stop insertion above. This type of fracture reoccurs throughout the database and throughout the studies in the volume for a number of constituency domains.

Another type of fracture is a distinction between stricter and laxer interpretations of a criterion. The most obvious instance where this is relevant is in the context tests of selection. The reason is that selection is a matter of degree. An element with high selectivity, might only combine with verbs. One with lower selectivity might only combine with nouns. An element might display an intermediate status in that it can appear in nonverbal predicates, but not strictly combine with nouns, however. For instance, the assertive morpheme *rá* in Chácobo requires there to be a verbal predicate. The reportative only requires there to be a predicate, verbal or nonverbal. We can, thereby, define domains based on laxer and stricter definitions of selection.[30]

There are also fractures which relate to specific constructions of a language. The most obvious cases relate to recursion based diagnostics, or 'subspan repetition.' These have to be fractured according to what appear to be very language specific subtypes (e.g. same vs. different subject clauses in Pano languages, "word-internal" complementation structures in Inuit–Yupik–Unangan languages; compounding and/or serial verb constructions in Zapotec languages). Each of these constructions can be constituent identifying in different ways, but often they are distinguished according to highly specific structural criteria. This does not mean that the different instances of subspan repetition cannot be taxonomized into different subgroups eventually (Bickel 2015). Future research might reveal that different construction types can be further broken down into codable properties for typological investigation (Bickel 2010).[31]

[30] Javier Carol (see paper in this volume) in particular is to be credited with highlighting this point, which was not initially obvious to me (Tallman 2021a)

[31] A comment at this point is necessary to avoid confusion. It has been suggested to me that somehow fracturing involves abandoning 'comparative concepts.' I do not think this is correct. Fracturing in the context of this research project simply means that each collaborator is responsible for developing and applying comparative concepts in the process of database development. Attention to concrete details not subsumed under a comparative concept does not entail abandonment of comparative concepts. For instance, we can code the domain which is

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

A final way that tests can be fractured is *by analysis*. This situation arises when interpretation of a test is contingent on whether some set of formatives are interpreted as being allomorphs of a single morpheme or diachronically related forms. The structure of syntagmatically defined distributional classes is contingent on such analytic decisions and tests that refer to linearization can, thereby, be affected as well. A clear example comes from the causative *-chi* in South Bolivian Quechua. Camacho-Rios (2022) splits occurrences of the morph into cases where the suffix is 'lexicalized' with a verb base and cases where it is not. Muysken (1981), among others, does not adopt such an analysis, and, in fact, argues against it. These analytic differences matter for the interpretation of constituency tests since they change facts about the relative (non)permutability of elements in the Quechua verb complex. Fracturing according to analysis here implies reporting different tests depending on which of the analyses of the *-chi* morphs is adopted. Fracturing by analysis provides us with important information about analytic ambiguity in the assessment of constituency tests.

8 Domains: Morphosyntactic, phonological and indeterminate

It is outside of the scope of this introduction to provide a full review of all the constituency tests and issues in their application. In this section I list the main test/domain types that we attempted to code across all the languages of this study. These can be classified into 'morphosyntactic', 'phonological' and 'indeterminate.' The morphosyntactic tests/domains are listed in 73. The phonological tests are listed in 74. For details on how to apply the relevant tests and how they are fractured the reader should consult the chapters of this volume.

(73) Morphosyntactic tests/domain types

- **Nonpermutability**: A span wherein the elements do not displlay variable ordering with respect to one another.
- **Noninterruptability**: A span that cannot be interrupted by an element of a certain type.
- **Ciscategorial selection**: A span whose elements are ciscategorial selective with respect to a particular part of speech.
- **Recursion-based/Subspan repetition**: For a specific construction that involves repetition of positions in in the planar structure (e.g.

not interruptable by a single free form in Chácobo and Hup, but note that the relevant interrupting elements are morphemes with different semantics.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

conjunction, reduplication), the largest possible span where size is calculated as *R–L*, where R is the right edge and L is the left edge of positions filled out by elements in each of the conjoined spans of structure.

The phonological domains are divided into two overarching types. We also annotate these with the classifications provided in (Bickel et al. 2007) as well as these are largely appropriate for our purposes.

(74) Phonological tests/domain types

(75) **Segmental**: A span wherein a segmental phonological process applies.

(76) **Suprasegmental**: A span wherein a suprasegmental process applies.

A number of coded domains do not fall straightforwardly into either the morphosyntactic or phonological categories. We refer to these as indeterminate domains they are listed in 77 below.

(77) Indeterminate domains

(78) **Free Occurrence**: A span which is a single free form.

(79) **Deviations from biuniqueness**: A span which displays deviations from biuniqueness.

(80) **Recursion-based/Subspan repetition**: For a specific construction that involves repetition of positions in in the planar structure (e.g. conjunction, reduplication), the span wherein none of the elements can display wide scope over the conjoined spans of structure.

Free occurrence is sometimes listed described as a morphosyntactic test (Haspelmath 2011) and sometimes as a phonological one (Zingler 2020). Deviations from biuniqueness (e.g. circumfixation, domains for the cells of inflectional classes etc.) mix phonological and morphosyntactic properties in such a way that straightforward classification as morphosyntactic or phonological is problematic. Finally, conjunction of spans of structure is used as a test for constituency, but accounts differ on whether wide-scope phenomena are a product of ellipsis or not. On an ellipsis based account the relevant domain could be phonological, which is why this domain is coded as indeterminate (see Osborne (2006) for background).

It is worth stressing that coding a domain as indeterminate simply reflects agnosticism in the stage of coding data, rather than a theoretical commitment.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Furthermore, I would like to emphasize that the constituency tests applied in this volume do not exhaust what one *could* code as a constituency test in this approach. There are other aspects of constituency structure that have not yet been operationalized to a point where they can be coded in a cross-linguistic study. An example of this would be constraints related to islandhood which form an important part of the insights achieved in the syntactic literature. Hopefully future research will fill in the relevant gaps. The planar-fractal method is *extensible* in the sense that new tests can be added as we learn more about constituency and expand the scope of the project to new domains.

9 Papers of this volume

The papers on this volume contribute to the description and analysis of wordhood and constituency phenomena in the languages of the Americas. We attempted to do this by applying a unified methodology, the planar-fractal method. Researchers are also encouraged to critique the method: this allows for the development a cross-linguistic database in the short term, but also for the development of ideas about how to improve or expand te coverage of the methodology in the long term.

In Chapter 2, Woodbury provides a description of constituency in Central Alaskan Yupik (Inuit–Yupik–Unangan, USA). Cup'ik displays a relatively high degree of convergence around the 'word' domain, as it is understood in Inuit–Yupik–Unangan studies. Out of the studies of this volume, the evidence for wordhood based on convergence is perhaps the most impressive in this language. However, Woodbury identifies a number of word 'slivers' inside the traditional word that could also be identified as 'words' if other criteria were rallied. Woodbury provides a number of incisive comments on the definition of wordhood in Cup'ik. He points out that 'conventionalized coherence and meaning', while specified as a wordhood diagnostic in Dixon & Aikhenvald 2002, identifies lexemic verb bases in Cup'ik. Woodbury also critiques Tallman's 2021a notion of the base, used as a type of non-moveable anchor in the construction of a planar structure. Tallman (2021a) tried to use Croft's notion of semantic head to define this construct, but Woodbury points out that the criteria for semantic headedness give competing results in Cup'ik. This raises the question as to whether planar structures 'bake in', so to speak, an assumption about language structure (one semantic head per part of speech domain) that does not apply in all cases.

In Chapter 3, Hiroto Uchihara provides a description of Oklahoma Cherokee (Iroquian, USA). He shows a high number of convergences around the traditional

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Iroquian word in this language. He provides a detailed discussion of how the domains identified in Cherokee relate to categories of the prosodic hierarchy. While previous research has reanalyzed the Iroquian 'word' as a phrase, Uchihara points out that this depends on what criterion or set of criteria are rallied to support domain labeling. Based on a the relatively high number of convergences found in Cherokee, he points out that apparent cases of domain misalignment could arise as much from looking at an arbitrarily low number of criteria (e.g. Bickel & Zúñiga (2017)). While certain languages may show a relatively high amount of domain misalignments, 'emergentist' explanations still need to explain high convergences where they occur.

In Chapter 4, Miller applies the planar-fractal method to Kiowa (Tanoan, USA). She argues that the methodology provides further support (in addition to Miller & Sande 2021) for Tri-P mapping, a phase-based theory of the syntax-phonology interface. In this approach phonological domains are the output of morphosyntactic phases, defined in terms of derivations in syntax. Empirically the results suggest that for every phonological domain there is at least one converging morphosyntactic one. Miller's chapter shows that the planar-fractal method might be helpful in testing competing theories of the syntax-phonology interface since it 'strips away theoretical assumptions' that can lead to noncommeasurability between linguistic analyses.

Nakamoto provides a detailed analysis of constituency in Ayautla Mazatec (Popolocan, Oto-Manguean, Mexico) in Chapter 5. Nakamoto shows a relatively low amount of convergence in phonological domains. He shows that interesting analytic issues arise with domain (mis)alignment assessment because of the presence of concatenative floating tones. This suggests again more potential problems in assessing domain (mis)alignment cross-linguistically.

In Chapter 6, Auderset, Hernandez Martinez and Ventayol-Boada provide a description of constituency tests applied to Duraznos Mixtec (Baja Mixteca, Oto-Manguean, Mexico). Duraznos Mixtec displays the most striking misalignments out of any languages in the volume, including a high number of bracketing paradoxes. The authors suggest the high degree of ambiguity in identifying the word is reflected in the literature, with authors representing Mixtec languages with different degrees of synthesis orthographically. In general, the results could be regarded as evidence for Pike's contention that the morphology-syntax and word-phrase distinctions are weak or unmotivated in Mixtec languages, yet we should refrain from claiming that all Mixtec is the same in this regard.

In Chapter 7, Gutierrez and Uchihara apply the planar-fractal to the analysis of nominal and verbal domains in Teotitlán del Valle Zapotec (Zapotecan, Oto-Manguean, Mexico). They argue that there is some support for the morphosyn-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

tactic and phonological words independently based on the clustering assumption (i.e. words are domains of high clustering). Based on the clustering assumption, TDZ Zapotec would appear to be closer to isolating than is has been described in previous literature, at least morphosyntactically. Assessment of the clustering of phonological domains is less clear, however. The highest domain appears to be the one with the strongest convergences. The results suggest that a clustering assumption cannot be used to divide words from phrases: higher utterance/sentence level domains might be just as likely to show high convergences.

Campbell provides a description of constituency in Zenzontepec Chatino (Chatino, Oto-Manguean, Mexico) in Chapter 8. He shows a high degree of convergence in Zenzontepec Chatino on (morpho)phonological grounds around a small span of structure, which he described as the word in previous work. The situation is reminiscent of Central Alaskan Yupik in terms of convergences, but for a smaller (isolating?) word domain. However, in Zenzontepec Chatino identifying a morphosyntactic word is more problematic. Nevertheless, a question arises in such cases as to how an emergentist approach would explain high convergences in phonological processes found in Chatino.

In Chapter 9, Duzerol provides a description of the French-based creole Martinican (Martinique). According to Duzerol there are not many phonological criteria that can be used to motivate a notion of phonological word in the language, thus most of the criteria that one can rally to analyze Martinican structure are morphosyntactic. Duzerol discusses the results in light of orthographic conventions and practices in Martinican. While the results do not line up with official orthographic conventions for delineating words, Duzerol suggests they might line up more with actual writing practice.

In Chapter 10, Epps provides a description of Hup using the planar-fractal method. By focusing on the difference between Hup and its sister language Daw, she argues that one could characterize the Hup and/or Daw structures as isolating or synthetic depending on which criteria are prioritized. Either both languages are isolating, Hup is polysynthetic and Daw is isolating or both languages are polysynthetic depending on which criteria are considered. Epps suggests that the key difference between Daw and Hup, the phonological integration of elements in a fixed order into a larger phonological unit in the latter but not the former, arose due to contact with Tukanoan languages. Epps discussion also reveals that noninterruption as a test is not obviously informative. There are many different noninterruptable domains depending on which element is chosen. Epps suggests a diachronic explanation for this situation.

Lemus-Serrano provides a description of constituency in Yukuna (Arawak, Colombia) in Chapter 11. Overall, Lemus-Serrano reports extremely low levels

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

of convergence in Yukuna (somewhat surprising given that Yukuna's template also requires a relatively lower number of positions compared to that of other languages). This raises questions about the applicability or relevance of categories such as morphosyntactic and phonological word for the language. The synthetic status of Yukuna is likewise unclear because it depends on which criteria are prioritized. On the other hand Lemus-Serrano argues that the results support current diachronic scenarios about the evolution of person prefixes/proclitics in Arawak.

Salanova provides a description of Mẽbêngôkre (Ge, Brazil) in Chapter 12. Salanova argues that the planar-fractal analysis provides further support for the notion of word that was adopted in his previous analysis. That the relevant constituent is a word is also supported by the fact that a number of structure preserving morphophonological processes occur within the same span. Interestingly, Salanova suggests that the language has few obvious post-lexical processes. Apart from this Salanova shows that Mẽbêngôkre displays a number of striking bracketing paradoxes that are mostly related to the possibility of incorporating postpositions into a span of structure left-adjacent to the word.

In Chapter 13, Tallman describes the application of constituency tests to Araona (Takanan, Bolivia). I argue that whether we find convergences within the phonological or morphosyntactic domains depends on how certain 'indeterminate' domains are classified. It is unclear whether deviations from biuniqueness, minimal subspan repetition and free occurrence domains should be classified as morphosyntactic or phonological. How to relate the results to claims about morphosyntactic and phonological structure is contingent on how we treat these indeterminate domains. Overall there is a way of interpreting the results with respect to common assumptions about wordhood in Takanan languages, but the planar-fractal method shows that such analyses are partially arbitrary. Whether Araona is isolating or (poly)synthetic depends on which of the diagnostics we assume are word identifying versus phrase identifying.

In Chapter 14, Camacho-Rios and Tallman provide an analysis of Uma Piwra South Bolivian Quechua (SBQ) (Quechua, Bolivia). We find some support from the planar-fractal method for the orthographic word in SBQ. SBQ is interesting because of the number of complex morphemes that replace spans of structure internal to the word, but without covering the root (semantic head). In the phonological domain, there are no convergences in SBQ. We contextualize the results in relation to debates about the morphology-syntax distinction in Quechua.

In Chapter 15, Carol provides a description of constituency in Chorote (Matacoan, Argentina). Carol discusses his results in terms of the high degree of 'transcategoriality' of elements in Chorote. Transcategoriality is relevant for the way

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

we have formulated selection in this project. A selection domain is one which contains elements which can only combine with a single part of speech class. Carol argues that this domain is, in fact, ambiguous because it depends on whether we are concerned with 'selection of a predicate' versus 'selection of a verb' in its assessment. He breaks down the criteria further to capture this difference. Chorote also displays cases where the nominal must be partially interspersed (tangled) with the verbal one because the distribution of noun phrases in Chorote depends on whether these occur with a demonstrative or not: nominal demonstratives also incorporate into the Chorote verbal structure. Overall the results for Chorote suggest a highly ambiguous situation without obvious support for the word bisection thesis.

In Chapter 16, Juarez provides a description of Mocovi (Guaycuruan, Argentina) using the planar-fractal method. Juarez shows that the constituency test results in Mocovi support a graded notion of word. Minimal fractures of domains overall suggest a much smaller word constituent than has been described for Guaycuruan languages, whereas maximal domains come closer to supporting a larger word constituent.

Chapter 17 provides an overview of the results of the volume. We focus on the structure of the database and the workflow for its development. We target three assumptions in linguistics that we think need to be revised in light of the results of this volume. This chapter calls for reassessment of the notion of synthesis, wordhood test, and claims about the relative reliability of tests in the linguistic literature.

Chapter 18 provides a critical and retrospective commentary on the project of comparing wordhood and constituency cross-linguistically by Kristine Hildebrandt. Hildebrant compares the methodology of the Word Domains project to the Constituency-Convergence project, commenting on areas that still require future research.

In Chapter 19, Taylor Miller further assesses the planar-fractal method in relation to a one of the current theories of syntax-phonology interaction: Tri-P mapping with Cophonologies by Phase. Taking some select examples from this volume, she argues that the model makes successful predictions concerning the patterns found in Araona and Ayautl Mazatec. She shows that a description of the data in terms of the planar-fractal method permits a relatively stream-lined assessment of how well data fit syntax-phonology interface theories, thus opening the door to more rigorous intertheoretic comparison.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

References

Aikhenvald, Alexandra Y., R.M.W. Dixon & Nathan M. White. 2020. The essence of 'word'. In *Phonological word and grammatical word*. Oxford: Oxford University Press.

Anderson, Mark. 2019. *From boas to black power: racism, liberalism, and american anthropology*. Stanford: Stanford University Press.

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

Anderson, Stephen R. 2015. Dimensions of morphological complexity. In Matthew Baerman, Dunstan Brown & Greville Corbett (eds.), *Understanding and measuring morphological complexity*, 11–28. Oxford: Oxford University Press.

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

Baker, C.L. 1995. *English syntax (second edition)*. Cambridge: MIT Press.

Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.

Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge.

Bennett, Ryan & Emily Elfner. 2019. The syntax-prosody interface. *Annual Review of Linguistics* (5). 151–171.

Bickel, Balthasar. 2010. Capturing particulars and universals in clause-linkage. In Isabelle Bril (ed.), *Clause linking and clause hierarchy: Syntax and pragmatics*, 51–104. Amsterdam: John Benjamins Publishing Company.

Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *Oxford handbook of linguistic analysis*, 2nd edn., chap. 37, 901–924. Oxford University Press.

Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Prasad Paudyal, Ichchha Purna Rai, Manoj Rai, Novel Kishore Rai & Sabine Stoll. 2007. Free prefix ordering in Chintang. *Language* 83(1). 43–73.

Bickel, Balthasar, Kristine A. Hildebrandt. & René Schiering. 2009. The distribution of phonological word domains: a probabilistic typology. In Janet Grijzenhout & Kabak Baris (eds.), *Phonological domains: universals and deviations*. De Gruyter Mouton.

Bickel, Balthasar & Johanna Nichols. 2002. Autotypologizing databases and their use in fieldwork. In *Proceedings of the international LREC workshop on resources and tools in field linguistics, Las Palmas*, vol. 2627.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Bickel, Balthasar & Fernando Zúñiga. 2017. The 'word' in polysynthetic languages: Phonological and syntactic challenges. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford handbook of polysynthesis*, 158–185. Oxford: Oxford University Press.

Blevins, James P. 2016. *Word and paradigm morphology*. Oxford.

Bloomfield, Leonard. 1914. Sentence and word. *Transactions and Proceedings of the American Philological Association* (45). 65–75.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart & Winston.

Boas, Franz. 1911. Introduction. In *Handbook of american indian languages, bulletin 40, part 1*, 1–83. Washington D.C.: Bureau of American Ethnology.

Bögel, Tina. 2021. Function words at the interface: a two-tiered approach. *Languages* (6). 197.

Booij, Geert. 1996. Cliticization as prosodic integration: the case of dutch. *The Linguistic Review*. 219–242.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan, Parmotima Chakraborti, Dagmar Jung & Joanne Scheibman. 1998. Prosody and segmental effect: some paths of the evolution of word stress. *Studies in Language* 2(22). 267–314.

Bybee, Joanna, Joan; Scheibman. 1999. The effect of usage on degrees of constituency: the reduction of don't in english. *Linguistics* 4(37). 574–596.

Camacho-Rios, Gladys. 2022. *Verbal complexity in south bolivian quechua: insights from the speech of monolingual elders*. Austin: University of Texas at Austin. (Doctoral dissertation).

Carnie, Andrew. 2000. On the definition of x0 and xp. *Syntax* 2(3). 59–106.

Carnie, Andrew. 2010. *Constituent structure*. New York: Oxford University Press.

Chao, Yuen Ren. 2011. *A grammar of spoken chinese*. Beijing: The Commercial Press.

Cheng, Laura J., Lisa Lai-Shen; Downing. 2021. Recursion and the definition of universal prosodic categories. *Languages* 3(6). 125.

Croft, William. 2001a. *Radical Construction Grammar*. Oxford, U.K.: Oxford University Press.

Croft, William. 2001b. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2010. Ten unwarranted assumptions in syntactic argumentation. In Kasper Boye & Elisabeth Engberg (eds.), *Language usage and language structure*, 313–350. Berlin, New York: De Gruyter Mouton.

Croft, William. 2022. *Morphosyntax: constructions of the world's languages*. Cambridge: Cambridge.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in information-based morphology. *Journal of Linguistics* 52(2). 311–374.

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press.

de Souza, Livia de Camargo Silva Tavares. 2020. *Switch-reference as anaphora: a modular account*. New Jersey: Rutgers, The State University of New Jersey. (Doctoral dissertation).

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A Cross-linguistic Typology*, 1–41. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486241.002.

Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2010a. *Basic linguistic theory: violume 1 methodology*. Oxford University Press: Oxford.

Dixon, R.M.W. 2010b. *Basic linguistic theory: volume 2 grammmatical topics*. Oxford University Press: Oxford.

Downing, Laura J. & Maxwell Kadenge. 2020. Re-placing pstem in the prosodic hierarchy. *The Linguistic Review* 3(37). 433–461.

Dryer, Matthew S. 2006. Functionalism and the theory-metalanguage confusion. In Grace Wiebe, Gary Libben, Tom Priestly, Ron Smyth & Sam Wang (eds.), *Phonology, morphology, and the empirical imperative: papers in honour of bruce derwing*, 27–59. Taipei: The Crane Publishing Company.

Féry, Caroline. 2017. *Intonation and prosodic structure*. Cambridge University Press: Cambridge.

Fleck, David. 2003. *A grammar of matses*. Rice University. (Doctoral dissertation).

Fleck, David. 2013. *Panoan languages and linguistics*. Washington: Anthropological Papers of the American Museum of Natural History.

Forker, Diana. 2014. A canonical approach to the argument/adjunct distinction. *Linguistic Discovery* 12(2). 27–40.

Good, Jeff. 2016a. *The linguistic typology of templates*. Cambridge: Cambridge University Press.

Good, Jeff. 2016b. *The linguistic typology of templates*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2008. The oxford handbook of grammatical analysis. In chap. Framework-free grammatical theory. Oxford: Oxford University Press.

Haspelmath, Martin. 2010. Comparative concpts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–689.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45. 31–80. DOI: 10.1515/flin.2011.002.

Haspelmath, Martin. 2020. The morph as a minimal linguistic form. *Morphology* (30). 117–134.

Haspelmath, Martin. 2023. Defining a word. *WORD* 69.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Gilbert Kiparsky Paul; Youmans (ed.), *Phonetics and phonology, volume 1*, 201–260. New York: Academic Press Inc.

Hildebrandt, Kristine A. 2007. Grammatical and phonological domains in limbu. *Himalayan Linguistics* (8). 1–34.

Hockett, Charles F. 1947. Morphology: the descriptive analysis of words by eugene a. nida. *Language* 23(3). 273–285.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: The Macmillan company.

Hyman, Larry M. 1993. Structure preservation and postlexical tonology in dagbani. In Ellen Hargus Sharon; Kaisse (ed.), *Studies in lexical phonology*, 235–254. San Diego: Academic Press.

Ishihara, Shinichiro & Sara Myrberg. 2023. Match theory and the asymmetry problem: an example from stockholm swedish. *Languages* (8). 102.

Ito, Armin, Junko; Mester. 2021. Recursive prosody and the prosodic form of compounds. *Languages* 2(6). 65.

Kahneman, Daniel, Olivier Sibony & Cass R. Sunstein. 2021. *Noise: a flaw in human judgment*. New York: Little Brown Spark.

Kroeger, Paul R. 2005. *Analyzing syntax: a lexical-functional introduction*. Cambridge: Cambridge University Press.

Kügler, Frank. 2023. Phrase-level atr vowel harmony in anum - a case of recursive prosodic phrasing. *Languages* (7). 308.

Levine, Robert D. 2017. *Syntactic analysis: an hpsg-based approach*. Cambridge: Cambridge University Press.

Loos, Eugene E. 1999. Pano. In *The amazonian languages*, 227–251. Cambridge: Cambridge University Press.

Louagie, Dana. 2021. *Noun phrases in australian languages: a typological study*. Boston/Berlin: De Gruyter.

Luís, Ana & Andrew Spencer. 2004. A paradigm function account of 'mesoclisis' in european portuguese. In Jaap Booij Geert; van Marle (ed.), *Yearbook of morphology 2004*, 177–228. Springer.

Martinet, André. 1962. *A functional view of language*. Oxford: Clarendon Press.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Matthews, Peter. 2002. What can we conclude? In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A Cross-Linguistic Typology*, 266–281. Oxford: Oxford University Press.

Mayo, Deborah. 2018. *Statistical inference as severe testing: how to get beyond the statistics wars*. Cambridge University Press: Cambridge.

McCawley, James D. 1988. *The syntactic phenomena of english (volume 1)*. Chicago: University of Chicago Press.

McGregor, William B. 2021. *Neo-firthian approaches to linguistic typology*. Sheffield, U.K.: Equinox Publishing.

Michaud, Alexis. 2017. *Tone in Yongning Na: lexical tones and morphotonology*. Berlin: Language Sciences Press.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Miller, Taylor. 2018. *The phonology-syntax interface and polysynthesis: a study of kiowa and saulteaux ojibwe*. University of Delaware. (Doctoral dissertation).

Miller, Taylor L. & Hannah Sande. 2021. Is word-level recursion actually recursion? *Languages* 100(6). 1–27.

Muysken, Pieter. 1981. Quechua word structure. In Henry Frank (ed.), *Binding and filtering*, 279–327. Cambridge: MIT Press.

Neely, Kelsey. 2019. *The linguistic expression of affective stance in yaminawa (pano, peru)*. Berkeley: University of California, Berkeley. (Doctoral dissertation).

Nespor, Irene, Marina; Vogel. 2007. *Prosodic phonology (with a new foreword)*. 2nd edn. Berlin: Mouton de Gruyter.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven & David T. Mellor. 2018. The preregistration revolution. *PNAS* 115(11). 2600–2606.

Osborne, Timothy. 2006. Shared material and grammar. *Zeitscrift für Sprachwissenschaft* (25). 39–93.

Osborne, Timothy. 2018. Tests for constituents: what they really reveal about the nature of syntactic structure. *Language Under Discussion* 1(5). 1–41.

Ouhalla, Jamal. 1999. *Introducing transformational grammar: from principles and parameters to minimalism (second edition)*. London: Edward Arnold.

Partee, Barbara, Alice ter Meulen & Robert E. Wall. 1990. *Mathematical methods in linguistics*. Kluwer Academic Publishers: Dordrecht.

Peperkamp, Sharon. 1996. On the prosodic representation of clitics. In Ursula Kleinhenz (ed.), *Interfaces in phonology*, 102–127. Berlin: Akademie Verlag.

Peperkamp, Sharon. 1997. *Prosodic words*. The Hague: Holland Academic Graphics (HIL dissertations 34).

Pike, Kenneth. 1943. Taxemes and immediate constituents. *Language* 2(19). 65–82.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Pike, Kenneth L. 1972. A problem in morphology-syntax division. In Ruth M. Brend (ed.), *Kenneth l. pike selected writings*, 74–84. First published in: *Acta Linguistica* 5:3 (1949), 125-138. Reprinted by permission. Mouton.

Pitman, Donald. 1980. Bosquejo de la gramática araona. In *Notas lingüísticas de Bolivia: No. 9*, 126. Cochabamba: Summer Institute of Linguistics.

Plag, Ingo. 2014. Phonological and phonetic variability. *Italian Journal of Linguistics / Revisti di Linguistica* (26). 209–228.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge: MIT Press.

Risen, Jane & Thomas Gilovich. 2006. Informal logical fallacies. In *Critical thinking in psychology*, 110–130. Cambridge University Press: Cambridge. DOI: https://doi.org/10.1017/CBO9780511804632.008.

Rodseth, Lars. 2022. Reality remodeled: practical fictions for a more-than-empirical world. *HUA: Journal of Ethnographic Theory* (12). 217–234.

Russell, Kevin. 1999. What's with all these long words anyways? In Leora Barel, Rose-Marie Dechaine & Charlotte Reinholtz (eds.), *Mit occasional papers in linguistics*, 119–130.

Sadock, Jerrold M. 1991. *Autolexical syntax: a theory of parallel grammatical representations*. Chicago: University of Chicago Press.

Sadock, Jerrold M. 1980. Noun incorporation in Greenlandic: A case study of syntactic word formation. *Language* 56. 300–319.

Scheer, Tobias. 2010. *A guide to moprhosyntax-phonology interface theories: how extra-phonological information is treated in phonology since Trubetzkoy's grenzsignale*. Amsterdam: De Gruyter Mouton.

Schiering, Réné. 2006. *Cliticization and the evolution of morphology: a cross-lilnguistic study on phonology in grammaticalization*. Universität Konstanz. (Doctoral dissertation).

Schiering, René, Balthasar Bickel & Kristine Hildebrandt. 2012. Stress-time = word-based? Testing a hypothesis in prosodic typology. *STUF-Language Typology and Universals* 65(2). 157–168.

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.

Seidl, Amanda. 2001. *Minimal indirect reference: a theory of the syntax-phonology interface*. London: Routledge.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Seifart, Frank, Jan Strunk, Swintha Danielson, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nikolaus P. Himmelmann & Balthasar Bickel. 2021. The extent and degree of utterance-final word lengthening in spontaneous speech from 10 languages. *Linguistics Vanguard*.

Selkirk, Elisabeth. 1996. The prosodic structure of function words. In Katherine Morgan James L.; Demuth (ed.), *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Mahwah: Lawrence Erlbaum Associates.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge: The MIT Press.

Selkirk, Elizabeth. 2011. The syntax-phonology interface. In John Jason Riggle Goldsmith & Alan C.L. Yu (eds.), *The handbook of phonological theory (second edition)*, Second, chap. 14, 435–484. Wiley-Blackwell: Blackwell Publishing.

Smaldino, Paul E. 2017. Models are stupid and we need more of them. In Robin R. Vallacher, Stephen J. Read & Andrzej Nowak (eds.), *Computational social psychology*, 311–333. Milton Park: Routledge.

Sobin, Nicholas. 2008. "do so" and vp. *Linguistic Inquiry* 39(1). 147–160.

Sonderegger, Oriana Kilbourn-Ceron & Morgan. 2018. Boundary phenomena and variability in japanese high vowel devoicing. *Natural Language and Linguistic Theory* 36.

Stocking Jr., George W. 1974. Introduction: the basic assumptions of boasian anthropology. In *A Franz Boas reader: the shaping of American Anthropology, 1883-1911*, chap. 1. Chicago: The University of Chicago Press.

Tallman, Adam J. R. 2020. Beyond grammatical and phonological words. *Language and Linguistics Compass* 14(2). e12364.

Tallman, Adam J. R. 2021a. Constituency and coincidence in Chácobo (pano). *Studies in Language* 45(2). 321–383.

Tallman, Adam J.R. 2017. *Nahuapaxahua: cuento tradicional chacobo*. San Bernadino: Open Space.

Tallman, Adam J.R. 2018. *A grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon*. University of Texas at Austin. (Doctoral dissertation).

Tallman, Adam J.R. 2021b. Analysis and falsifiability in practice. *Theoretical Linguistics* 47(1-2). 95–112.

Tallman, Adam J.R. 2021c. Caroline féry, intonation and prosodic structure (key topics in phonology). cambridge: cambridge university press, 2017. pp. ix 374. *Journal of Linguistics* 1(57). 208–214.

Tallman, Adam J.R. 2021d. Constituency and coincidence in Chácobo (Pano). en. *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal. http://www.jbe-platform.com/content/journals/10.1075/sl.19025.tal (20 April, 2023).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Tallman, Adam J.R. & Sandra Auderset. 2023. Measuring and assessing indeterminacyand variation in the morphology-syntax distinction. *Linguistic Typology* 27(1). 113–156.

Tallman, Adam J.R. & Gabriel Gallinate. accepted. The prefix e- in Araona (Takana). *Amerindia*.

Valenzuela, Pilar Bismark. 2003a. *Transitivity in shipibo-konibo grammar*. University of Oregon. (Doctoral dissertation).

Valenzuela, Pilar M. 2003b. *Transitivity in Shipibo-Konibo in Grammar*. University of Oregon. (Doctoral dissertation).

Vater, Heinz. 1978. On the possibility of distinguishing between complements and adjuncts. In *Valence, semantic case, and grammatical relations*. John Benjamins Publishing: Amsterdam.

Vigário, Marina. 2010. Prosodic structure between the prosodic word and the phonological phrase: recrusive nodes or an independent domain. *The Linguistic Review* (27). 485–530.

Vogel, Irene. 2008. The morphology-phonology interface: isolating to polysynthetic languages. *Acta Linguistica Hungarica* (1-2). 205–226.

Vogel, Irene. 2019. Life after the strict layer hypothesis: prosodic structure geometry. In Youyong Zhang Hongming; Qian (ed.), *Prosodic studies: challenges and prospects*, 9–60. Oxon: Routledge.

Vogel, Irene. 2023. Is there foot structure in isolating languages? In Jeroen van de Weijer (ed.), *Part ii syllable, stress, and sign*, 99–113. Berlin: De Gruyter Mouton.

Werle, Adam. 2009. *Word, phrase, and clitic prosody in bosnian, serbian, and croatian*. University of Massachusetts Amherst. (Doctoral dissertation).

Wiltschko, Martina. 2014. *The universal structure of categories: towards a formal typology*. Cambridge University Press: Cambridge.

Wimsatt, William C. 2007. *Re-engineering philosophy for limited beings*. Harvard University Press: Harvard.

Woodbury, Anthony C. 2002. The word in Cup'ik. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 79–99. Cambridge, UK: Cambridge University Press.

Woodbury, Anthony C. 2011. Atkan Aleut "unclitic" pronouns and definiteness: a multimodularanalysis. In Etsuyo Yuasa; Tista Bagchi; Katharine Beals (ed.), *Pragmatics and autolexical grammar: in honor of Jerry Sadock*, 125–142. Amsterdam: John Benjamins.

Zingg, Philipp. 1998. *Diccionario chacobo-castellano y castellano-chacobo*. La Paz: Ministerio de Desattollo Sostenible y Planificacón - Viceministerio de Asuntos Indinas y Pueblos Originarios.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Zingler, Tim. 2020. *Wordhood isses: Typology and grammaticalization*. The University of New Mexico. (Doctoral dissertation).

Zsiga, Elizabeth C. 2020. *The phonology/phonetics interface*. Edinburgh: Edinburgh University Press.

Zúñiga, Fernando. 2014. (anti)-cliticization in mapudungun. *Morphology* (24). 161–175.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 2

Constituency and convergence in the Americas – Results and discussion

Sandra Auderset[a], Caroline de Becker[b], Gladys Camacho Rios[c], Eric W. Campbell[d], Javier Carol[e], Minella Duzerol[f], Patience Epps[g], Ambrocio Gutiérrez[h], Cristian R. Juárez[i], Magdalena Lemus Serrano[j], Stephen Francis Mann[k], Taylor L. Miller[l], Shun Nakamoto[m], Zoe Poirier Maruenda[i], Andrés Salanova[n], Hiroto Uchihara[o], Natalie Weber[p], Anthony C. Woodbury[g], Dennis Wylie[g] & Adam J. R. Tallman[b]

[a]University of Bern [b]Friedrich Schiller Universität Jena [c]State University of New York at Buffalo [d]University of California Santa Barbara [e]University of Buenos Aires [f]Laboratoire Dynamique du Langage – CNRS [g]University of Texas at Austin [h]University of Colorado, Boulder [i]Max Planck Institute for Evolutionary Anthropology [j]Aix-Marseille Université & Laboratoire Parole et Langage [k]Universitat de Barcelona & Max Planck Institute for Evolutionary Anthropology, Leipzig [l]State University of New York at Oswego [m]Universidad Nacional Autónoma de México [n]University of Ottawa [o]Tokyo University of Foreign Studies [p]Yale University

This paper provides a basic conceptual introduction to the planar-fractal method. The methodology is then contextualized with respect to multivariate typology. The structure of the database based on this method is then described and an illustration of what the database can be used for is also provided. Four issues related to contextualizing constituency in typological are then assessed in relation to the data gathered in the current volume: (i) the index of synthesis; (ii) the absence of *a priori* wordhood tests; (iii) the relative reliability of wordhood tests; (iv) the word bisection dogma.

Sandra Auderset et al. 2024. Constituency and convergence in the Americas – Results and discussion. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 85–128. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

1 A synopsis of the planar-fractal method

The planar structure is a template over which constituency tests/domains can be coded (Tallman 2021c; Tallman This Volume). It constitutes an attempt to apply the ideas of multivariate or distributional typology to the problem of constituency (Section 2). It was constructed in order to assess the degree to which logically distinct constituency tests/domains align and/or nest with each other and how much typological variation there is in this regard.

The planar structure can be conceptualized as a template, built out of a 'lumping' strategy Good 2016, meaning that as the template is outfitted to describe aspects of linear stipulation in as many constructions as possible, or as a species of phrase structure grammar with constraints imposed on what types of nonterminal node are admissible (see Tallman This Volume). We should point out again that the device is not a 'theory of grammar' in the sense of Chomsky (1965). It is a comparative concept used to study a very specific aspect of linguistic structure. In other words, it is a measuring device that could be constructed with different constraints and coding properties for different research questions (for example Good 2016). If we do not use a planar-structure or some such measurement technology, we will not have any way of keeping track of when diagnostics align and when they do not.

The 'fractal' aspect of the planar-fractal method runs off of the premise that constituency tests, stated in the abstract, can have ambiguous interpretations when applied to actual language data. When a constituency test is applied to a given language we cannot and do not apply the test *as is* to another language. Rather there is a process of abstraction and then reconcretization in the application of the 'test' to a new system. We 'lift' the test from its language specific context, making it abstract, and then add details to apply it to a new language, reconcretizing the test in the process. Every constituency test must be recycled in this fashion if it is to be applied beyond the context for which it was originally developed and used.

We note, for instance, that some span of structure which we call 'words' cannot be interrupted by other elements we have already identified as words in some language: English let's say. We abstract away from this property, and claim that 'noninterruption' is a general diagnostic for the identification of 'words.' But noninterruption by what? Surely we cannot use words of English to test whether a given span of structure in Hup is a 'word' based on noninterruption. So, we attack the problem by reconcretizing the test, adding a different Hup specific interrupting element to the equation. The epistemic leap might seem so trivial that it passes above conscious awareness.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

It is in this reconcretization of recently abstractified 'tests' where fracturing comes into play. The problem is that there is often more than one way in which a given constitutency test can be reconcretized when it is applied to a new domain. This aspect of linguistic analysis can pass below conscious awareness, especially when linguists are told to 'find' specific categories or structures in novel data, but not told how one could possibly ever justify saying that the category or structure is not present in some linguistic system (see Tallman's This Volume discussion of basic linguistic theory). Therefore, we seek to develop a method that makes the reconcretization explicit and compels us to not discard competing interpretations surreptitiously as a consequence of cognitive biases (Ackermann 1985).

If we apply the process repeatedly to more and more languages, we will find that our original 'test' has expanded into a number of sub-types. We view this as an application of the autotypologizing methodology (Bickel & Nichols 2002, Witzlack-Makarevich et al. 2022) to the problem of constituency. The goal of the project is to articulate a taxonomy of domains organized hierarchically from their abstract to their more concrete instantiations. The typology is constructed to discern whether there is some statistical order to the patterns we find with these domains in specific languages.

A planar structure can be defined as follows:

(1) **Planar structure**: a template of consecutively ordered positions from 1 to n. There is a planar structure for each part of speech which is open class. Each planar structure has a position for a core element. All other positions are for non-core elements.

Positions can be 'fit out' by core or non-core elements. But for a given planar structure there is just one position for a given core element. The core element can be defined as follows:

(2) **Core element**: A core element is an open class element. Any sentence that is fit out by a planar structure needs to have an overt core element. For instance, a verbal planar structure will have one position for a core element and all sentences that contain that core element should be able to be mapped to a planar structure. The core element functions as the semantic head (see Croft 2001: 241–280 and Croft 2022: 35–37) of a planar structure and the sentences that it outputs (see Tallman 2021c and Tallman introduction and Woodbury This Volume for discussion)

For instance, a verbal planar structure will have a position for a verb core. A nominal planar structure will have a position for a noun core. If it is necessitated

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

by the facts of the language, we could also add adjectival or adverbial planar structures as well. Currently we have limited the scope of the study to verbal planar structures, although one chapter provides preliminary nominal planar structure (Guttierez and Uchihara This Volume).[1]

As stated above a planar structure is composed of a number of positions. A position has a number, contains elements and is associated with a specific planar structure. Each position is either categorized as a slot or a zone. Slots and zones are defined below.

(3) slot: A position which can only be filled by one element at a time.

(4) zone: A position which can be filled by more than one element and the elements can occur in any order in the zone.

For expository purposes I provide a simple planar structure below. I have placed a superscript ^{c} over the core elements of the planar structure. This means that the position that contains such elements is always obligatorily filled.[2]

Table 1: Example planar structure

1	slot	a,b,h
2	slot	c
3	slot	d^{c},e^{c}
4	zone	f,g
5	slot	h

In position 1, there are 3 elements (a,b,h). In this position, only one of these elements can occur for a given sentence. This means that *acdfh* is an admissible string according to the planar structure above, but *abcdfgh* or *ahcdfgh* is not.

[1] A given core element might be fit out in more than one position. But there can be no positions which can contain the same core element in them. For instance, our planar structures are not allowed to have a position 3 and a position 5 both of which could output a core element (e.g. a verb root). However, a planar structure could have a core element which is composed of two pieces one of which occurs in position 3 and other in position 5. The reason for this restriction as described in (Tallman This Volume) is to make the reporting of constituency tests more manageable.

[2] The reverse is not true. We cannot determine that a position is a core position because it must always be filled.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

However in position 4 the elements, *f* and *g* can variably order. Thus, *acd**fg**h* and *acd**gf**h* are both admissible strings.

Positions can be obligatorily filled or optional (as with categories in a phrase structure grammar). Positions can be **open** or **closed** contingent on the presence of specific elements or whether a given position is filled. For instance, if we find that element *b* never co-occurs with *h*, we can add a stipulation that position 5 is closed if position 1 is filled with element *a*.

Note that there are two ways of describing the variable ordering of elements in a planar-structure. If the variable ordering is *local* in the sense that there is no intervening elements between the elements that variably order, then a zone is posited as with the elements *f* and *g* above. Zones of this type are useful for defining cases where affixes variably order with one another (Bickel et al. 2007) in a traditional 'word' or where adverbs or particles do so.

If the elements variably order but *around* an element which displays a fixed order, then we simply place the relevant element in more than one position as with the element *h* above. Allowing *h* to be in position 1 or position 5 means that we can have the order *hd* and *dh*. A typical example of this type of variable ordering is with noun phrases around a complex verb structure in nonconfigurational languages (Austin & Bresnan 1996). A subject NP for instance, can be given a position on each side of a span of verbal elements.

Finally, we need to define an element.

(5) **Element**: an element is a morph (Haspelmath 2020) another planar structure or a well-defined subspan of a planar structure.

As a consequence, a nominal planar structure can be in an element of a verbal planar structure or some subspan of a nominal planar structure can be an element of a verbal planar structure, and vice versa. The ability to have elements which are planar structures is necessary to make planar-structures practically useful: if this condition was not allowed recursion would mean that planar structures would not be finite. In other words, we do not flatten out phrase structure without limit.

While a planar-structure grammar imposes some hierarchical structure by allowing planar structures to embed within each other, notions such as 'word' and 'phrase' are prohibited.

With the planar structure in hand, we apply autotypology as a research method in the application of constituency tests. Constituency tests can now be operationalized as 'variables' which code spans over planar structures of specific languages (see Tallman This Volume for more details).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2 Multivariate typology and the constituency variables

Autotypology as a method emerged in the early 2000s as part of the larger AUTOTYP research program, which aims at systematically analyzing variation in the languages of the world as well as explaining this variation both quantitatively and qualitatively (Bickel & Nichols 2002, Bickel et al. 2017, Witzlack-Makarevich et al. 2022). It has also been referred to as 'Multivariate Typology' and 'Distributional Typology' (Bickel 2015), although these labels could be seen as more appropriately describing a whole research agenda, rather than only a typological method. However, they share the same approach, so in the remainder of this chapter we will use the label Autotypology as cover term for the methodology and theory behind the AUTOTYP project.

Typological variables always involve a certain degree of abstraction and generalization from language-specific details. In most typological approaches and the resulting databases, the variables as well as the possible values they can realize are determined *a priori*, usually based on tradition, theoretical assumptions, and convenience (for more details and examples see Witzlack-Makarevich et al. 2022: 632). Even approaches that try to circumvent the issues with categorization based on tradition and theory by relying on known variation and pilot studies still define the variables top-down. This is also the case for the two largest typological databases currently available, WALS (Dryer & Haspelmath 2013) and Grambank (Skirgård et al. 2023).

Autotypology differs from these more traditional typological approaches in that the variables and their values are arrived at in a bottom-up fashion and constantly adapted to capture the variation present in the data at hand. The idea behind this methodology is to invest in coding fine-grained variables that adequately account for the diversity of the world's languages and can be used to investigate a variety of research questions across different theoretical frameworks. While initially more time-consuming than relying on pre-defined, aggregated variables, the methodology ensures that the resulting database can be expanded on and (re-)used by other researchers. In the following, we will expand on the methodology and how it was used in developing the diagnostics of the constituency database. As in other frameworks, the starting point for developing variables in Autotypology is usually found in earlier typological studies or theoretical discussions relating to the research question. In the case of constituency, we can draw on a wealth of literature proposing or evaluating diagnostics for constituency as such or for wordhood and phrasehood (see Tallman This Volume). These starting point variables are not seen as static, but rather they are re-evaluated and adjusted with each new language being coded. One type of

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

adjustment frequently encountered with constituency diagnostics is fracturing, that is, the splitting of a diagnostic into multiple diagnostics.

The constituency variables are united in that they follow the definition (comparative concept) of a constituency test provided by Tallman 2021a which we cite below:

> **Constituency test**: "... a generalization within or across constructions that targets or crucially refers to some subspan of a planar structure. A constituency test can only be applied in a given language if it is specific enough such that it refers to a **well-defined** subspan. A subspan is well-defined if it contains a single left-edge (e.g. position 3) and a single right-edge (position 8)."

We start with those constituency tests which are frequently found in the literature (displacement, interruption). The diagnostics as they are found in the literature typically require a great deal of refinement to meet the definition provided above, however. Thus much of the intellectual work in developing constituency variables amounts to operationalizing relatively vague heuristics in the morphology, phonology and syntax literature so that they can be applied consistently. This often involves making finer distinctions than what is found in the literature. For instance, noninterruption can be cut into different tests depending on what we choose as the interrupting element. The converse situation also arises. There are cases where the literature attests of apparently distinct diagnostics, but, upon closer scrutiny, it is revealed that they are the same, but described or conceptualized somewhat differently. An example of this is with the distinction between noninterruption in the morphology/wordhood literature and displacement in the syntax literature. The identity between diagnostics that are often described as if they were distinct arises when we think hard about whether convergences between diagnostics might be a spurious consequence of the how such diagnostics are formulated.

The formulation of a constituency test above and the operationalization of these tests as variables often elicits a protest from certain linguists. It has been claimed that some of the tests used in this study are (or might be) 'junk' tests that should be discarded for linguistic analysis, usually on the basis of them not churning a clear result in favor of some or another syntactic theory. This point is actually partially valid. Many of the constituency tests developed in this book might very well be 'junk.' However, the protest misses an important point about database construction, measurement, and its relationship to hypothesis testing (Ackermann 1985: 125–149). By coding a constituency test in a database we are

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

not thereby claiming that the test necessarily identifies a constituent in any specific linguistic theory (let alone all theories). A linguist researching within a perspective whereby one of the coded tests is considered useless is free to discard the test and assess what the results show after they have trimmed the data of unwanted tests. What the database allows, or better yet, compels, the researcher to do is to be consistent and explicit about how such data are trimmed. For instance, they cannot discard a test in one language yet regard the same test as an important piece of evidence for their analysis in another. The worry has been that linguists treat a test as reliable only insofar as it confirms a given prejudged analysis and that they discard it otherwise (Croft 2001, 2010, Haspelmath 2011, Tallman 2021b) and constructing a database which samples tests independently of the researcher's analyses attenuates this problem. Another reason we think that the protest against junk tests misses the mark is because it presupposes that we know which tests are going to turn out interesting generalizations and which ones are not going to *a priori*. Further justifying this perspective is the fact that protests about junk tests are not mutually consistent. It turns out that one linguist's trash is another linguist's treasure, a point we return to in Section 6. Rather, whether a test turns out to be junk for language description or cross-linguistic generalization is an empirical question.

3 The structure of the database and use cases

The constituency tests and the planar structures are collected in an interlinked database based on AUTOTYP principles such as modularity, autotypology (see Section 2), separation of definition and data files, and late aggregation (Witzlack-Makarevich et al. 2022). As mentioned above, AUTOTYP is a typological database that has been continuously developed over the past twenty-five years as part of a large-scale research program in order to address problems that have arisen from the creation of more traditional typological databases. One of these issues is the use of fixed, *a priori* categories determined by theoretical considerations, or simply by traditional usage, that often fail to adequately capture a phenomenon across a large and diverse set of languages. The application of the AUTOTYP principles also facilitates the later re-use and expansion of the database. Another design principles concerns the separation of information across multiple files which are linked together via a common, standardized identifier. This flexibility makes it possible to address an larger number of different questions with one data set. As such, these design principles integrate well with the approach taken in this volume. The constituency test results are coded in a bottom-up fashion and we want to make the data usable for future studies.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The workflow for gathering the data and collecting it in the database is illustrated in Figure 1. It starts with the elaboration of the planar structure by the language expert based on data collected through fieldwork and collaboration with speakers. The planar structure then serves as the basis for applying constituency diagnostics as described in Section 2. The results are then written up, including discussion of issues with the methodology or application of specific tests that came up during analysis. Finally, the results are entered into the constituency database for cross-linguistic comparison. Given how autotypology works, the structure of the database and the variables are informed by the language-specific analyses and vice versa. In practice, this means that the database and variables are adjusted to accommodate language-specific facts not previously considered, but also that the exact application of a test in a language can be refined or adjusted based on what we learn from other languages.

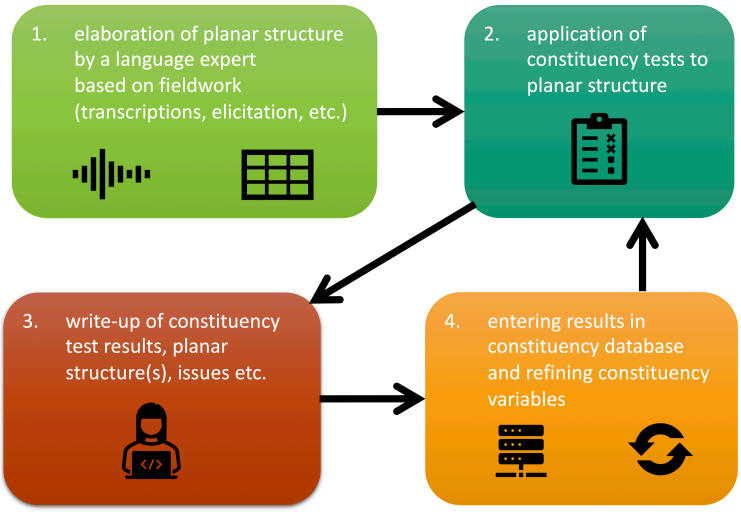


Figure 1: Schematic illustration of the workflow

The structure of the interlinked database is depicted in Figure 2. In the following, we will discuss the modules and the variables in more detail, following the outline from left to right and top to bottom. The **sources** file contains bibliographic information and can be linked to the metadata file with the citekey. The **metadata** file contains information about languages and contributors, such as commonly used language names, Glottocodes (if available), geographic information, as well as contributor names and the form of the contribution. The planar structures are collected in the **planar** file, where each planar structure and each

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

position within it receives a unique identifier. The positions are listed together with the position type (slot vs. zone) and the language-specific elements that can appear in each position. For analyzing convergences and other aspects of test domains, we need to know in which position the base of the planar structure occurs. This information is provided in the **overlaps** file, which can be linked to the planar file by the planar ID and to the other files by the language ID.

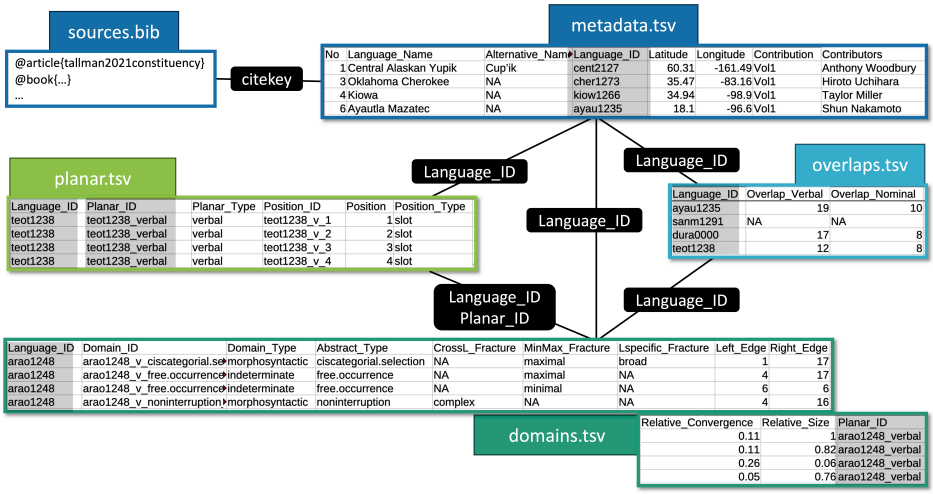


Figure 2: Illustration of the structure of the constituency database with file excerpts. Lines indicate which modules can be connected to each other. Black boxes show the unique identifier(s) that link(s) two modules together and the respective columns in the file excerpts are shaded in grey.

Finally, the test results are recorded in the **domains** file. This file can be linked to the other files via the language ID and additionally to the planar structure file with the planar ID. For each reported test in a language, we record the position numbers that delimit the respective span, as well as information about the type of test applied and measures derived from that, such as span size and the number of other tests the span converges with in this language. Below, we briefly summarize the contents of this file:

Domain Type: the linguistic level that the test applies to. Values:

- phonological: The test makes reference to phonological criteria. An example of this is a domain where vowel elision applies.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

- morphosyntactic: The test makes reference to morphosyntactic criteria. An example of this is domain delimited by elements that are ciscategorial with the verb.
- indeterminate: The test can be interpreted as either making reference to phonology or morphosyntax or both. An example of this is free occurrence, as it could be seen as resulting from a phonological constraint or a morphosyntactic one.

Abstract Type: standardized classification of constituency tests into abstract classes. Values:

- Ciscategorial selection: domains where the non-base elements are selectionally restricted to a specific base (e.g. verbal affixes which only combine with the verb).
- Deviations: domains where elements display a specific type of deviation from biuniqueness (e.g. extended exponence)
- Free occurrence: domains that identify spans which are free forms.
- Non-interruption: domains that cannot be interrupted by some element.
- Non-permutability: domains which exhibit fixed ordering of elements
- Segmental: domains that undergo some segmental phonological process.
- Suprasegmental: domains defined by some suprasegmental phonological processes
- Repair: domains that are identified by repair strategies.
- Pausing: domains that can be delineated by a pause
- Proform: domains that can be replaced by a proform
- Play language: tests that identify spans which are targeted in play language
- Idiom: domains which contain elements that typically form idioms or noncompositional constructions

Cross-linguistic fracture: a fracture that can be applied across languages with a standardized set of labels. Such fractures can be subtypes of phonological processes, for example, consonant and vowel deletion as fractures of a segmental domain. Our current data set contains 45 such fractures.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Language-specific fracture: a fracture that only applies within a specific language. Those fractures are thus not standardized. Our current data set contains 178 language-specific fractures.

Minimal-Maximal fracture: a fracture for the smallest and largest span where a test applies.

(Right/Left) Edge: The boundary of the span, i.e the first and last positions where the test applies. This is recorded by the position number.

Size and Relative size: The size of the span in number of positions and the relative size of the span in number of position divided the by the largest span in the respective language.

Convergence and Relative convergence: Number of other spans in the language that this span converges with. The relative convergence is the convergence number divided by the total number of tests applied in the language.

Largest: The largest span identified in a language.

Position total: The size of the planar structure in number of positions.

Tests total: The total number of tests applied in a language.

Due to the modular structure of the database it can be easily expanded upon in the future. The data collected in this volume are available on Zenodo as version 1.0 (https://doi.org/10.5281/zenodo.10076550) (Auderset & Tallman 2023).

The database is designed in such a way that it can be used for investigating a variety of research questions and for providing overviews and summaries regarding constituency. We provide a few examples relevant to the volume here. The sample languages are plotted on a map in Figure 3, which additionally displays the maximum relative convergence found in each language. The map shows that Cup'ik displays the highest relative convergence, while Chorote has the lowest one. It also shows that even languages spoken in the same geographical area, such as Hup and Yukuna, do not necessarily exhibit the same degree of convergence.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

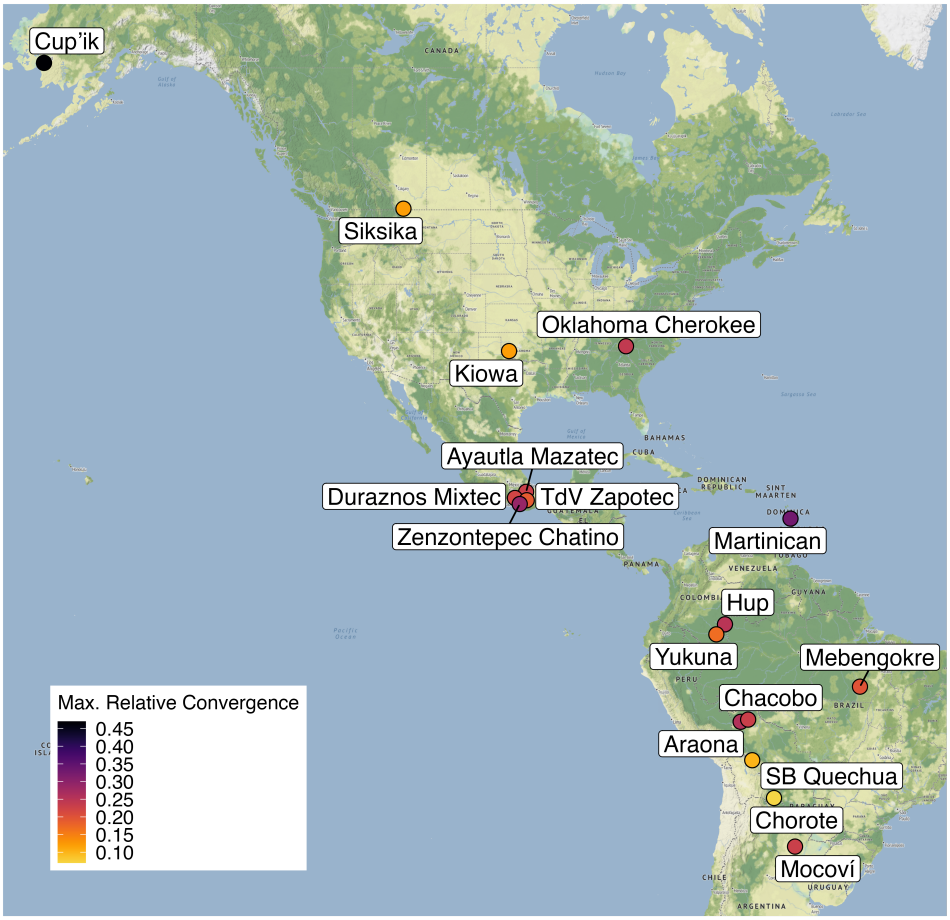


Figure 3: Location of the sample languages with maximum relative convergence (= the maximum number of test convergences per language divided by the total number of tests) represented as color gradient.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The database also allows one to compare layer sizes and convergences across languages. Figure 4 displays relative convergences versus relative layer size and shows that there is great cross-linguistic variation in this domain. In terms of relative layer size, most of the languages described here have spans of various sizes, ranging from targeting only one position to the whole planar structure, as in South Bolivian Quechua and Chorote. In others, the spans identified by the constituency diagnostics cluster around a few layer sizes, as in Martinican, or are skewed to either relatively small spans, as in Kiowa, or relatively large spans, as in Oklahoma Cherokee. In terms of relative convergences, the languages also exhibit vast differences. In a few languages, a clear 'winner' emerges, that is a span that is identified by many diagnostics, while all other spans show no or very little convergences. This is the case for example in Cup'ik, where almost half of the diagnostics converge on a span of size 0.79 (covering 15 out of 19 positions). Martinician and Zenzontepec Chatino both have spans that are targeted by about a third of the diagnostics, but these are much smaller. In Zenzontepec Chatino, the span is of size 0.19 (covering 4 out of 21 positions) and in Martinican it is even smaller at 0.16 (covering 4 positions out of 25). This shows that even if we find cross-linguistically strong convergences, the spans that emerge from them can be very different depending on the language. Furthermore, in some languages, there are no strong convergences at all, as in Chácobo, Hup and Siksika (Blackfoot). These languages approach a situation where each test targets a different span.

The database can also be used to explore tendencies of certain test types across languages. Figure 5 displays the distribution of relative span size according to the type of constituency test. Many of the test types have similar bimodal distributions, with a larger peak targeting a smaller span and a smaller peak targeting a larger span. This reflects the minimum and maximum fractures. Deviations from biuniqueness, however, exhibit a different distribution: they overwhelmingly target small spans (peak around 0.15), with very few tests resulting in larger spans above 0.5. This could explain why deviations from biuniqueness are often seen as good wordhood tests – they capture almost exclusively small spans that can easily be interpreted as 'words.'

4 The index of synthesis reconsidered

Classically languages are described as varying in terms of their degree of synthesis. The degree of synthesis that a language displays refers roughly to its tendency to pack more or less concepts into a single word (Sapir 1921). For typological comparison the notion has been operationalized by counting the number

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

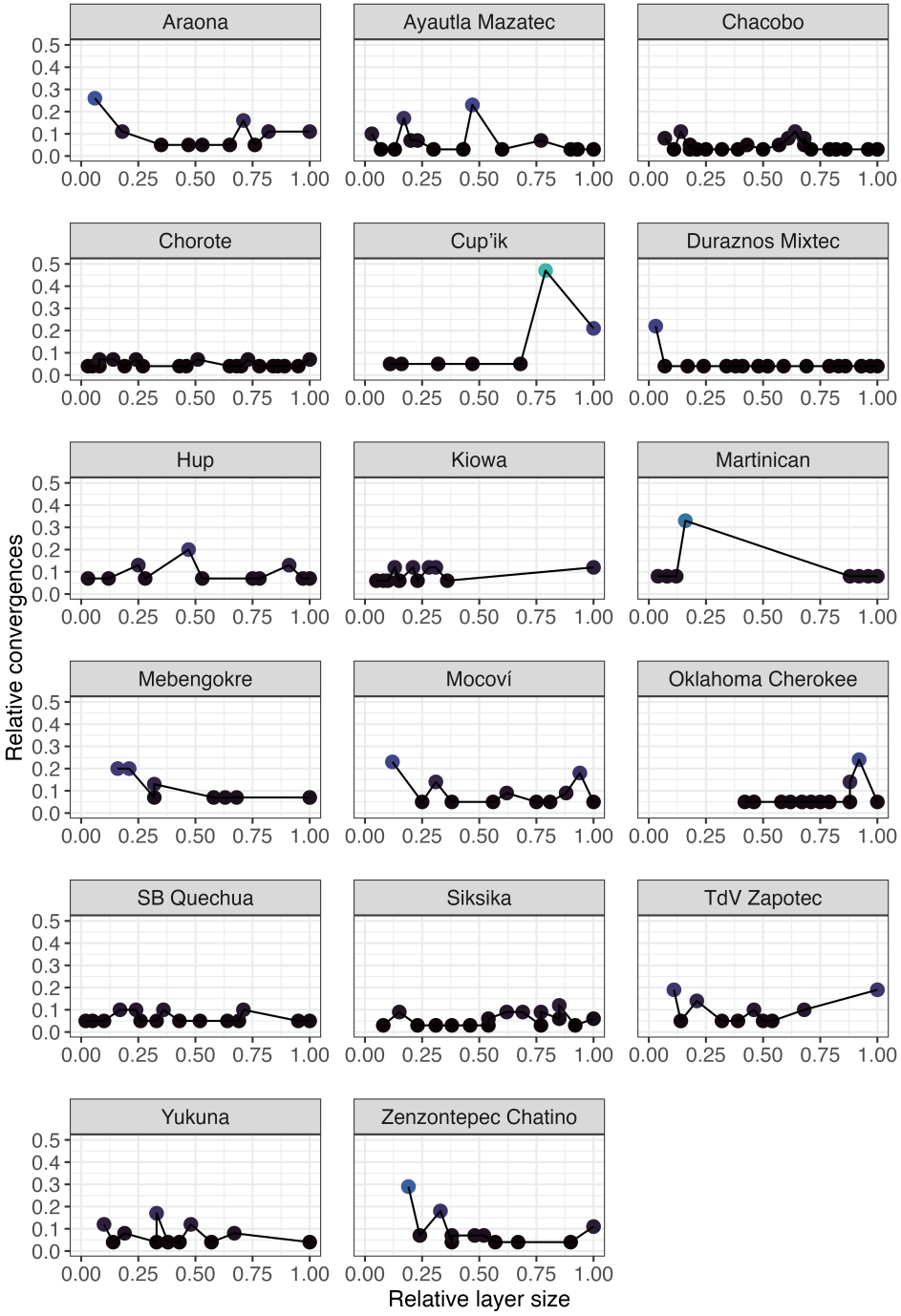


Figure 4: Visualization of relative convergences per relative layer size across the languages of the sample in the verbal domain

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

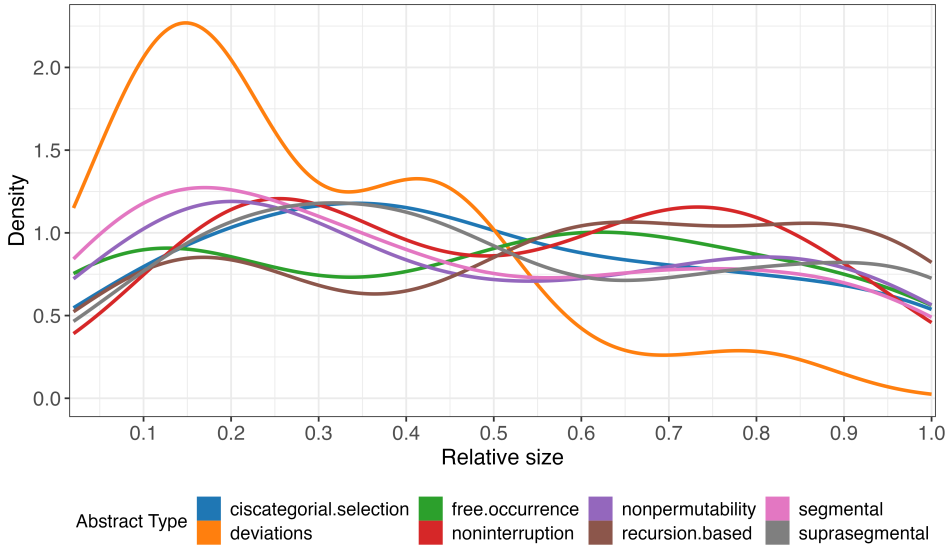


Figure 5: Density of relative size of spans by abstract type across sample languages

of segmentable morphemes that occur within each orthographically spaced out word on average over some text (Greenberg 1954, Easterday et al. 2021). Such studies rely on orthographic words and they rest on the assumption that either orthographic words are legitimate units of comparison or are approximations to some unit of comparison across languages. The analytic-synthetic continuum also forms an important aspect of describing variation and change in certain language families (e.g. Schwegler 1990, Ledgeway 2017 for Romance; Arcodia & Basciano 2020 for Sino-Tibetan).

The results of the studies of this volume highlight the fact that the synthetic status ascribed to a language might be contingent on which constituency tests are deemed to be appropriate wordhood diagnostics. There is no unified notion of synthesis, but a multiplicity of different notions that may or may not align in a given language.

Even in languages with a high degree of convergence it should be noted that not all diagnostics pick out what is traditionally considered a word. Clear examples comes from Cup'ik and South Bolivian Quechua, which have both been described as polysynthetic languages in the literature. However, if we take the criterion of 'conventionalized coherence' as definitional Dixon & Aikhenvald 2002a, Dixon 2010a, Aikhenvald et al. 2020, both of the languages are much closer to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

being analytic. These languages contain pockets of word-like chunks or clusters within their traditionally defined words. This is not because linguists working with these languages have simply ignored wordhood criteria. On the contrary: free occurrence, noninterruption, phonological criteria such as 'stress', or syllabification all hit a domain of structure attached to the notion of 'word' used by Inuit–Yupik–Unanganists and Quechuanists respectively.

The observation that so-called polysynthetic languages have word-like pockets inside their grammatical words is not a theoretically innocuous observation: many morphologists and researchers in corpus linguistics propose that morphological structure emerges as a distinct component from syntax via 'chunking.' (Bybee 2001, 2010, Lorenz & Tizón-Couto 2019), the process whereby multiple pieces of structure are gradually reinterpreted as a single unit for processing and production, presumably the basis of 'conventionalized coherence.' If such a theory of morphological development and maintenance is maintained, then it follows that the traditional 'word' in these languages is a phrasal (or post-word) constituent. But the convergences of wordhood tests around this domain still provide evidence for dichotomous structuring of some sort. Insisting that morphology is defined through conventionalized coherence does not result in the purported 'morphological complexity' of polysynthetic languages disappearing but simply displaces it to a different terminological realm: we would now claim that many polysynthetic languages display dichotomous patterning in their 'syntax.'

Other languages pose even starker problems for the traditional analytic-synthetic distinction. It can be observed that in Chácobo, Duraznos Mixtec and Hup, a shift in perspective regarding which criterion we regard as defining the word, can result in these languages being categorized as isolating or (poly)synthetic. In both Chácobo and Hup, a focus on nonpermutability (contiguity, fixedness of order) and certain interpretations of noninterruption would result in the classification of these languages as isolating or at least highly analytic. If we shift our focus to free occurrence domains, the languages become (poly)synthetic, and the facts that were rallied to argue that they were isolating now become indicative of the languages displaying a 'syntax-like' morphology (Payne 1990, Tallman & Epps 2020). Moreover it is too simplistic to claim that this is a difference between only two types of criteria free: occurrence vs. nonpermutability or noninterruption. Mixtec is isolating or polysynthetic depending on how free occurrence is treated as a constituency test, the minimal fracture providing an isolating result and the maximal fracture providing a highly synthetic result. We are reminded of Boas' observation that in some languages (Tsimshian was his example) the division or combination of forms into separate or single words can be fairly arbitrary (Boas

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

1911: 28), but importantly languages may vary in terms of how arbitrary this division is (Boas 1911: 26; see Bazell 1953: 68 as well).

Claims about synthetic status usually make reference to morphological complexity (e.g. Easterday et al. 2021). But synthesis could also be discussed in terms of phonological domains - in terms of segmental morphemes per phonological word. This approach would run into the same problem, however, as there are competing definitions of the phonological word for many of the languages of the study. The notion of a phonological word is not unified behind a single criterion either and so couching synthesis in terms of phonological integration does not necessarily simplify the notion of synthesis.

These considerations do not mean that the analytic-synthetic notion should be abandoned for typological research, but rather that it should be least refined. As a language has less and less converging wordhood criteria the notion of synthesis becomes more complex and graded in that language. In this way we could understand the index of synthesis that is not only multidimensional (as it can be decomposed into a number of logically distinct variables) but as one which interacts with other architectural properties of a language, as in how strongly the languages displays dichotomous patterning or how fuzzy the boundary between morphology and syntax is in the language (e.g. Tallman & Epps 2020 for this perspective).

5 No a priori wordhood tests

In a sense, the notion that there are wordhood tests presupposes that there are words to begin with (Lara 2004). If we claim that wordhood tests are not always picking out a unified notion of word, then what are these wordhood tests picking out? The apparent paradox is resolved once we recognize that words are a species of constituent which we assign special status because it represents a cut-off point between two different realms of structural organization. From this perspective it is somewhat misleading to even refer to 'wordhood' tests as such. Rather, if the whole idea of 'word' is interesting because it indexes our belief in the modular[3] structure of grammar, then wordhood emerges from patterns of structural groupings over utterances, not from singular diagnostics applied in the abstract.

[3]Note that claiming that languages display modular structure does not entail that the modular structure is innate, nor that there are some fuzzy boundaries between domains. In cognitive science and biology generally it is well recognized that modularity is a matter of degree (Rasskin-Gutman 2005; Carruthers 2006: 14) and that it can be emergent (Coltheart 1999, Zerilli 2020).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

There are reasons to think there should be no coherent notion of 'wordhood test', as distinct from phrasehood test, at least not a priori.

The fact that there is no clear distinction between wordhood and phrasehood tests can be discerned in two ways. First, we can consider the fact that when we put formulations of wordhood and phrasehood tests side by side, we find that they are difficult to distinguish. Tallman (Introduction This Volume) gives the examples of noninterruption as a wordhood test versus displacement as a phrasehood test.

We can also discern that the lack of clear distinction between wordhood and phrasehood constituency test by considering the fact that in numerous cases a diagnostic that hits a 'word' according to its definition in one language (or linguistic tradition), hits a subword unit in second language, and yet an apparently phrasal unit in a third. For instance, noninterruptability by a free form lines up with the traditional word in Cherokee (the orthographic and what is considered by a word by Iroquianists) (Uchihara This Volume). The same is true of noninterruptability in Martinican (Duzerol This Volume). The derivational prefix, the verb root and two pronominal indexes make up the orthographic word in this language as long as the pronominal indices are second or third person. However, if take the way the word is described in Araona (and the Takanan tradition generally) the same interruption test identifies a subword unit, in fact, just the verb root, rather than the large polysynthetic structure described as a word by some linguists who have described the language (Pitman 1980, Emkow 2019). The converse problem is also attested. Noninterruption by a single free form identifies a span of structure *higher* than what Gutierrez & Uchihara argue is the best candidate for phonological wordhood in Teotitlán Zapotec. Therefore noninterruption by a single free form identifies a word, a subword or phrasal domain depending on the language. Should we still consider noninterruption a 'wordhood test?'

Another example is extended exponence. In Araona, extended exponence lines up with the traditional word, but in Central Alaskan Yup'ik, the same diagnostic identifies a subword constituent with respect to the traditional word of this language. Again, should extended exponence be identified with a word or a subword?

In the phonological domain these issues are so endemic that it is difficult to know where to start. In their study, Bickel et al. 2009 show that there is no overall tendency for phonological domains to cluster within one another around a universal 'prosodic word'. Furthermore, once prosodic words are classified for the type of phonological generalization that defines them (e.g. rhythm, epenthesis etc.), there is no overall tendency for any specific phonological process to

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

identify higher or lower domains, except for 'stress' which shows a tendency to identify higher one (Schiering et al. 2012).

'Words' refer to boundaries between domains of different structural organization. But it is doubtful that a 'wordhood test', abstracted from the rest of the structure of a language, is a useful starting point for typological investigation. Constituents, domains or groupings are a better starting point since they do not presuppose that we know *a priori* the properties of the modules we are interesting in investigating, which may be subject to cross-linguistic variation.

6 Reliable and unreliable tests

6.1 Introduction

The literature on wordhood and constituency often implies that certain tests are better or more reliable than others. For instance, Dixon & Aikhenvald 2002c distinguish certain 'main criteria' (cohesiveness, fixed order, conventionalized coherence) as distinct from others. But the test of 'isolatability', for example, only identifies words as a matter of 'tendency' (Dixon & Aikhenvald 2002d: 25). Similarly, Payne (2006: 162) claims regarding coordination that it "can't be the major way of determining constituent structure", compared to the other constituency tests he discusses (Adger 2003: 125 and Carnie 2010: 21 for related claims).

Writers on these topics apparently do not agree with each other. Dixon & Aikhenvald (2002b: 25) state that "the principle of uninterruptability .... is only a tendency - which may apply more to phonological than to grammatical words - but can be a useful support for the other criteria". Bauer (2017: 17) has a discussion concerning "criteria involving structural integrity", which appears to be similar if not the same as noninterruption. He makes nearly the opposite claim regarding the reliability of the wordhood test: "The uninterruptability of the word is, in general terms, a much stronger criterion". Martinet 1962: 92 states "[a]s a matter of fact, inseperability is one of the most useful criteria for distinguishing what is formally one word from what is a succession of different words"(see Brown & Miller 1980: 164–165 as well). Booij (2005: 185-187) describes noninterruption as definitional of word constituents. Some of the apparent disagreement could be a result of authors interpreting the criteria in different ways[4], but the point

[4]For instance, Dixon & Aikhenvald (Dixon 2010b) make a distinction between cohesiveness and noninterruption that the other authors do not make to my knowledge. Noninterruption seems to also involve a pause, whereas 'cohesiveness' is the more general term for any non-interruptable piece of structure. The ambiguity regarding how to interpret the diagnostics as they are formulated in the literature is perhaps one of the reasons why it appears so difficult

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

remains that there is a reoccurring tendency to regard some tests as better or more reliable than others in some sense, yet it is unclear from the literature which ones should be regarded as more reliable.

It is worth asking on what basis such claims about the relative reliability of tests could be made. In the literature the relative superiority of some tests over others is generally asserted without any justification. In some cases it is pointed out that the test is unreliable because it does not converge with a predefined or established constituency analysis (e.g. Payne 2006: 162 Carnie 2010: 21), which appears to be a circular argument. More charitably what some of these researchers might mean is that unreliable tests are just those tests that are prone to not be applied correctly (presumably by linguists who are not as skilled at syntactic analysis as they are). Yet an articulation of the proper interpretation of a potentially unreliable test is never given, except insofar as it means "in line with my own theoretical expectations."[5]

In the context of the literature on word identification, we might speculate that the widespread sense that there are some tests which are better than others is probably based on how well a given test lines up with prescriptive orthographic conventions within some speech community. Given that prescriptive orthographic practices are socially constructed (not all languages/speech communities have them), it is not clear that they would correlate to the *same* degree with the *same* diagnostics cross-linguistically. Disagreements between linguists with respect to the reliability of some diagnostics together with the widespread feeling that some tests are better than others might be a reflection of the languages (or perhaps even constructions in languages) that these linguists are most familiar with and the degree to which the orthographic conventions of these languages line up with this or that diagnostic stated in literature.

We might, however, consider 'convergence' to be a more empirically grounded, and perhaps theoretically grounded, way of assessing the relative reliability of tests. The convergence of logically distinct diagnostics has been used to justify the validity of 'word' and 'phrase' as valid linguistic units, as the quotations from Matthews 2002 and Levine 2017 below illustrate respectively.

to refute them. If I find that a diagnostic is not working, I can be accused of misinterpreting it. Indeed as we have shown throughout the papers of this volume, the diagnostics have multiple interpretations.

[5]In the context of coordination tests Phillips states "Traditionally, the results of movement tests have tended to be taken more seriously, and the results of other tests have been made to fit with these." (Phillips 1996: 27). As Phillips shows, one ends up with a quite distinct view of constituency structure if coordination is put on par with the other tests (see Osborne 2018 as well for relevant discussion).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

> *For words:*
> 'The issue for us is whether the word has any other status. No criterion is either necessary or sufficient, as Bazell ... made clear long ago. But they are relevant insofar as, in particular languages, they do tend to coincide. A form which is cohesive need not logically consist of elements whose order is fixed.' (Matthews 2002: 276)

> *For phrasal constituents:*
> 'The two phenomena which appeal to unithood must, in other words, be fundamentally independent. Normal methodological considerations then make it highly unlikely that the joint appeal to syntactic unithood from two independent sources envisioned here arose from coincidence.' (Levine 2017: 13)

If we work our way backwards from such statements, then tests are reliable insofar as they tend to converge, because insofar as they tend to converge they are identifying (abstract?) constituents.

In what follows we attempt to assess the relative reliability of certain tests by assessing the degree to which they converge with tests in general. We report two findings: (i) there are some clear correlations between certain specific tests (e.g. free occurrence and segmental phonological processes); (ii) there is no overall tendency for any constituency test to be more reliable than another as judged by convergence. What this means is that for some tests, one can predict with some degree of accuracy what other tests it is likely to converge with. However, for a given test one cannot say whether it is more likely to converge than any other test in general. In the former case, where possible, we point to some fairly straightforward functional motivations which have been pointed out in the literature. Overall the results suggest that edges ('junctures', 'boundaries') might be a source of more meaningful generalizations as opposed to span defined units such as 'word' or 'phrase'.

6.2 Results – correlations between domains

Before presenting the results, some remarks regarding comparison are in order. The comparison of word/constituency tests cross-linguistically is complicated by a number of factors, two of which should be mentioned. First, we can compare constituency convergence in terms of convergence at individual edges of structure (e.g. left or right edge) or at both edges simultaneously. We will refer to the former as 'edge convergence' and the latter as 'span convergence.'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Secondly, constituency domains can be compared on different levels of abstraction. For instance, we could ask how well noninterruption, regardless of whether and how it is fractured into subtests, converges with domains related to accent/stress marking. If we wanted to get more granular we could ask how well noninterruption by a single free form converges with the minimal fracture of an accent-based domain. We will, therefore, be presenting results at different layers of abstraction.

We exclude discussion of some tests types that only have one example in our data set (e.g. 'language play' in Zenzontepec Chatino).[6] We note that our results are preliminary as they only contain 463 test results from 17 languages. Furthermore, future research might involve applying and or further operationalizing more constituency domains which could change the results.

Below, we will also ignore fractures of recursion-based diagnostics such as those based on whether the marking is syndetic or asydetic etc. This is done in order to simplify the discussion.

In what follows we assess the relationships between individual domains using correlation matrices. Correlation matrices present the correlations between different tests. In order to present these correlations all variables are coded as binary variables. We use Kendall's tau as our correlation metric.

The meaning of a correlation metric in relation to constituency test convergence requires some commentary. Imagine that we have two tests x and y. If x always converges with y, the correlation coefficient will be 1. If these tests never converge with one another, the correlation coefficient will be -1. Constituency domains which tend to correlate with one another will, therefore, show positive correlations. Note that two constituency tests can be non-convergent on their spans, but convergent on their edges.

For expositional reasons we present correlations in terms of percentages (i.e. a 0.5 correlation is presented as '50'). In order to focus our discussion we will focus on those correlations which are above 0.1 in significance. Those which do not meet that threshold are crossed out (with an 'X') in the correlation matrices below.

The correlation matrix in Figure 6 shows the correlations between tests in terms of span convergence. The correlation matrices in Figure 7 and Figure 8 provide correlations for left and right edges. These figures provide overviews of the tests coded by 'Abstract Type.' This means that the results pool fractures of constituency tests (e.g. minimal and maximal domains of free occurrence are

[6]This does not mean that we think this test is irrelevant. Rather it means that future research is needed in order to compare the relevant domain cross-linguistically.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

coded together).[7] Note that there are very few positive correlations when considering span convergences. In fact, tests at an abstract level are more likely to misalign than not if we consider only span convergence.

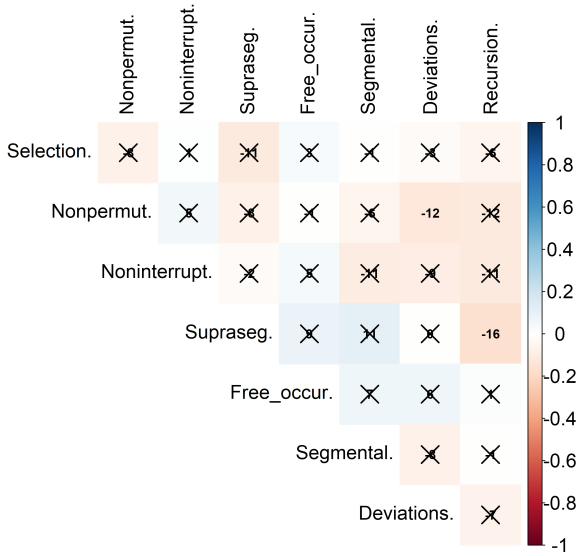


Figure 6: Span convergence for domains organized according to 'Abstract Type'

The notable generalization to observe from the matrices in 6-8, is that there are no obvious positive correlations between domains when the left and right edge of the domains are considered together. When we consider span convergence, therefore, tests in the abstract are more likely to converge than not. The correlations become positive in the aggregate and statistically stronger when we consider edges by themselves. Thus, there is a strong difference between span and edge convergence. On the left-edge convergences, there is a relatively strong correlation between noninterruption and phonological domains defined by segmental processes. For right-edge convergence, domains defined by segmental and suprasegmental processes tend to align more strongly than other domains.

[7]The plot contains short-hands for a number of tests: (i) Deviations = 'Deviations from biuniqueness'; (ii) Noninterrupt. = 'Noninterruptaility'; (iii) Nonpermut. = 'Nonpermutability'; (iv) Free_occur. = 'Free occurrence'; (v) Selection = 'Ciscategorial selection'; (vi) Segmental. = 'Segmental phonological processes/domains'; (vii) Supraseg. = 'Suprasegmental phonological processes/domains'.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

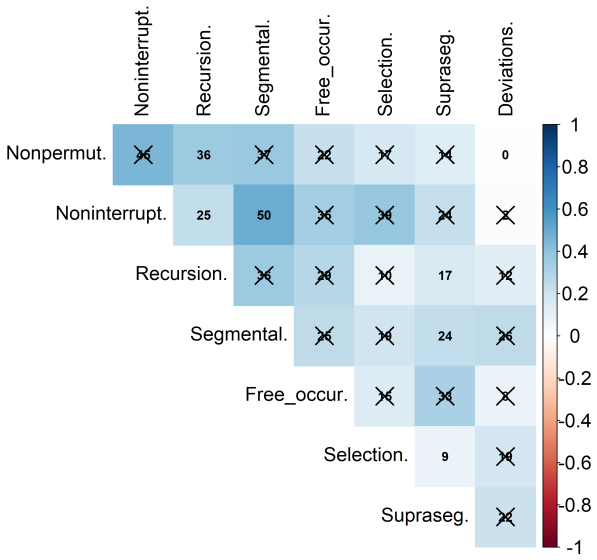


Figure 7: Left edge for domains organized according to 'Abstract Type'

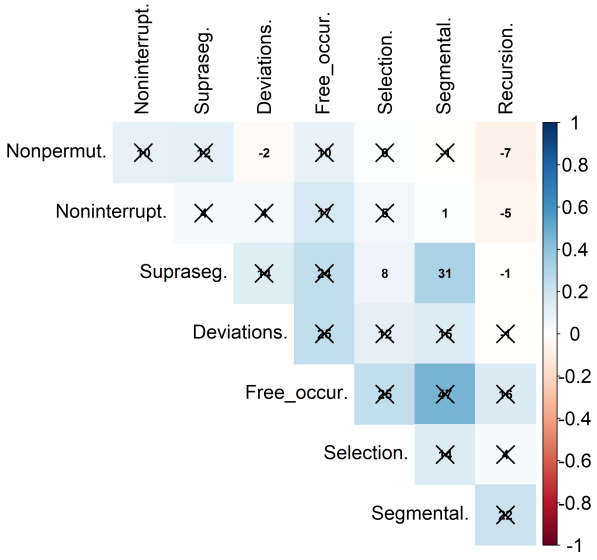


Figure 8: Right edge for domains organized according to 'Abstract Type'

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

On diachronic grounds it is not surprising that suprasegmental and segmental processes should line up on an edge. The presence of a prominent syllable can result in segmental changes over time (Bybee et al. 1998), for instance, but prominent syllables are almost always attached to the edge of their domains.

Overall the domains reflected in 6-8 are perhaps too abstract to develop very specific explanations. Next we consider more fine-grained domains. We consider domains fractured according to whether they are minimal or maximal, where this fraction is available. If not, we break apart tests by highly frequent cross-language fractures. In the case of noninterruption this means the distinction between simplex, complex and multipositional interruptors. In the case of nonpermutability, we break them apart according to whether the tests are scopal or non-scopal interpretations.

When we consider span convergence, correlations are relatively weak even where the domains are fractured and made more specific as illustrated in Figures 9 and 10 for minimal and maximal domains respectively. In these cases, the strongest correlation arises between the minimal fracture of the free occurrence domain and minimal domains of segmental phonological processes, and minimal domains of recursion based diagnostics. A freely occurring unit of structure has a weak tendency to also contain no elements that can wide-scope under any type of conjunction. Also, segmental phonological processes perhaps display a weak tendency to blocked around such domains as well.

There are no strong tendencies with the maximal domains, however. The weak tendencies that are present below are just as likely to be negative.

Once again, when we consider edge convergences, stronger relationships appear as can be seen from Figures 11 through 14. Nearly all of the variables are highly associated with each other on the left-edge for minimal domains with the exception of the scopal fracture of nonpermutability. The convergences are somewhat weaker on the right-edge.

Maximal domains overall tend to have lower convergences than the minimal ones. A few chapters of this volume have suggested that maximal domains might be more likely to indicate phrase-level structures. If convergences are more likely to hit edges of structural shift from morph to utterance (i.e. words), this difference is potentially understandable.

We find again that free occurrence has a relatively strong tendency to converge on the edges of certain domains, despite displaying a low span convergence.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

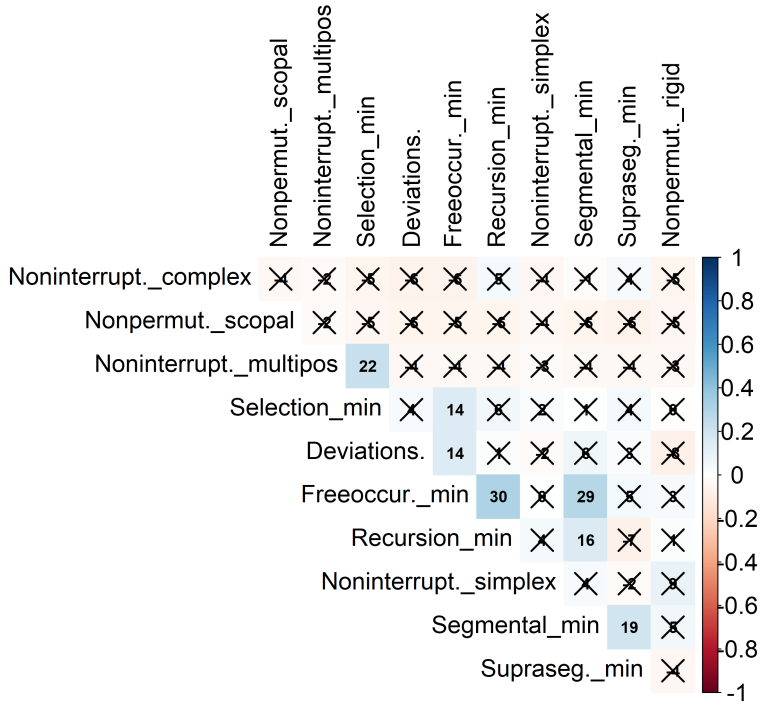


Figure 9: Span convergences for fractured tests (minimal and cross-language fractures)

6.3 Results - predicting convergence

In this section we attempt to discern whether there is any overall tendency for some domains to converge more than others. First we need to discuss some metrics of convergence.

One can discern the relative importance of domains based on how often they converge with other diagnostics. Each coded domain or test result can be coded with a an **absolute convergence** number. If a domain convergences with no other tests in a language, its absolute convergence is 1. We assign each domain a **relative convergence metric** by language. This takes the absolute convergence level and divides it by the total number of tests applied in a language. Thus a domain which converges with no domains in a language for which 10 tests were applied has a relative convergence of 0.1. In a given language the relative convergence level is perhaps a more accurate metric of the convergence strength of a given

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

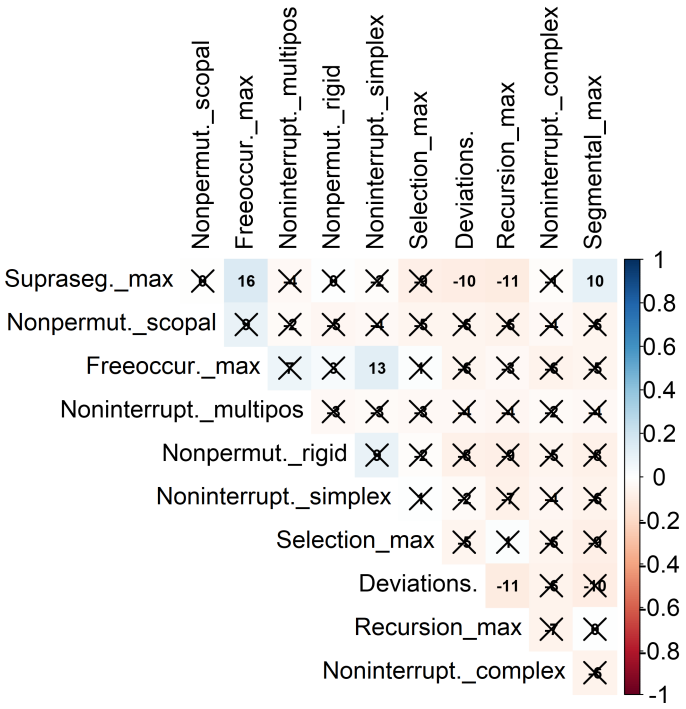


Figure 10: Span convergences for fractured tests (maximal and cross-language fractures)

test. The reason is that we expect overall convergence to increase as a matter of chance as the number of tests increases in a given language (Tallman 2021c).

Note there are three types of absolute and relative convergence: span convergence, left edge convergence and right edge convergence.

Figure 15 provides density distributions showing span convergence (pink), right edge convergence (green) and left edge convergence (blue). Span convergence is heavily skewed leftwards. Most domains do not span converge. Right edge convergence is less skewed to lower relative convergence values, and left edge convergence presents something approaching a uniform distribution (or else weakly distinguished bimodality).

One way we can discern whether certain domains are more convergent than others would be through comparing their distribution along relative convergence compared to the distribution of all the domains pooled. A more convergent test would have its distribution more skewed to the right compared to the distribu-

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

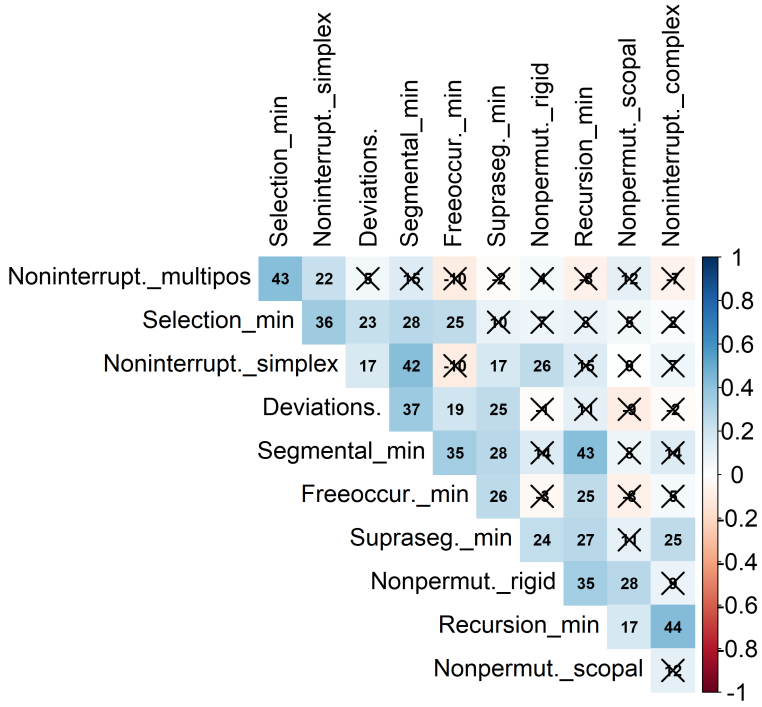


Figure 11: Left edge convergences (minimal and cross-language fractures)

tions of the domains as a whole. Figures 16 through 17 suggest that none of the tests are obviously more convergent than any others as overall they all seem to display the same left-skewed distribution.

A statistical test of reliability might attempt to predict convergence level from the domain type. A more reliable domain should predict a higher convergence level than a less reliable one.

For span convergence a random forest was constructed (iterations =1000) in order to assess whether any of the domains might be predictors of convergence level. In this model absolute convergence was a predictor and the classification of different domains at all levels of abstraction that are implied by fracturing by the Cross-language fractures and the minimal-maximal fracture were used as predictors together with the prosodic-word domain classifications. The random forest always predicts level 1 convergence for all domains. The baseline classification rate for the random forest is 0.4060475 and the accuracy is also 0.4060475.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

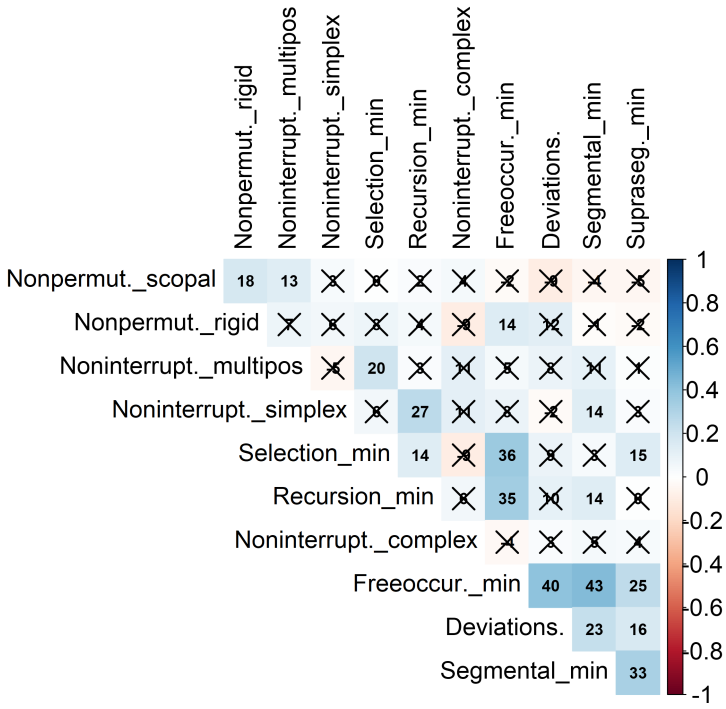


Figure 12: Right edge convergences (minimal and cross-language fractures)

Thus, The model does not predict better than chance when the baseline is taken as chance.

7 The word bisection dogma

Another hypothesis that the data structures developed through this volume can test is the (empirical) word bisection dogma. In the introduction to this volume, Tallman (This Volume) notes that there are two versions of the word bisection dogma. The fiat-based word bisection dogma states that a universal distinction between morphosyntactic and phonological words can be maintained because diagnostics for the relevant constituents can be concocted. There is no sense in arguing about this claim because it has the status of a tautology. The empirical word bisection dogma is more interesting because it maintains that the relevant

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

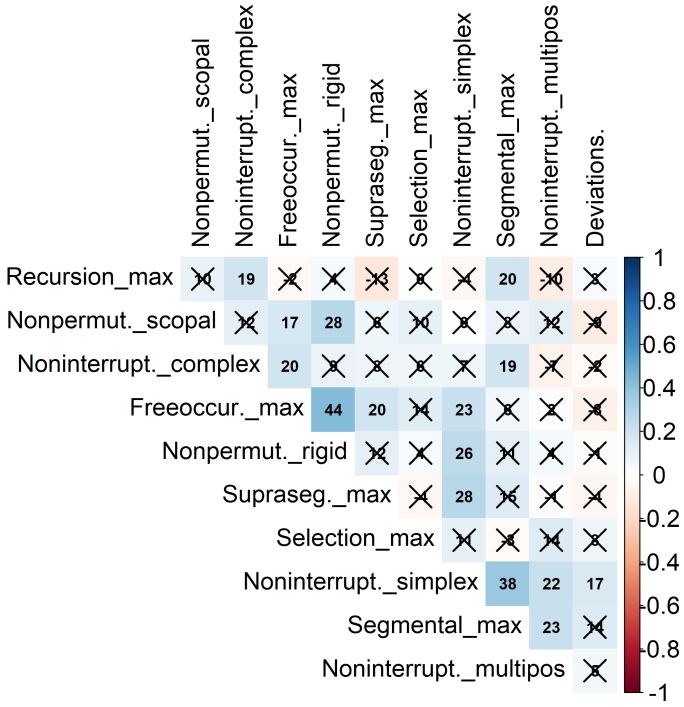


Figure 13: Left edge convergences (maximal and cross-language fractures)

diagnostics tend to converge with one another to support the morphosyntactic versus phonological word dichotomy in language after language.

Tallman (2021c) attempts to test the empirical word bisection dogma with data from Chácobo. He shows that there are few convergences within morphosyntactic domains and within phonological domains. The paper attempts to articulate a the word bisection dogma as a falsifiable hypothesis concerning the (mis)aligment of wordhood tests – phonological and morphosyntactic tests may misalign with other, but morphosyntactic tests should tend to align with other morphosyntactic tests and phonological tests should tend to align with other phonological tests. Based on this methodology, convergence between tests is not meaningful by itself, however. As the number of tests increases, the probability that two or more tests align by chance increases. Some notion of 'chance convergence' has to be constructed in order to assess an empirically contentful notion of word bisection thesis.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

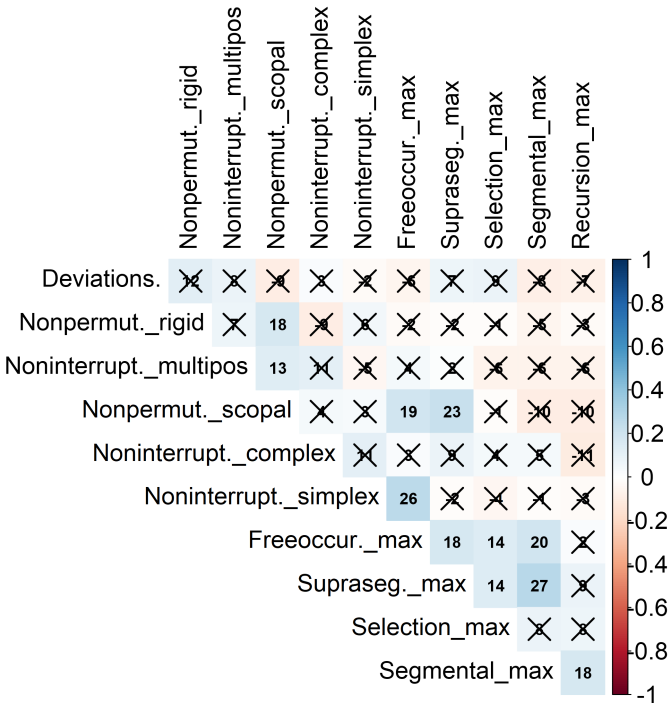


Figure 14: Right edge convergences (maximal and cross-language fractures)

Hypotheses which are falsifiable in principle are not necessarily falsifiable in practice if methods cannot be designed to test them. Much work of the serious work in the natural sciences is not narrowly concerned with theory construction, but also in designing experimental ideas, analytic techniques and new kinds of technology that can be used to test hypotheses (Hacking 1983: 214; Mayo 2018). Tallman uses a methodology for calculating chance convergence between wordhood tests that relies on a simulated null distribution. The results suggest no support for the version of the word bisection hypothesis he constructs.

But there are a number of problems with the methodology. First, there was no way of assessing whether some tests should be considered more reliable than others. Secondly, the simulation method has some unrealistic aspects to it as it depends on each test being independent, a logical impossibility in the context of the planar-fractal method (minimal fractures cannot converge with maximal fractures by definition). The testing methodology is also overly simplistic in another

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

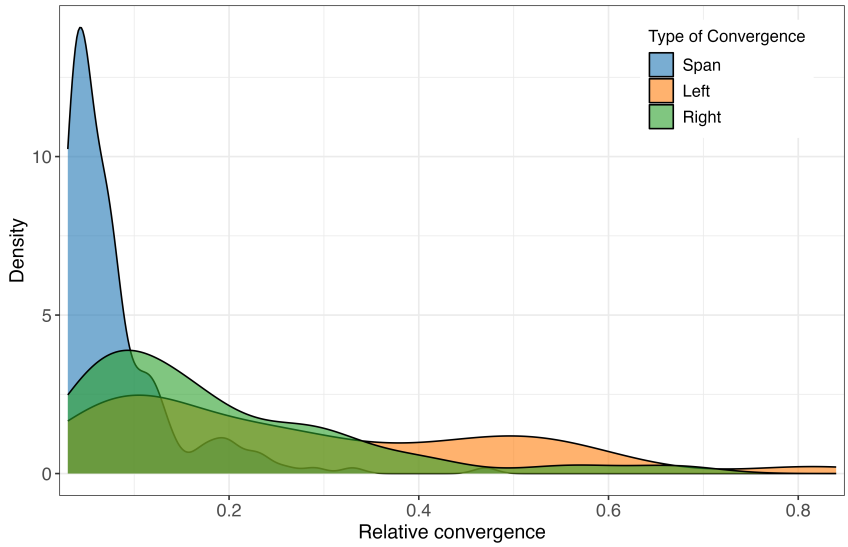


Figure 15: Density distributions of relative convergence at the right and left edge and across the whole span

sense. We expect constituency diagnostics to display a trend towards nesting, but the simulation method assumes no such trend or constraint.[8] Finally, only one language is assessed (Chácobo, Pano), which may not be generally representative of cross-linguistic trends.

The current database allows us to give a first pass assessment of the word bisection hypothesis with more languages. Ideally a method would also be used to construct chance probability, but we will leave that for future research. Here we will present simpler metrics that can be derived from basic arithmetic. There are two main results from the current study that we wish to emphasize: (i) there is interesting language variation with respect to how strongly convergent word constituents are supported (see Figure 4 above). (i) while there are some constituents that are strong word candidates for 'word' in terms of convergence, cases where morphosyntactic *and* phonological words appear to be motivated based on convergence are less common and/or less obvious.

Figure 19 displays the relative convergence levels for phonological tests across planar structures of this study. Each language is coded with its Glottocode and then an indicator for whether the planar structure is verbal or nominal. The y-axis has the absolute convergence level and the x-axis contains the relative layer

[8]To be fair it would be difficult to articulate what such a constraint should look like in the absence of typological data.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

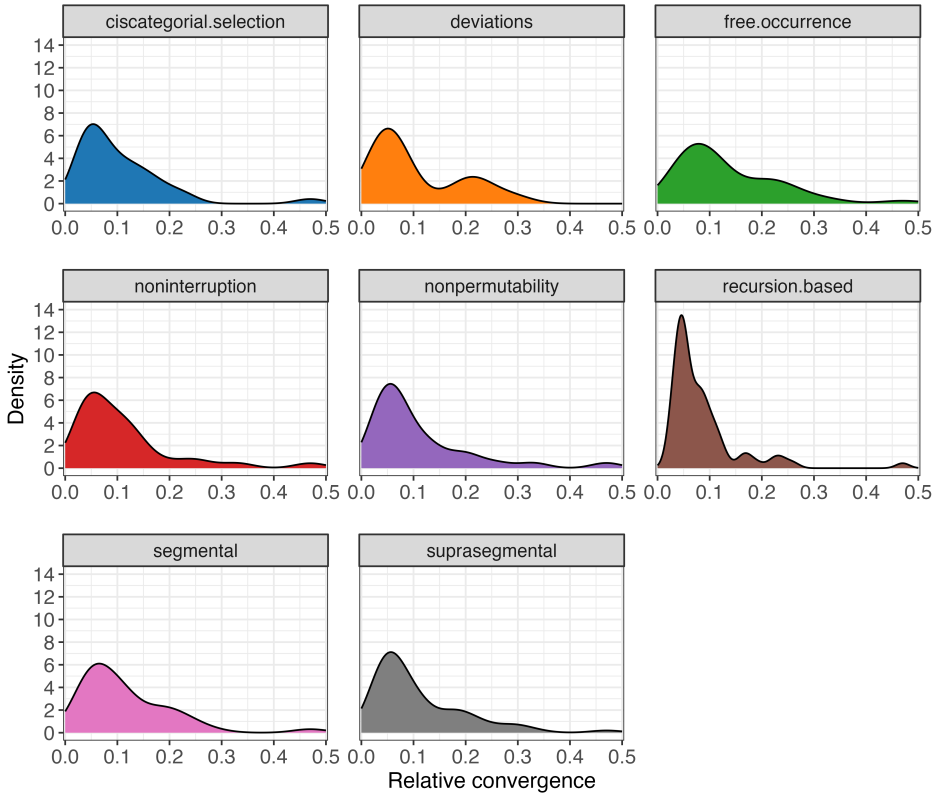


Figure 16: Density distributions of abstract types on relative span convergence in the verbal domain. Types with fewer than 5 data points are excluded.

size. We can give a preliminary assessment of the strength of a wordhood proposal based on a combination of absolute convergence and the number of tests that were applied in each language. An ideal case where phonological wordhood would be supported would a spike upward (high convergence) in relation to a relatively low number of tests applied.

To the extent that convergence supports phonological wordhood, the strongest case appears to be the p-word in the verbal domain of Zenzontepec Chatino (with 5 convergences). The Chácobo nominal domain also displays some evidence for phonological wordhood (contrast this with the verbal domain, Tallman 2021c). The case of Teotitlan del Valle Zapotec is somewhat difficult, because although there are a relatively large number of covergences these appear in a domain that most authors would consider to be an utterance phrase (Gutierrez & Uchihara

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

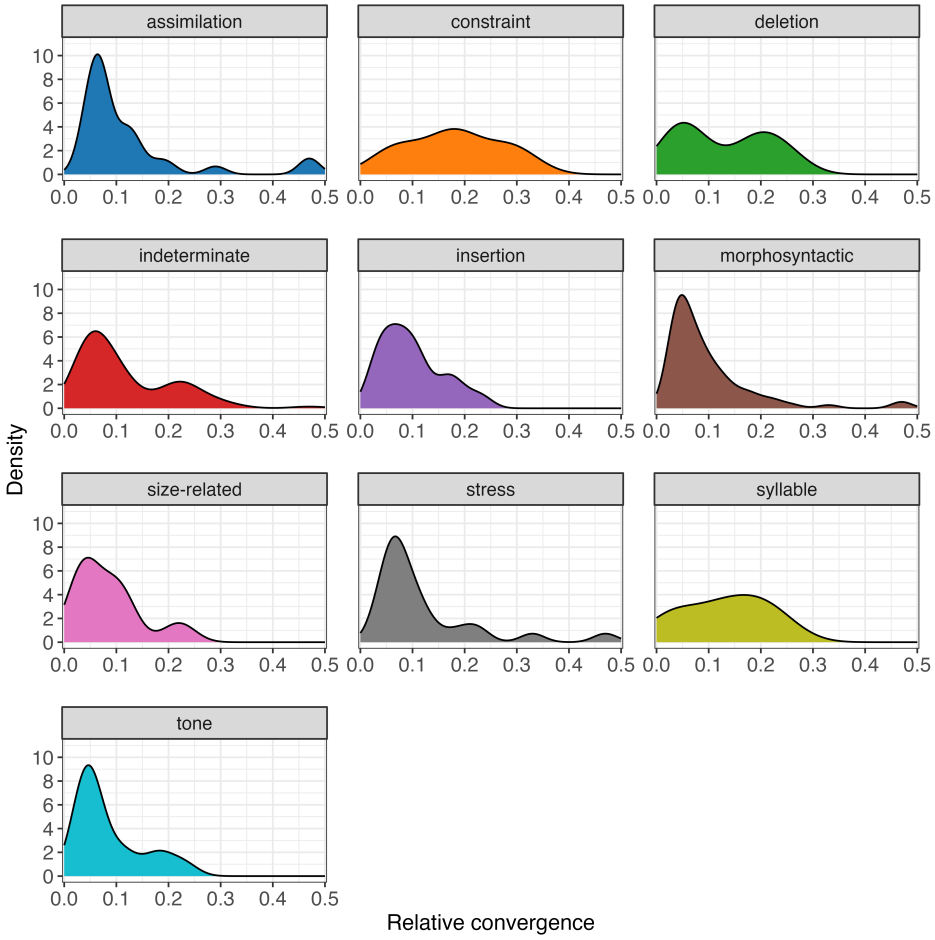


Figure 17: Density distributions of prosodic word domains based on (Bickel et al. 2009) on relative span convergence in the verbal domain. Domains with fewer than 5 data points are excluded.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

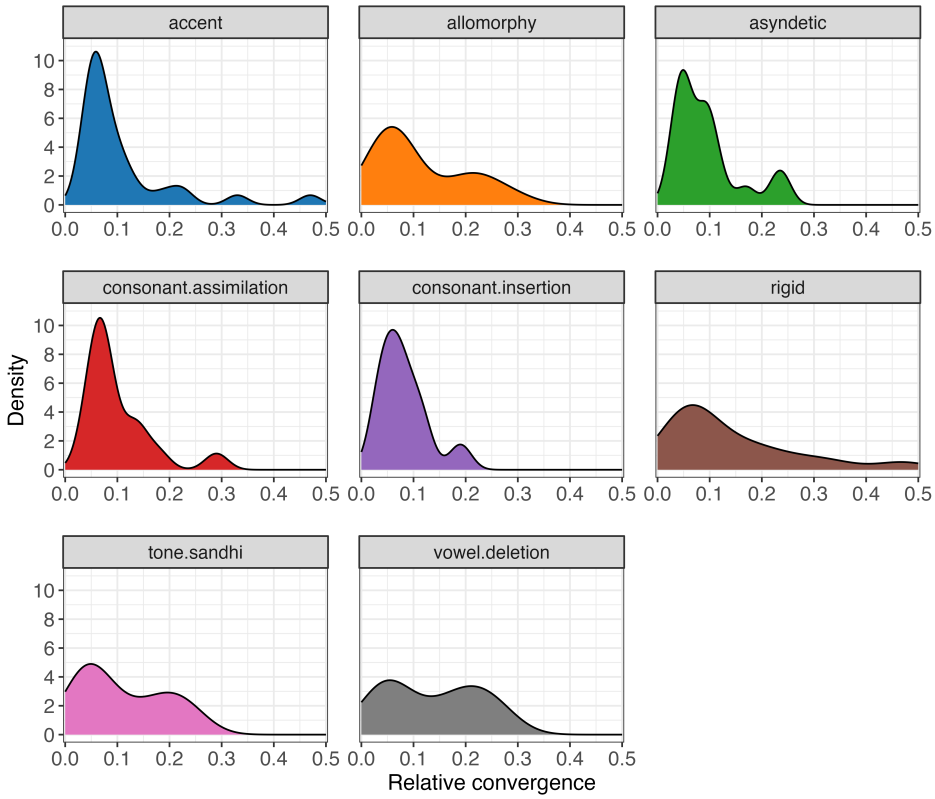


Figure 18: Density distributions of relative convergence across cross-linguistic fractures with 10 or more tokens in the verbal domain

for discussion). We would also say that the p-word in the Central Alaskan Yupik verbal domain is relatively well supported. While the convergence level is only 3, only 4 tests were applicable.

In the morphosyntactic domain (Figure 20), no layer of structure goes beyond 4 layers of structure. Central Alaskan Yupik, zenzeontepec Chatino, and Duraznos Mictex seem to display the strongest candidates for morphosyntactic wordhood. Note that the latter is somewhat weaker because in Duraznos a larger repertoire of morphosyntactic tests could be applied. Slightly weaker domains appear for Cherokee, Blackfoot, Mocovi and Mẽbêngôkre verbal domains.

There are only two languages that provide some type of support for the word-bisection thesis: Central Alaskan Yupik and Zenzeontepec Chatino. In both cases there are domains with relatively high convergences in both moprhosyntax and phonology. While some degree of convergence appears to be the norm, the more

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

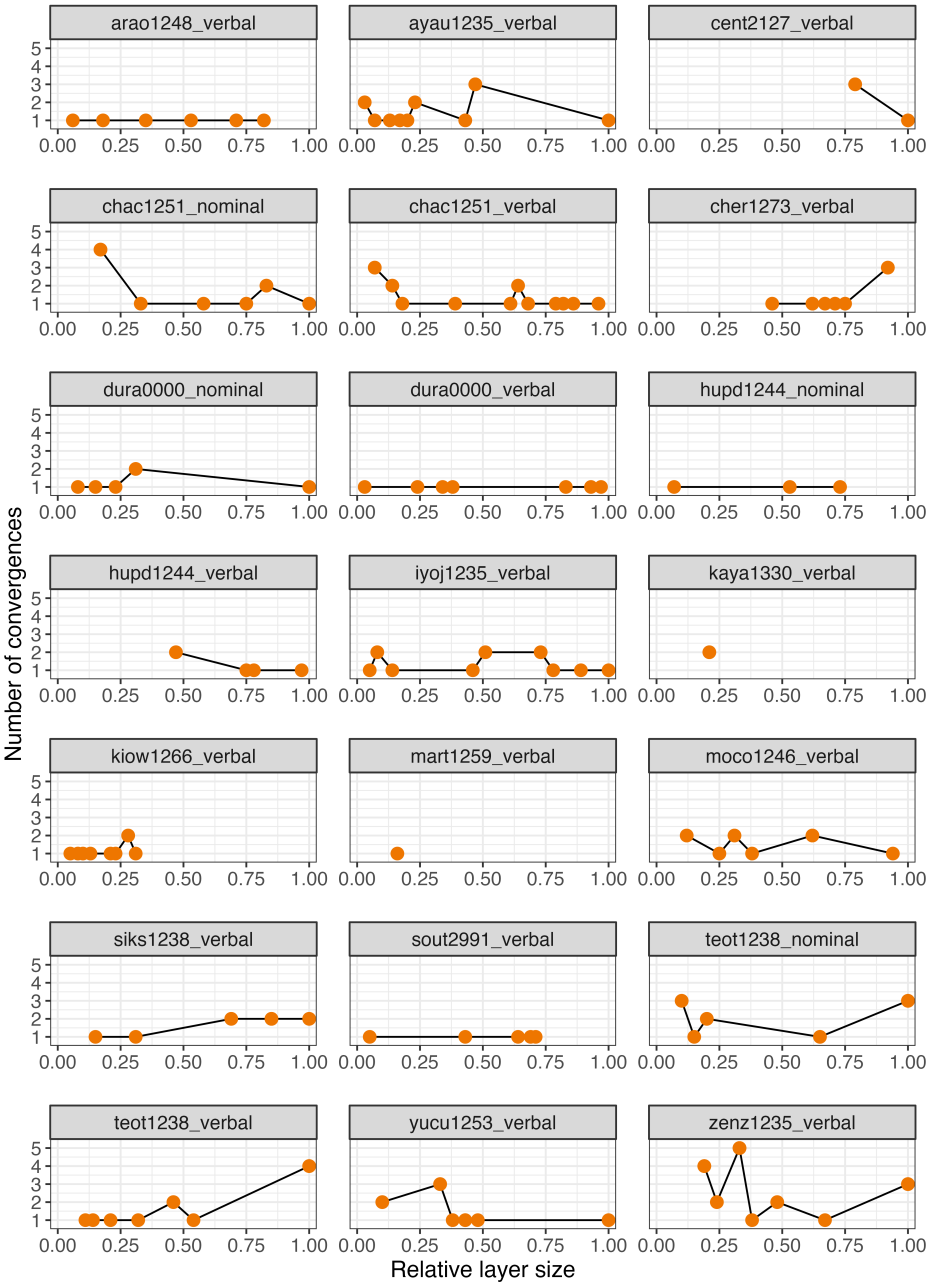


Figure 19: Distribution of relative convergence versus layer size in phonological domains by planar structure

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

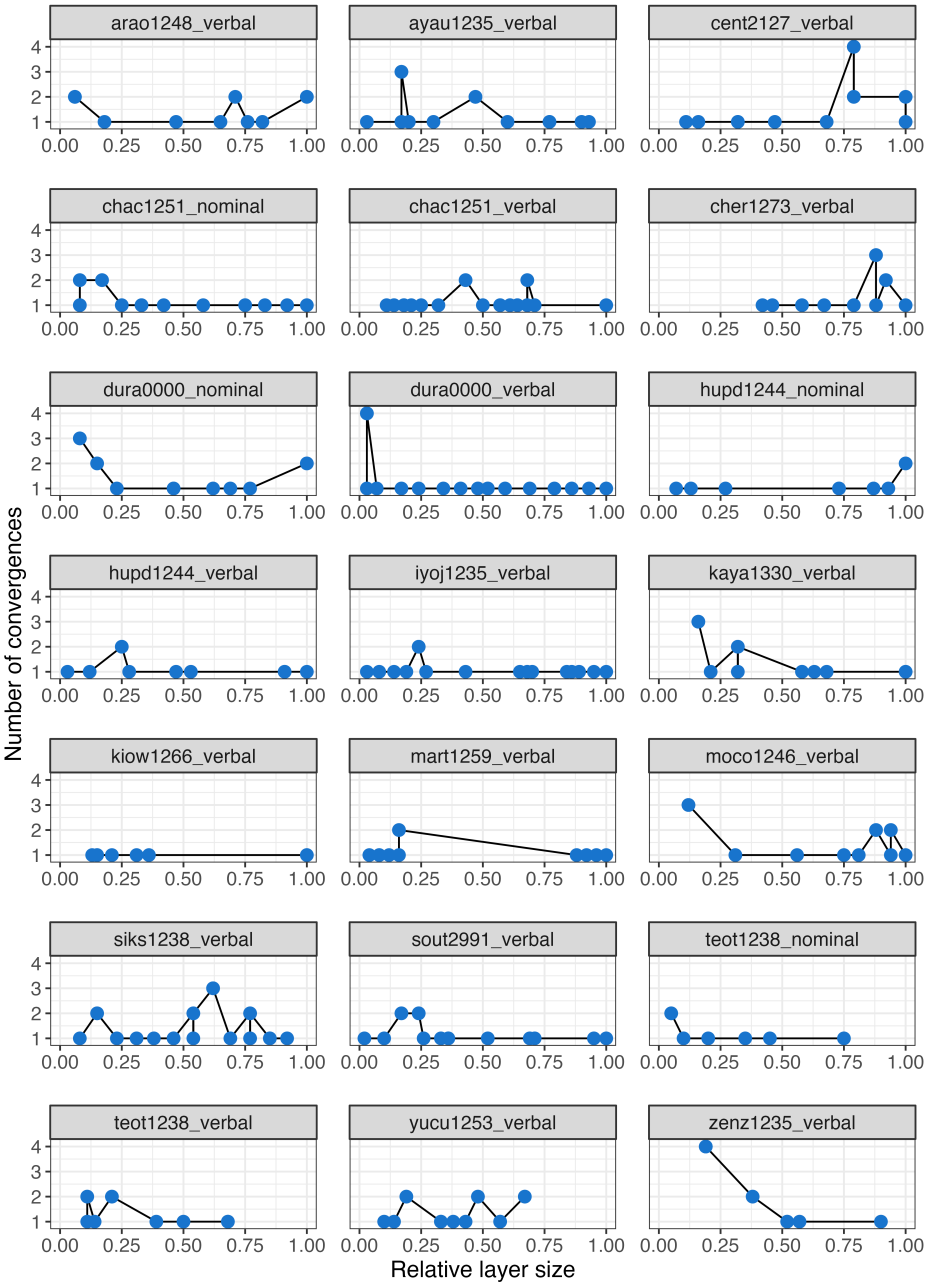


Figure 20: Distribution of relative convergence versus layer size in morphosyntactic domains by planar structure

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

typical pattern thus far is that either there is a highly convergent phonological domain or a highly convergent morphosyntactic one, but not both.

We emphasize again that the meaningfulness of the (non)convergences across languages is an open question both on methodological and theoretical grounds. On methodological grounds, more realistic simulation methods might find that the apparently highly convergent patterns are not surprising given factors such as the number of tests applied, the number of languages considered, the tendency for tests to nest, and the hypothesis space for test alignment (e.g. the planar structure). On theoretical grounds, researchers could challenge the idea that convergence is the right notion to use to assess the word bisection thesis, for instance. We might also find independent reasons to consider some tests are more theoretically relevant than others. There are other tests that have not yet been included in the database (e.g. proform replacement), but whose inclusion might change the picture as well.

8 Summary and conclusion

The goal of this paper is to summarize the construction, conceptualization and structure of the constituency and convergence data set. We also, in a general sense, show how the data set can be used to investigate typological questions in linguistics.

Apart from developing simulation methods as described in the previous section, future research can be concerned with developing more constituency tests, attempting to tease out the distinction between wordhood and phrasehood tests (or levels in general) more explicitly. A fuller account of convergences in nominal domains also needs to be provided. In this book we focused mostly on the verb, because we viewed this as more consistently associated with problems of wordhood, probably because of its relatively high syntagmatic complexity compared to the noun. If both verbal and nominal domains are considered, an actual assessment of the degree to which verbal and nominal constituency structure is actually homologous could be given (e.g. some version of X' theory could actually be tested empirically).

A number of phonological domains are also likely missing across the languages. For instance, there is a relative absence of claims or information concerning utterance level phenomena in this study. This is a natural consequence of the project starting with a focus on 'wordhood', but now that it has been revealed that a focus uniquely on wordhood is likely incoherent, higher level prosodic domains ought to be included.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Deviations from biuniqueness are also relatively superficially considered in the current approach as well. In the current approach domains wherein their are elements that display some definable notion of deviation from biuniqueness are identified, fractured according to the type of deviation (e.g. extended exponence, suppletion etc.). A great deal of complexity and variation is hidden behind such designations. Future research might be concerned with findings some way of syncing current studies on paradigmatic complexity and morphomic structure (e.g. Borja 2023) with a broader study of constituency.

References

Ackermann, Robert John. 1985. *Data, instruments, and theory: a dialectical approach to understanding science*. New Jersey: Princeton University Press.

Adger, David. 2003. *Core syntax: a minimalist approach*. Cambridge: Oxford University Press.

Aikhenvald, Alexandra Y., R.M.W. Dixon & Nathan M. White. 2020. The essence of 'word'. In *Phonological word and grammatical word*. Oxford: Oxford University Press.

Arcodia, Gieorgio Francesco & Bianca Basciano. 2020. Morphology in sino-tibetan languages. In *Oxford research encyclopedias, linguistics*. Oxford. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.530.

Auderset, Sandra & Adam James Ross Tallman. 2023. Constituency and convergence/constituency database: 1.0.0 [data set]. *Zenodo*. DOI: https://doi.org/10.5281/zenodo.10076550. https://doi.org/10.5281/zenodo.10076550.

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

Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge.

Bazell, C.E. 1953. *Linguistic form*. Instanbul: Isanbul Press.

Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *Oxford handbook of linguistic analysis*, 2nd edn., chap. 37, 901–924. Oxford University Press.

Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Prasad Paudyal, Ichchha Purna Rai, Manoj Rai, Novel Kishore Rai & Sabine Stoll. 2007. Free prefix ordering in Chintang. *Language* 83(1). 43–73.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Bickel, Balthasar, Kristine A. Hildebrandt. & René Schiering. 2009. The distribution of phonological word domains: a probabilistic typology. In Janet Grijzenhout & Kabak Baris (eds.), *Phonological domains: universals and deviations*. De Gruyter Mouton.

Bickel, Balthasar & Johanna Nichols. 2002. Autotypologizing databases and their use in fieldwork. In *Proceedings of the international LREC workshop on resources and tools in field linguistics, Las Palmas*, vol. 2627.

Bickel, Balthasar, Johanna Nichols, Taras Zakharko, Alena Witzlack-Makarevich, Kristine Hildebrandt, Michael Riessler, Lennart Bierkandt, Fernando Zúñiga & John B. Lowe. 2017. *The AUTOTYP typological databases. version 0.1.0*. https://github.com/autotyp/autotyp-data/tree/0.1.0.

Boas, Franz. 1911. Introduction. In *Handbook of american indian languages, bulletin 40, part 1*, 1–83. Washington D.C.: Bureau of American Ethnology.

Booij, Geert. 2005. *The grammar of words*. Oxford & New York: Oxford University Press.

Borja, Herce. 2023. *The typological diversity of morphomes: a cross-linguisti study of unnatural morphology*. Oxford: Oxford.

Brown, E.K. & J.E. Miller. 1980. *Syntax: a linguistic introduction to sentence structure*. London: Hutchinson University Library.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, Joan, Parmotima Chakraborti, Dagmar Jung & Joanne Scheibman. 1998. Prosody and segmental effect: some paths of the evolution of word stress. *Studies in Language* 2(22). 267–314.

Carnie, Andrew. 2010. *Constituent structure*. New York: Oxford University Press.

Carruthers, Peter. 2006. *The architecture of mind: massive modularity and the flexibility of thought*. Oxford: Clarendon Press. DOI: https://doi.org/10.1093/acprof:oso/9780199207077.001.0001.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

Coltheart, Max. 1999. Modularity and cognition. *Trends in Cognitive Science* 3(3). DOI: 10.1016/s1364-6613(99)01289-9.

Croft, William. 2001. *Radical Construction Grammar*. Oxford, U.K.: Oxford University Press.

Croft, William. 2010. Ten unwarranted assumptions in syntactic argumentation. In Kasper Boye & Elisabeth Engberg (eds.), *Language usage and language structure*, 313–350. Berlin, New York: De Gruyter Mouton.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Croft, William. 2022. *Morphosyntax: constructions of the world's languages*. Cambridge: Cambridge.

Dixon, R. M. W. 2010a. *Basic linguistic theory: Grammatical topics*, vol. 2. Oxford: Oxford University Press.

Dixon, R. M. W. & Alexandra Y Aikhenvald. 2002a. Word: A typological framework. In R. M. W. Dixon & Alexandra Y Aikhenvald (eds.), *Word: A cross-linguistic typology*, 1–41. Cambridge: Cambridge University Press.

Dixon, R. M. W. & Alexandra Y Aikhenvald. 2002b. Word: A typological framework. In R. M. W. Dixon & Alexandra Y Aikhenvald (eds.), *Word: A cross-linguistic typology*, 1–41. Cambridge: Cambridge University Press.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002c. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: Across-linguistic typology*, 1–34. Cambridge: Cambridge University Press.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2002d. Word: A typological framework. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: Across-linguistic typology*, 1–34. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2010b. *Basic linguistic theory: volume 2 grammmatical topics*. Oxford University Press: Oxford.

Dryer, Matthew S. & Martin Haspelmath. 2013. *WALS online*. Matthew S. Dryer & Martin Haspelmath (eds.). http://wals.info/. Leipzig.

Easterday, Shelece, Matthew Stave, Marc Allassonnière-Tang & Frank Seifart. 2021. Syllable complexity and morphological synthesis: a well-motivated positive complexity corerlation across subdomains. *Frontiers in Psychology* 12. DOI: https://doi.org/10.3389/fpsyg.2021.638659.

Emkow, Carola. 2019. *A Grammar of Araona* (Outstanding grammars from Australia 19). Munich: LINCOM.

Good, Jeff. 2016. *The linguistic typology of templates*. Cambridge: Cambridge University Press.

Greenberg, Joseph. 1954. A quantitative approach to the morphological typology of language. In Robert F. Spencer (ed.), *Method and perspective in anthropology*, 192–220. University of Minnesota Press.

Hacking, Ian. 1983. *Representing and intervening: introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45. 31–80. DOI: 10.1515/flin.2011.002.

Haspelmath, Martin. 2020. The morph as a minimal linguistic form. *Morphology* (30). 117–134.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Lara, Luis Fernando. 2004. ¿es posible una teoría de la palabra? *Lexis* 1-2(XXVII). 401–427.

Ledgeway, Adam. 2017. Syntheticity and analyticity. In Andres Dufter & Elisabeth Stark (eds.), *Manual of romance morphosyntax and syntax (mrl 17)*, Günter Holtus and Fernando Sánchez Miret, chap. 23, 839–886. Berlin: Walter de Gruyter.

Levine, Robert D. 2017. *Syntactic analysis: an hpsg-based approach*. Cambridge: Cambridge University Press.

Lorenz, David & David Tizón-Couto. 2019. Chunking or predicting - frequency information and reduction in the perception of multi-word sequences. *Cognitive Linguisics* 4(30). 751–784.

Martinet, André. 1962. *A functional view of language*. Oxford: Clarendon Press.

Matthews, Stephen. 2002. What can we conclude? In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A Cross-Linguistic Typology*, 266–281. Cambridge, U.K.: Cambridge University Press.

Mayo, Deborah. 2018. *Statistical inference as severe testing: how to get beyond the statistics wars*. Cambridge University Press: Cambridge.

Osborne, Timothy. 2018. Tests for constituents: what they really reveal about the nature of syntactic structure. *Language Under Discussion* 1(5). 1–41.

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.

Payne, Thomas. 2006. *Exploring language structure: a student's guide*. New York: Cambridge University Press.

Phillips, Colin. 1996. *Order and structure*. Cambridge Massachusetts: MIT. (Doctoral dissertation).

Pitman, Donald. 1980. Bosquejo de la gramática araona. In *Notas lingüísticas de Bolivia: No. 9*, 126. Cochabamba: Summer Institute of Linguistics.

Rasskin-Gutman, Diego. 2005. Modularity: jumping forms within morphospace. In Wener Callebaut, Diego Rasskin-Gutman & Herbet A. Simon (eds.), *Modularity. universtanding natural complex systems*, 207–219. MIT Press.

Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace & World.

Schiering, René, Balthasar Bickel & Kristine Hildebrandt. 2012. Stress-time = word-based? Testing a hypothesis in prosodic typology. *STUF-Language Typology and Universals* 65(2). 157–168.

Schwegler, Armin. 1990. *Analyticity and syntheticity: a diachronic perspective with special reference to romance languages*. Berlin: Mouton de Gruyter.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Skirgård, Hedvig, Hannah J Haynie, Damián E Blasi, Harald Hammarström, Jeremy Collins, Jay J Latarche, Jakob Lesage, Tobias Weber, Alena Witzlack-Makarevich, Sam Passmore, et al. 2023. Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. *Science Advances* 9(16). eadg6175.

Tallman, Adam J. R. 2021a. Constituency and coincidence in Chácobo (pano). *Studies in Language* 45(2). 321–383.

Tallman, Adam J. R. & Patience Epps. 2020. Morphological complexity, autonomy, and areality in western amazonia. In Francesco Gardani & Peter Arkadiev (eds.), *The complexities of morphology*. Oxford: Oxford University Press.

Tallman, Adam J.R. 2021b. Analysis and falsifiability in practice. *Theoretical Linguistics* 47(1-2). 95–112.

Tallman, Adam J.R. 2021c. Constituency and coincidence in Chácobo (Pano). en. *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal. http://www.jbe-platform.com/content/journals/10.1075/sl.19025.tal (20 April, 2023).

Witzlack-Makarevich, Alena, Johanna Nichols, Kristine Hildebrandt, Taras Zakharko & Balthasar Bickel. 2022. Managing AUTOTYP data: design principles and implementation. In Eve Koller Lauren B. Collister Andrea L. Berez-Kroeker Bradley McDonnell (ed.), *The open handbook of linguistic data management*. Cambridge: MIT Press.

Zerilli, John. 2020. *The adaptable mind: what neuroplasticity and neural reuse tell us about language and cognition*. Oxford University Press: Oxford.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Chapter 3

Commentary

Taylor L. Miller
State University of New York at Oswego

The Planar-Fractal Method is meant to provide a theory-neutral way to evaluate linguistic theories. In this commentary, I do this for the Combined Model, a new phonology-syntax interface theory which combines Tri-P Mapping and CoPhonologies by Phase. The model successfully predicts and accounts for the patterns in Araona and Ayautla Mazatec, highlighting several strengths of the Planar-Fractal Method and opening issues for future direction.

1 Introduction

As mentioned in the Introduction, the Planar-Fractal Method is meant to provide a theory-neutral way to compare constituency across languages and may be used to evaluate linguistic theories. A model of the phonology-syntax interface, for example, should successfully predict prosodic constituents that align with and explain the phonological patterns and convergences in a given language. In Chapter 4, I identified five wordhood candidates in Kiowa (Tanoan) and found the results neatly coincide with constituents predicted by a new phase-based model (Miller & Sande 2021, 2023) which combines Tri-P Mapping (Miller 2018, 2020) and CoPhonologies by Phase (Sande 2019, 2020; Sande & Jenks 2018; Sande et al. 2020). Here, I test and confirm the Combined Model's success for two other languages from this volume: Araona (Takanan) and Ayautla Mazatec (Popolocan). The results highlight several strengths of the Planar-Fractal Method and open issues for future directions.

In Section 2, I introduce the details of the Combined Model. In Section 3, I test Tri-P's predictions against the wordhood candidates identified in each language. In each language, there is evidence for a Phonological Word (ω), Phonological

Taylor L. Miller. 2024. Commentary. In Adam J.R. Tallman, Sandra Auderset & Hiroto Uchihara (eds.), *Constituency and convergence in the Americas*, 129–141. Berlin: Language Science Press. DOI: ??

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Phrase (φ, and Intonational Phrase (ι). There is also evidence of the Prosodic Stem (ρ) in both languages and Constituent χ in Ayautla Mazatec, though neither constituent has been precisely defined yet under the Combined Model. In Section 4, I discuss the results and conclude.

2 The Combined Model

2.1 Tri-P Mapping

Tri-P Mapping (or Phase-based Prosodic Phonology)[1] is a model of the phonology-syntax interface, which builds on Miller 2018's findings that current interface models (i.e., Relational Mapping as in Nespor & Vogel 1986; Vogel 2019, Syntax-Driven Mapping as in Selkirk 2011, and Syntactic Spell-Out Approaches as in Sato 2006, Pak 2008; Samuels 2011) fall short when tested against data from languages with extreme morpho-syntactic complexity. Relational Mapping and (Direct Reference) Syntactic-Spell Out Approaches alone correctly predicted verb-internal domains in languages like Kiowa and Saulteaux Ojibwe, but neither provided full accounts for either language. Arguing a combined approach with assumptions from both models is necessary, and Miller 2018, 2020 advanced such a model in Tri-P Mapping.

Tri-P Mapping uses an Indirect Reference strategy for mapping prosodic constituents from morpheme- and clause-level phases (Miller 2018, 2020). Phonology may reference any spelled-out phase to map to prosodic structure, but phonology itself does not apply cyclically. This allows for domains of smaller sizes, as opposed to work like Cheng & Downing 2016 which assumes phonology applies after all Spell-Out operations. As in other Indirect Reference Spell-Out accounts (Ahn 2015; Cheng & Downing 2007; Compton & Pittman 2007; Dobashi 2003, 2004a, 2004b; Ishihara 2007; Kratzer & Selkirk 2007; Piggott & Newell 2006), morpheme-level phases (those headed by a categorizing head) map to ω and clause-level phases little v/VOICE map to φ. C's phase maps to ι. Phonologically motivated restructuring may then occur including or excluding various elements within the tree.

Recursion is banned below φ, as in Vogel's Composite Prosodic Model (Vogel 2019). This suggests at least one intermediate constituent between ω and φ is necessary: Constituent χ. This constituent is not yet formally defined, but it is expected to be mapped referencing prosodic and not syntactic structure.

[1]The three Ps of Phase-based, Prosodic, and Phonology are abbreviated as Tri-P.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

2.2 Cophonologies by Phase

Cophonologies by Phase (CbP) is a model of the interface between morphosyntax and phonology, which assumes late insertion of vocabulary items, spell-out at syntactic phase boundaries, and a constraint-based phonology (Sande 2019, 2020; Sande & Jenks 2018; Sande et al. 2020). The innovation of CbP is in the content of vocabulary items, or lexical items. Specifically, in addition to their phonological feature content ($\mathcal{F}$), vocabulary items also contain a prosodic subcategorization frame $\mathcal{P}$ (Inkelas 1990, Paster 2006), and a morpheme-specific constraint ranking adjustment $\mathcal{R}$ (1).

(1) Example CbP vocabulary entry

$$[n] \longleftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{in} \\ \mathcal{P}: & [_{\omega}\ \text{X-} \\ \mathcal{R}: & \textsc{NasalPlaceAssimilation} \gg \textsc{Ident-Place} \end{Bmatrix}$$

The segmental and suprasegmental content of the plural marker in (1) is /in-/, the prosodic subcategorization frame says it should be a prefix within a prosodic word, and the constraint adjustment tells the phonological grammar to rank NasalPlaceAssimilation above Ident-Place. In the spell-out domain containing the morpheme in (1), the default ranking of Ident-Place ≫ NasalPlace-Assimilation will be reversed, resulting in assimilation in this domain, even if assimilation is not a general process in the language. That is, similar to traditional Co-Phonology Theory (Orgun n.d., Anttila 2002, Inkelas & Zoll 2007), there are multiple phonological rankings of constraints within the same language, which vary with the specific morphemes present in a spell-out domain. The key difference is that, in CbP, phonological evaluation applies at phase boundaries, rather than on the addition of each morpheme.

The result of adding morpheme-specific constraint ranking adjustments to vocabulary items is a specific mechanism of communication between the morphology and phonology, such that the phonology knows which grammar or cophonology to apply in a given instance of phonological evaluation. Additionally, the fact that CbP assumes spell-out at syntactic phase boundaries means that morpheme-specific effects are predicted to apply within the phase in which they are introduced, but they are not predicted to affect morphemes introduced in higher phase boundaries (2).

(2) Phase containment principle (Sande & Jenks 2018, Sande et al. 2020):
Morphophonological operations conditioned internal to a phase cannot affect the phonology of phases that are not yet spelled out.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

The phase containment principle, which is related to previous predictions of level-ordering theories and cophonologies (cf. Orgun & Inkelas 2002) holds of morpheme-specific constraint rankings, but also of morpheme-specific prosodic subcategorization effects.

Previous work in CbP has shown that this framework can account for morpheme-specific phonological effects that apply in domains smaller than a word (Sande 2019), larger than a word (Sande & Jenks 2018, Sande et al. 2020), competing morpheme-specific specifications within a phase (Sande et al. 2020), category-specific effects (Sande & Jenks 2018; Sande et al. 2020), and morpheme-specific phonology conditioned by two simultaneous morphological triggers within a phase domain (Sande 2020).

3 Analysis

The two languages presented and analyzed below were selected for no other reason than they were first alphabetically from the list of languages discussed in the present volume (Table 1). The languages are unrelated genetically and aerially and thus offer an interesting test for the Combined Model. In the following subsections, I will present analyses for Aranoa (Takanan) as first analyzed by Adam Tallman in Chapter 12 and Ayautla Mazatec (Popolocan) as first analyzed by Shun Nakamoto in Chapter 5. Both languages are argued to present challenge for any prosodic analysis, but the Combined Model provides a principled account for both. I have included my own chapter's results for Kiowa (Tanoan) in the table below, though interested readers are directed to that chapter itself for the relevant analysis and discussion.

3.1 Araona

Tallman identifies six phonological domains that show no convergence at all. He, however, finds some convergence when including constituency tests which are indeterminate as to whether they fall under phonology or morphosyntax like Free Occurrence, Subspan Repetition, and Extended Exponence. In the end, Tallman only finds two domains that show some convergence: Pos. 4–17 'Prefix'–Connector which is the domain for Maximal Pitch Accent and Maximal Free Occurrence domain and Pos. 4–14 'Prefix'–TAM which is the domain for Minimal Subspan Repetition (*tso* 'prior'), Extended Exponence (Negation), and E-selection. Tallman posits that we may need to ignore span convergence and instead examine the strongest structural edges. In Araona, this is the 'Prefix' (Pos. 4) and the Core Verb Root (Pos. 6).

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Table 1: Summary of Phonological Results

Language	Domain	Reanalysis
Araona	Pos. 6 Verb Core	ρ?
	Pos. 4–15 Prefixes–tam	ω
	Pos. 4–17 Prefixes–Linkage	φ
	Pos. 1–17 Full Clause	ι
Ayautla Mazatec	Pos. 19 Stem	ρ?
	Pos. 15–19 Prog–Stem	ω
	Pos. 15–28 Prog–Pronom.	χ
	Pos. 6–28 Ant./Post.–Pronom.	φ
	Pos. 1–31 Full Clause	ι
Kiowa[a]	Pos. 30–34 Stem–hsy	ω
	Pos. 30–37 Stem–sub	χ
	Pos. 26–37 Pronom–sub	φ
	Pos. 2–40 Full Clause	ι

[a]In the original chapter, there are a total of five wordhood candidates identified via convergence. The fifth candidate is not listed here, as it consists of everything but the initial pronominal in the verb complex. This seems to be a reflex of the phonological separation of the pronominal from the rest of the verb complex and is therefore unrelated to the structure itself.

Tallman ultimately argues for a gradient and more fine-grained view of phonological patterns in the language and cross-linguistically. Therefore, we should move past formalist terminology and constituents used in the literature like 'phonological word' or the rest of the Prosodic Hierarchy. While I agree that the results do look unclear at first glance, Tri-P's independent mapping criteria give us a much clearer picture with three predicted constituents confirmed in the analysis: ω, φ, and ι. There is also evidence for a Prosodic Stem (ρ) constituent, which is yet to be formally defined in Tri-P Mapping.

First, consider the ω domain. Tri-P Mapping predicts categorial heads' phases

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

map to their own ω and may adjust phonologically to include or exclude elements that phonologically cohere or not. For verb complexes, this typically means that the verb stem and any suffixes tend to map to a ω. In Araona, there has been an apparent phonological adjustment to also include material preceding the verb stem. Inflectional prefixes, incorporated noun stems, and inflectional TAM suffixes join the verb core in the ω (Pos. 4–15 as seen in 3 below). This is the domain for E-Selection and Minimality, and it is the Maximal Subspan Repetition (Auxiliary). There is convergence with one morphosyntactic constituency diagnostic; the same subspan is the Maximal Ciscategorial Selection domain. None of these are surprising as ω-level processes and properties.

(3) Araona ω Domain[2]

'Prefix'-	N-	Root	-Aspect	-Margins	-TAM
4	5	6	7–9	10–13	14–15

It's interesting that Araona includes the 'prefixes', which are reportedly complex morphological elements in and of themselves. Cross-linguistically, prefixes are often phonologically separate from the rest of the verb complex due to boundaries of the verb stem's ω and any intervening incorporated stems that also form ωs. These boundaries don't appear to be happening in this case. Though incorporated nouns are typically not included in an ω with another root/stem, bare roots coming together into a single ω is not unattested. In Greek, for example, compounds do not form two ωs to make a new, larger constituent (Athanasopoulou & Vogel 2014). The inflectional prefixes and bare noun roots thus seem to be included in the same domain as the verb core. Both modify the verb (part-to-whole) but are not semantically transparent for transitivity or any other syntactic process. Thus, I am comfortable assuming that the incorporated noun is included in the ω via phonological adjustment. The details of that adjustment are left to future research.

The verb core itself is clearly a domain, too (Pos. 6). I posit that it forms a Prosodic Stem (abbreviated here as ρ), but this constituent has not been formally defined within the framework of Tri-P Mapping. Let us adopt an analysis in the spirit of Downing & Kadenge 2015 and 2020. The ρ in Araona is the Minimal Vowel Syncope domain, as well as the Minimal Free Occurrence and Minimal Subspan Repetition (Auxiliary) domain. There is convergence with two syntactic constituency diagnostics: Minimal Noninterruptibility and Minimal Nonpermutability.

[2]This is a simplified template provided for ease of understanding. The abbreviations used combine and adjust Tallman's verbal planar structure and Pitman's analysis of the Araona verb.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

(4) Araona ρ Domain = Core Verb Root (6)

Tri-P Mapping predicts that a φ will minimally consist of the little v/VOICE phase head's spelled-out phase. In Araona, this domain spans from the prefixes through to the auxiliary and connector at the end of the verb complex (Pos. 4–17). As expected, the language's rather free constituent order means the following XP in Position 18 is not included in the φ domain. The φ in Araona is the Maximal Pitch Accent Domain. There is convergence with two other constituency diagnostics: MAXIMAL FREE OCCURRENCE and MAXIMAL NONINTERRUPTIBILITY.

(5) Araona φ Domain

'Prefix'-	N-	Root	-Asp.	-Margins	-TAM	-Endings
4	5	6	7–9	10–13	14–15	16–17

Finally, Tri-P Mapping predicts that the entire clause will map to an ι because it is the C's phase. There is no positive evidence for the full clause (Pos. 1–17) forming a phonological domain, but it is domain for MAXIMAL SUBSPAN REPETITION (-tso-) and MAXIMAL CISCATEGORIAL SELECTION (broad). Though empty categories with no clear explanation are undesirable, I suspect future research will find ι-level phonological patterns. This is likely a result of the types of phonological processes documented and analyzed rather than a sign there is no ι in Araona.

3.2 Ayautla Mazatec

Nakamoto identifies six wordhood candidates via convergence. Candidate 1[3] consists of the verb root itself (Pos. 19). Three diagnostics converge to identify the domain, all of which are phonological (MINIMAL *ε.J and MINIMAL *3.24) or indeterminate (MINIMAL MINIMUM FREE FORM). Candidate 2 is comprised of all prefixes and the verb root (Pos. 15–19), and it is identified by 5 diagnostics: one is phonological (MINIMAL SANDHI-BLOCKING TONE SEQUENCES) and two are indeterminate (REDUPLICATION and MINIMAL DEVIATION FROM BIUNIQUENESS).

Candidate 3 spans from the prefixes through to the comitative suffix (Pos. 15–20). In other words, this domain spans all non-clitic elements in the verb complex. Of the two diagnostics that converge, only one is phonological. This is the domain for MAXIMAL STRESS ASSIGNMENT. Candidate 4 is just one position larger and includes the focus tonal marker (Pos. 15–21). Two phonological diagnostics converge to identify this domain: MAXIMAL SANDHI-BLOCKING TONE SEQUENCES

[3]Nakamoto refers to Candidates 1–6 and Layer 1, 4, 5, 6, 9, and 11, respectively.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

and Maximal *ε.j. Candidate 5 spans from prefixes through all enclitics (Pos. 15–28), and it shows the highest level of convergence with 7 diagnostics; two are phonological (Obligatory Sandhi and Minimal Possible Sandhi) and two are indeterminate (Maximal Deviation from Biuniqueness and Maximal Minimal Free Form). Candidate 6 (Pos. 6–28) consists of the virtually the entire verb complex. The only position excluded is the focus marker in Position 5. This domain is only identified by two morphsyntactic diagnostics, though.

Because most convergences in Ayautla Mazatec are morphosyntactic and not phonological, Nakamoto concludes that prosodic domains must not be universal as in Schiering et al. 2010. He notes that the fine-grained differences between Candidates 1–6 often hinge on the tonal focus markers in Positions 5 and 21. Their tonal nature poses challenges for most phonological diagnostics. It is therefore separate and forms an incrementally larger domain (e.g. Candidate 4 versus Candidate 3) or left out entirely as in Candidate 6. As in the previous section, however, the Combined Approach (Tri-P Mapping and CoPhonologies by Phase) provides a principled account of what we observe in Ayautla Mazatec.

First, Tri-P Mapping predicts that the ω will coincide with the categorial verb head's phase (i.e. stem and cohering suffixes) with optional phonological adjustment. In Ayautla Mazatec, there is clear phonological adjustment as the ω consists of Pos. 19 (the Stem) and its *preceding* phase (i.e. the inflectional prefixes (Pos. 15)) instead of the phase below like expected. Thus, the ω coincides with Nakamoto's Candidate 2, and it is the domain for phonological processes like Minimal Sandhi Blocking Tone Sequences and Minimal Deviation from Biuniqueness.

(6) Ayautla Mazatec ω Domain

Prog.-	Asp./Mode-	Assoc. Motion-	Caus., Incoh.-	root(s)
15	16	17	18	19

Recall that Nakamoto identified Candidate 3, which includes the ω plus an additional position: the comitative suffix *-ko*^{13} in Position 20. There is indeed a clear separation between the Stem (Pos. 19) and the Comitative (Pos. 20) for Minimum Deviation from Biuniqueness, Total Reduplication, and Verbal Paralellism. In all three cases, the comititative is blocked from being involved. Additionally, Candidate 3 is identified as the domain for Maximal Stress Assignment and Non-permutability. The only phonological diagnostic here is stress assignment, but a re-analysis is possible.

Stress is predictably assigned to the verb root, but it will shift to the comitative suffix if it is present. While Nakamoto identifies the root and comitative as the

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

minimal domain for stress assignment, the maximal domain proceeds backward until the next element that may exhibit stress (i.e. independent pronouns in Pos. 14). It is possible to re-analyze stress as a ρ-final process where stress is applied to a verb root. The comitative's special nature can then be captured by a morpheme-specific overwriting stress assignment or a re-bracketing process. There is no need for an additional prosodic domain.

(7) Ayautla Mazatec ρ Domain = Verb Root (Pos. 19)

Next, Nakamoto's Candidate 4 consists of the ω, the comitative, and the focus marker in Pos. 21. This domain is associated with two phonological constituency tests: Maximal Sandhi Blocking Tone Sequences and Maximal *3.(2)4. In both cases, this domain is established by only negative evidence and no other convergence. While there is clear separation of the other enclitics, there is no way to tell whether the comitative and focus are also separated as they will never participate in either process. Thus, Candidate 4 is not actually a viable candidate and will be excluded from the present analysis. Sandhi Blocking Tone Sequences and *3.(2)4 are assumed to be restricted to ρ.

Turning again to Tri- Mapping, χ is not yet formally defined but seems to support spanning processes between the ω and φ (Miller & Sande 2021). In this language, this domain spans from the progressive prefix (Pos. 15) through the pronominal clitics (Pros. 28). Of the reported processes, only Obligatory Sandhi shows a spanning process across this domain. The remaining processes can be reanalyzed as edge-based phenomena that can be accounted for with a boundary-requirement or constraint instead of appealing to prosodic structure (e.g. pausability is likely referencing the right edge of φ).

(8) Ayautla Mazatec χ Domain

prefixes-root(s)	comitative	focus	enclitics
15–19	20	21	22–28

Next, Tri-P Mapping defines the φ as the little v or VOICE phase, which typically maps to the full verb complex. In Ayautla Mazatec, this domain spans from the anterior/posterior prefix (Pos. 6) to the pronominal clitics (Pos. 28). The adjacent positions are an NP's focus marker on the left, and another NP is on the right. Like the comitative suffix, these positions are mentioned as "prosodically variable" because there are few to no phonological contexts to check the domains of relevant phonological phenomena based on the shapes of the relevant morphemes. At this point, only morphosyntactic evidence converges on this domain (non-interruptibility 1< and coordination min.), but that does not rule it out as a

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

phonological domain. Future research may find a phonological phenomenon in the future that takes place at this level of the prosodic structure. In fact, given the reanalysis above, pausability references the right edge of φ.

(9) Ayautla Mazatec φ Domain

proclitics	adv.,pro.	prefixes.-root(s)	comitative	focus	enclitics
6–13	14	15–19	20	21	22–28

Finally, the ι consists of the full clause, and it is the maximal Possible Sandhi domain. Nakamoto initially lists this as only Pos. 2–31 but there is no reason not to include position 1. It is simply never going to take part in the process, as it never includes tone 4. This is not identified by convergence, but phonological evidence may yet be found.

4 Discussion

The Combined Model's success in Araona and Ayautla Mazatec is only possible because of the fine-grained and comprehensive analysis via the Planar-Fractal Method. First, justifying the planar structures and identifying each element as a zone or slot strips away theoretic decisions like morpheme type. Second, constituency diagnostics are defined precisely and may be fractured to formally account for different types of evidence that may identify subspans (e.g. positive vs. negative evidence). Miller 2018 offered a rudimentary attempt to do this by color-coding different types of evidence, but the final results became unwieldy and hard to follow. This, on the other hand, is quite elegant!

The above analysis raises issues related to convergence, however. Though the Combined Model successfully predicts the subspans in Ayautla Mazatec, most of the convergence is syntactic. In most cases, only one phonological diagnostic identifies each constituent. The fact that the Combined Model still successfully predicts the subspans provides support for convergence alone mattering, but I can see arguments against accepting such lean evidence. If two or more diagnostics of a particular type are required, we would also see issues of insufficient phonological diagnostics in order to satisfy the convergence requirement. Next, a subspan was identified in Ayautla Mazatec by two maximal fractures of tests. In other words, the subspan was identified entirely be negative evidence. This can be handled with a simple constraint that a subpsan cannot be exclusively identified by maximal fractures of diagnostics.

In all, the Planar-Fractal Method successfully enables cross-linguistic comparison and is suitable for testing models of the phonology-syntax interface. Future

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

research should focus on what exactly is expected for convergence across languages.

References

Ahn, Byron. 2015. *Giving reflexivity a voice: Twin reflexives in English*. UCLA. (Doctoral dissertation).

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory* 20(1). 1–42.

Athanasopoulou, Angeliki & Irene Vogel. 2014. Compounds in modern Greek. In *Proceedings of meetings on acoustics 168*, vol. 22.

Cheng, Lisa & Laura J Downing. 2007. The prosody and syntax of Zulu relative clauses. *SOAS Papers in Linguistics* 15. 51–63.

Cheng, Lisa Lai-Shen & Laura J Downing. 2016. Phasal syntax= cyclic phonology? *Syntax* 19(2). 156–191.

Compton, Richard & Christine Pittman. 2007. *Affixation by phase: Inuktitut word-formation*. The 81st Annual Meeting of the Linguistics Society of America. Anaheim.

Dobashi, Yoshihito. 2003. *Phonological phrasing and syntactic derivation*. Cornell University. (Doctoral dissertation).

Dobashi, Yoshihito. 2004a. Multiple spell-out, label-free syntax, and PF-interface. *Explorations in English Linguistics* 19. 1–47.

Dobashi, Yoshihito. 2004b. *Restructuring of phonological phrases in Japanese and syntactic derivation*. Manuscript.

Downing, Laura J & Maxwell Kadenge. 2015. Prosodic stems in zezuru Shona. *Southern African Linguistics and Applied Language Studies* 33(3). 291–305.

Downing, Laura J & Maxwell Kadenge. 2020. Re-placing PStem in the prosodic hierarchy. *The Linguistic Review* 37(3). 433–461.

Inkelas, Sharon. 1990. *Prosodic constituency in the lexicon*. New York: Garland Outstanding Dissertations in Linguistics.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133–171.

Ishihara, Shinichiro. 2007. Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24(2-3). 137–167.

Kratzer, Angelika & Elisabeth Selkirk. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24(2-3). 93–135.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Miller, Taylor L. 2018. *The phonology-syntax interface and polysynthesis: a study of Kiowa and Saulteaux Ojibwe*. University of Delaware. (Doctoral dissertation).
Miller, Taylor L. 2020. Navigating the phonology-syntax interface and Tri-p mapping. In *Proceedings of the annual meetings on phonology*, vol. 8.
Miller, Taylor L & Hannah Sande. 2021. Is word-level recursion actually recursion? *Languages* 6(2). 100.
Miller, Taylor L. & Hannah Sande. 2023. Recursion in Morphology. In *The Wiley Blackwell Companion to Morphology*, 1–43. John Wiley & Sons, Ltd.
Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Dordecht: Foris.
Orgun, Cemil Orhan. N.d. University of California, Berkeley. (Doctoral dissertation).
Orgun, Cemil Orhan & Sharon Inkelas. 2002. Reconsidering bracket erasure. In Geert Booij & Jaap Marle (eds.), *Yearbook of morphology 2001*, 115–146. Dordrecht: Springer.
Pak, Marjorie. 2008. *The postsyntactic derivation and its phonological reflexes*. University of Pennsylvania. (Doctoral dissertation).
Paster, Mary. 2006. *Phonological conditions on affixation*. UC Berkeley. (Doctoral dissertation).
Piggott, Glyne & Heather Newell. 2006. Syllabification and the spell-out of phases in Ojibwa words. *McGill Working Papers in Linguistics* 20(2). 39.
Samuels, Bridget D. 2011. *Phonological architecture: A biolinguistic perspective*. Oxford: Oxford University Press.
Sande, Hannah. 2017. *Distributing morphologically conditioned phonology: Three case studies from Guébie*. University of California, Berkeley. (Doctoral dissertation).
Sande, Hannah. 2019. A unified account of conditioned phonological alternations: Evidence from Guébie. *Language* 95(3). 456–497.
Sande, Hannah. 2020. Morphologically conditioned phonology with two triggers. *Phonology* 37.3. 453–493.
Sande, Hannah & Peter Jenks. 2018. Cophonologies by phase. In *NELS 48 Proceedings*, vol. 3, 39–53.
Sande, Hannah, Peter Jenks & Sharon Inkelas. 2020. Cophonologies by ph(r)ase. *Natural Language and Linguistic Theory* 38. 1211–1261.
Sato, Yosuke. 2006. Spelling-out prosodic domains: A multiple spell-out account. In *InterPhases Conference*.
Schiering, René, Balthasar Bickel & Kristine A Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46(3). 657–709.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 435–484. Hoboken: John Wiley & Sons, Inc.

Vogel, Irene. 2019. Life after the Strict Layer Hypothesis. In Hongmin Zhang & Youyong Qian (eds.), *Prosodic studies: Challenges and prospects*, 9–60. London: Routledge.

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Name index

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Subject index

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Proofreading version. Do not quote. Final version available from http://www.langsci-press.org

Constituency and convergence in the Americas

add blurb here