The distribution of phonological word domains: a probabilistic typology

Balthasar Bickel*, Kristine A. Hildebrandt#, and René Schiering+

*University of Leipzig, #University of Manchester, +University of Münster

1. Introduction

The past few decades have seen various attempts to develop descriptive frameworks that capture the range of phonological domains expected across languages. Most prominent among these is still the Prosodic Hierarchy framework (Nespor and Vogel 1986; Peperkamp 1997; Selkirk 1984). The framework posits a finite range of possible domains and predicts that phonological patterns universally converge on exactly this range of domains, e.g. there will be exactly one domain between foot and phrase: the phonological word ('ω'). We call this the Clustering Hypothesis. Another prediction of the framework is known as Strict Succession or Proper Headedness and states that each level *L* of the Prosodic Hierarchy dominates at least one level *L-1* until the terminal level *L=0*, e.g. there will always be at least one phonological word level ω dominated by a phonological phrase level. Proper Bracketing, finally, formulates the expectation that no language will exhibit non-stacking domains in its prosodic structure, e.g. the edge of a phonological phrase cannot be situated inside a phonological word.

As Inkelas and Zec (1995: 548) pointed out, although the theoretical predictions of the Prosodic Hierarchy framework are straightforward, the empirical evidence is often less so. When analyzing languages, one often encounters challenges to the predictions made by the Prosodic Hierarchy model. In principle, there are two routes for facing these challenges. One route is to adjust and refine the model, e.g. by extending the range of domains posited (e.g. admitting one or more additional levels such as the Clitic Group: Vogel, this volume), by weakening some claims (e.g. allowing violations of the Strict Succession requirement under certain conditions), or by restricting the scope of the framework (e.g. by limiting it to suprasegmental patterns only).

An alternative route is typological: instead of assuming a finite and universally fixed set of phonological domains, one may ask what kinds of domains are empirically evidenced by a sample of languages and explore the principles that explain the distribution of the attested domains. This chapter takes this second route and reports on the kinds of domains found at the word level – i.e. domains larger than the foot but smaller than the phrase – in a sample of languages. Rather than making an *a priori* assumption that there must be exactly one level of word, we instead define a typological variable whose values are the domains identified between foot and phrase.

We begin by briefly reviewing the kinds of empirical challenges that individual data present (Section 2), and from this we develop a universally applicable method of measuring the diversity observed (Section 3). In order to analyze our measurements, we adopt standard statistical tools, specifically Multi-Dimensional Scaling (MDS) (Section 4). The key result of this analysis is that stress-related domains tend to be larger than domains referenced by other phonological patterns (including any kind of phonology, from vowel harmony to consonant dissimilation). In Section 5 we submit this result to regression modeling and show that the impact of stress is independent of areal and genealogical affiliation, and is therefore a good candidate for a truly universal principle. Section 6 summarizes the findings and discusses its consequences.

2. Cross-linguistic challenges

Consider the prosodic domains in Limbu, an Eastern Kiranti (Sino-Tibetan) language of Nepal (van Driem 1987; Ebert 1994; Hildebrandt 2007; Schiering, Hildebrandt and Bickel 2007; Tumbahang 2007; Weidert and Subba 1985). At first sight, phonological patterns reference at least four distinct domains between the foot (φ) and the phonological phrase (P) (Figure 1).

INSERT FIGURE 1 ABOUT HERE

Figure 1: Prosodic Domains in Limbu (Phedappe dialect; Hildebrandt 2007)

The data in Figure 1 appear to challenge all predictions of the Prosodic Hierarchy framework: (i) prosodic domains cluster on more domains than provided by the Prosodic Hierarchy, i.e. at the word level we have to distinguish four distinct layers. With respect to Strict Succession/Proper Headedness, Limbu shows that a level ω may dominate another level ω, i.e. it does not directly dominate the next lower level, a foot. In this case, the domains ω_1 and ω_2 constitute non-stacking domains, i.e. overlapping domains, and thus violate Proper Bracketing.

However, under closer inspection, not all four domains pose an equal challenge. First, the Prosodic Hierarchy model only makes claims about purely phonological rules that apply generally across the lexicon, and makes no predictions about lexically limited phonological patterns. This discards the evidence for ω_1 and ω_2 because the processes which reference the morphological string in question are observable in two individual affixes, and we cannot exclude the possibility that these are lexically specified. But word domains ω_3 and ω_4 remain an issue because they are referenced by general rules.

Since the source for the violations of the Prosodic Hierarchy predictions is prefixes, one possible solution would be to analyze the prefixes as postlexical. However, since prefixes are obligatory elements of inflectional morphology and can only ever appear as parts of verb forms, there is no independent evidence that they could be analyzed as postlexical. More importantly, prefixes (and, for that matter, suffixes) contrast with genuine postlexical elements ('clitics') in Limbu, which syntactically attach to phrases and do not subcategorize for a specific part of speech or stem type. Clitics of this kind, e.g. the additive focus clitic *=aŋ* in Figure 1, are always included in both ω_3 and ω_4. Since prefixes are included in ω_4 but not in ω_3, analyzing them as postlexical would entail that two distinct postlexical domains need to be posited. As a result, one could at best re-label ω_3 and ω_4 as, for example, 'Clitic Group 1' and 'Clitic Group 2', but the challenge for the Prosodic Hierarchy framework remains.

Another analysis that could in principle be applied in order to bring Limbu in line with the Prosodic Hierarchy framework is in terms of recursion (e.g. via a low-ranking of the non-recursivity constraint proposed by Selkirk 1995). The larger domain ω_4 would then be analyzed as a recursive instantiation of the smaller domain ω_3. Following standard definitions of recursion, such an analysis would predict that the two domains in question have identical phonological properties (for examples, cf. Peperkamp 1996; Peperkamp 1997). However, this prediction is not borne out by the Limbu data, since the domains are motivated by a number of phonological patterns which are clearly distinct (Hildebrandt 2007). Primary stress assignment, for instance, operates at ω_4 only, and ω_3 has no stress-related properties at all.

In some cases, relativizing prosodic structure to different phonological tiers allows to account for data which otherwise seem to counter the predictions of the Prosodic Hierarchy framework. In Luganda (Bantu), for example, the domain for the presence of only one high-to-low pitch drop and the domain for a rule of final vowel shortening overlap in such a way that Proper Bracketing is violated (Hyman, Katamba and Walusimbi 1987). If we allow prosodic structure to be construed independently on different phonological tiers, in this case tone and quantity, the theoretical predictions are still borne out in a more fine-grained perspective on the architecture of prosodic structure. For Limbu, however, such an approach cannot solve the problem because the various phonological patterns which motivate each word domain come from such diverse phonological tiers and types as phonotactics, segmental, and suprasegmental phonology. In Figure 1, we illustrated ω_4 by stress assignment patterns. But the same domain is also referenced by a segmental rule of Coronal-to-Labial Assimilation:

(1) a. /mɛ-**n-m**ɛ**t-p**aŋ/ [mɛ**mm**ɛ**pp**aŋ]

nsA-NEG-tell-1s>3.PST

'I did not tell him'

b. /hɛ**n = p**hɛlle/ [hɛ**mb**hɛlle]

what-QUOT

'What?'

The rule states that the coronal phonemes /n/ and /t/ anticipate the place of articulation of the bilabial phonemes /m/ and /p/ within a phonological word. The relevant phonological word domain includes both prefix-stem and stem-suffix boundaries, as illustrated by (1a). Like the domain for Primary Stress Assignment, the domain also includes host-clitic boundaries, as shown in (1b).

In sum, regardless of the approach of adjustment or refinement that one may take, languages like Limbu provide unresolved challenges for the predictions contained in the Prosodic Hierarchy Hypothesis. In the following, we turn to an alternative way of looking at the data.

3. Measuring diversity in word domains

An alternative route to addressing challenges posed by languages like Limbu is typological: instead of reducing the observed diversity to a single universal model, we measure the actual distribution of word domains across languages and look for universal principles explaining the observed distribution. For this, we need a universally applicable working definition of word domains and use this to create a database which, for each language, specifies the precise nature of its phonological domains.

3.1. A working definition of phonological word domains

Our working definition of phonological word domains includes all sound pattern domains that are delimited by some morphological structure but do not include more than one lexical stem. Thus, we concentrate exclusively on specific combinations of stems, affixes and clitics/particles, and we exclude feet or syllable structure (where definitions do not reference morphological structure) as well as any larger domains like phrases and compounds. In line with the Prosodic Hierarchy framework, we also limit all our analyses to lexically general phonological patterns, i.e. we exclude data that are positively limited to a subset of affixes or stems, or only to loanwords.[1] For convenience, we refer to the patterns as defined here as ‘pw-patterns’.

3.2 The database

Our database[2] currently covers 70 typologically diverse languages, but for seven of these languages, we have not found any evidence for pw-patterns that are strictly sub-phrasal and are at the same time fully general across the lexicon. This reduces the dataset to 63 languages containing a total of 382 pw-patterns.

In one component of the database we collect, for each language, all pw-patterns, such as stress, tone, segmental rules and phonotactic constraints, with an exact description of their phonological properties. All 63 languages have between 1 and 19 distinct pw-patterns, and more than half of them have between 1 and 5 pw-patterns. For each pw-pattern we code the kind and number of morphemes which are included in its domain. In order to calibrate this information against the morphological set-up of the language, we also keep track of the different morpheme types that are relevant for a given language, such as ‘suffixes’ (defined as postposed grammatical markers that are subcategorized for a stem class), ‘proclitics’ (defined as preposed grammatical markers that are not restricted to a stem class or a part of speech), ‘stems’ etc. This information is included in a separate component of the database. Most languages have between 2 and 7 morpheme types, and almost 80% of them between 2 and 4 morpheme types.

---

[1] However, we systematically entered such data into our database in order to allow more comprehensive analyses later. Interestingly, the general results presented in this paper do not depend on whether lexically-specified patterns and domains are included or excluded from the analysis. Some statistical signals become stronger, some weaker, but the overall findings stay.

[2] For a project description and ancillary material, including analyses of individual languages, see www.uni-leipzig.de/~autotyp > projects and > language reports. The web site also makes available the database itself and a bibliography of its sources. [NOTE: this material will be released by the time of publication of this paper; advance copies can be requested by e-mail to bickel@uni-leipzig.de].

3.3. A measurement of word coherence

The intuition behind the Prosodic Hierarchy is that, independent of the morphological make-up of a language (e.g. whether a language happens to have prefixes or not), domains will distribute in a hierarchy of similarly or identically-sized levels — starting at the top with large domains that include all morphological material that a language has and ending at the bottom with narrowly defined domains like stem-suffix combinations that exclude prefixes and clitics, for example. Such a view suggests two probabilistic expectations across languages: (i) we expect some kinds of domains to be recurrently larger than others, so that, within languages, the larger domains will properly contain the smaller ones; (ii) these hierarchies of domains will tend to cluster on universal 'attractors' that are defined by some shared property. For example, there could be a cross-linguistic trend towards three nested sizes — say, 'large', 'medium' and 'small' — each characterized by the kind of phonology involved, e.g. stress placement vs. vowel harmony vs. consonant assimilation.

To find out, we need a way of measuring the size that domains have in relation to each other. A straightforward way of doing this is by determining how many of the morpheme types that a language has are included in a given phonological word domain. In principle, this could be the stem alone, the combination of stem, suffixes and clitics, as in the Limbu ω_3 domain, or even the combination of all available morpheme types, e.g., the prefix-stem-suffix-clitic string in the Limbu ω_4 domain. Since the number of morpheme types included in a domain depends on the number of morpheme types available in a given language, we define the relative word size as *c* of a pw-pattern *p* in a language *L* as (where *c* is mnemonic for 'coherence'):

(2) $c(\text{p,L}) = \dfrac{N(\text{morpheme types in domain referenced by } p)}{N(\text{morpheme types in language } L)}$

For example, a domain like ω_4 in Limbu is defined, among other patterns, by the Primary Stress Assignment listed in Figure 1 and the segmental rule of Coronal-to-Labial Assimilation discussed in example (1). The relative coherence *c* of these patterns is measured as follows:

(3) a. /mɛ-'thaŋ-e=aŋ/ 'they come up and …'

3ns-come.up-PST=and

b. /'ku-la:p/ 'its wing'

3POSS-wing

$\dfrac{4 \text{ (prefix-stem-suffix=clitic)}}{4 \text{ (prefix-stem-suffix=clitic)}}$

→ *c* (Limbu Primary Stress Assignment) = 1

(4) a. /mɛ-**n-m**ɛ**t-p**ɛŋ/ [mɛ**mm**ɛ**pp**aŋ] 'I did not tell him'

nsA-NEG-tell-1s>3.PST

b. /hɛ**n** = **p**hɛlle/ [hɛ**mb**hɛlle] 'What?'

what-QUOT

$\dfrac{4 \text{ (prefix-stem-suffix=clitic)}}{4 \text{ (prefix-stem-suffix=clitic)}}$

→ *c* (Limbu Coronal-to-Labial Assimilation) = 1

In (3a), only one primary stress is assigned to the combination of prefix-stem-suffix=clitic. In the given form, this primary stress is realized on the stem, but, as shown in (3b), in other forms the stress shifts to the prefix (cf. Hildebrandt 2007; Schiering, Hildebrandt, and Bickel 2007 for discussion). This suggests that the stress assignment rule of Limbu references all four available morpheme types of the language, and has therefore coherence *c* = 1. The examples in (4) repeat those of (1) and illustrate Coronal-to-Labial Assimilation. Combining the evidence from (4a) and (4b), the assimilation spans all four morpheme types available in the language and is thus coded as having coherence *c* = 1.

Of course, prosodic word domains often do not reference all available morpheme types of a given language, i.e. have *c* < 1. This point can be illustrated with the Limbu word domain ω_3 outlined in Figure 1. This domain is motivated by two phonological processes, i.e. Glottal Stop Insertion and the [l] ~ [r] alternation, and applies to only three of the four morpheme types available in the language.

(5) a. /k**u-eː**k/ [k**uʔeː**k] (3POSS-back) 'its/his/her back'

b. /a-mph**u-eː**/ [amph**ueː**] (1POSS-brother-VOC) 'Brother!'

c. /nu-b**a=iː**/ [nu-b**aiː**] (be.alright-NOM=Q) 'Is this OK?'

d. /**a-iː**r-ɛ/ [ʔ**aʔiː**rE] (1-wander-PST) 'We wandered.'

3 (stem-suffix=enclitic)

4 (prefix-stem-suffix=enclitic)

→ *c* (Limbu Glottal Stop Insertion) = .75

(6) a. /nɛ**l**ɛt/ [nɛ**r**ɛt] ‘heart’

b. /pha-**le** siŋ/ [pha-**re** siŋ] (bamboo-GEN wood) ‘the wood of bamboo’

c. /pe:g-i=**l**o:/ [pe:g-i=**r**o:] (go-p=ASS) ‘Come on, let’s go!’

d. /kɛ-**l**ɔʔ/ [kɛ-**l**ɔʔ] (2-say) ‘you say’

$\frac{\textit{3}\text{ (stem-suffix=enclitic)}}{\textit{4}\text{ (prefix-stem-suffix=enclitic)}}$

→ *c* (Limbu [l] ~ [r] domain) = .75

In (5a), a glottal stop is inserted between the prefix and the vowel-initial stem. In the segmentally identical hiatus situation at the stem-suffix boundary in (5b), the process does not apply to the vowel-initial suffix. The process also fails to apply with vowel-initial enclitics, as is shown in (5c). From this we can conclude that Glottal Stop Insertion targets the left edge of a vowel-initial prosodic word which references the stem, its suffixes and enclitics. Accordingly, three of the four available morpheme types are included in the domain. The resulting coherence value for this domain is therefore *c* = .75. Example (5d) demonstrates that the vowel-initial prefix constitutes its own prosodic word for the sake of this process.

The data in (6) illustrate the domain structure of the [l] ~ [r] alternation. Syllable-initial /l/ always surfaces as [l] in word-initial position, while in word-medial position it is realized as either [r] or [l], depending on the structure of the preceding syllable. The realization is [r] if the preceding syllable is open or ends in a glottal stop. In (6a), /l/ appears within a monomorphemic word and surfaces as [r] following an open syllable. For the sake of this alternation, the stem-suffix boundary in (6b) and the host-enclitic boundary in (6c) also provide an appropriate context for the realization of the allophone [r]. On the basis of these observations, we can formulate the generalization that, parallel to the case of Glottal Stop Insertion, the [l] ~ [r] alternation applies within a prosodic domain which encompasses the stem, its suffixes and enclitics, i.e. three of the four available morpheme types. By dividing 3 by 4 we again get a coherence value of c = .75 for this domain. The data in (6d) show that the prefix lies outside this domain.

Our definition of c abstracts away from the distinction between the kinds of morpheme boundaries involved. If a hypothetical language has prefixes, stems and suffixes as morpheme types, a prefix+stem domain and a stem+suffix domain will each score the same c-value of 2/3 although these involve different types of morpheme boundaries. However, both domains are smaller than an all-encompassing prefix-stem-suffix domain, and for finding out whether there are trends towards hierarchical arrangements of larger vs. smaller domains, the general size difference is more important than the exact alignment of domains.[3] This justifies our abstracting away from morpheme boundary types. Still, we checked whether equating prefix- and suffix-oriented domains might have a distorting effect. It does not: a comparison of the means and variances of c values when calculated only to the left of the stem (stem and prefix/proclitics) vs. when calculated only to the right of the stem (stem and

[3] In fact, the hypothetical case violates the principle of Proper Bracketing. The few languages that have this pattern in our dataset invite research beyond the scope of this chapter. Independently of this, however, we explored an alternative way of comparing domains, based on determining whether or not given morpheme boundary types (e.g. stem-suffix, prefix-stem, etc.) are each included in two domains (a criterion we use in Figure 2 below). This approach, suggested to us by an anonymous reviewer, seeks clusters with regard to specific boundary types and exact domain definitions, and not with regard to general size differences. In contrast to what we report below on c, a multidimensional scaling analysis based on such domain definitions did not reveal any systematic clustering (and Kruskal Stress values below 10% required at least 4 dimensions). It is possible that the exact regulation of domains in terms of boundary types is free of universal preferences, but we must again leave this question to future research. (The coding is available in the published database.)

suffix/enclitic) revealed no statistical differences in our data (Monte-Carlo permutation[4] t-test, $t = -1.79$, $p > .05$; Fligner-Killeen test, $\chi^2 < .01$, $p > .05$).

The measurements of *c* across pw-patterns and across languages define a typological variable with values starting near 0 (where only one out of many morpheme types is included) and an upper limit of 1 (where all available types are included). In our database, *c* ranges from .14 to 1 in eighteen intervals spaced between .02 and .14, which we take as approximating a continuous scale for statistical purposes.

3.4. Data summary and discussion

One key prediction of the Prosodic Hierarchy framework is that in each language, phonological patterns converge on a single domain between foot and phrase. After exhaustively searching the literature and, in many cases, having undertaken additional studies on available phonetic data or having queried fieldworkers, we observe that the majority of the 63 languages under study have more than a single domain.[5] Figure 2 plots how many languages (y-axis) have a given number of non-isomorphic domains (x-axis), where isomorphism is defined as identity in morpheme types and morpheme boundaries.

INSERT FIGURE 2 ABOUT HERE

Figure 2: Number of non-isomorphic domains referenced by lexically general phonological patterns (data from 63 languages). The x-axis represents the

[4] Since our data are not randomly sampled, we assess statistical differences by permutation methods throughout (cf. Everitt and Hothorn 2006; Janssen, Bickel and Zúñiga 2006; R_Development_Core_Team 2007). All approximations are based on 10,000 random permutations.

[5] For one language, Khalkha Mongolian, the search was not exhaustive and there may be more nonisomorphic domains than what we have coded in our database.

number of non-isomorphic domains, the y-axis the number of languages with that many non-isomorphic domains.

Figure 2 shows that only 9 out of the 63 languages considered here show the predicted clustering of word-related phonological patterns on one word domain. All other languages violate Clustering by having more than one non-isomorphic domain. In other words, Limbu is not an idiosyncratic isolated case of deviation but instead exemplifies a cross-linguistic trend for prosodic domains to multiply at the word level.

4. Probabilistic clusters

A question arising from the observation in the preceding section is whether, instead of a categorial cluster of size 1 (i.e. with exactly one domain on which all pw-patterns converge), there are probabilistic clusters of similarly-sized domains. If there are, it is likely that they depend on various factors. In this paper, we explore the possibility that the coherence of pw-patterns depends on the kind of sound pattern involved, e.g. it might be the case that across languages, tonal patterns target domains with different relative coherence than domains referenced by segmental patterns. This is much in line with what Hyman, Katamba, and Walusimbi (1987) suggested for Luganda, but re-interpreted as a probabilistic trend rather than any categorical constraint.

To find out whether there are such trends, we conducted a non-metrical Multi-Dimensional Scaling analysis (MDS, e.g. Cox 1994; Everitt and Hothorn 2006; Venables and Ripley 2002) on our dataset of pw-patterns.

4.1. Coding and methods

In preparation of the MDS analysis we coded each individual pw-pattern into a taxonomy of sound pattern types at a chosen level of resolution. For instance, the segmental pw-patterns can be broken down into three basic types: allophony, phonotactic constraint, and alternation process. Alternation processes can, in turn, be broken down into subtypes like assimilation, deletion, dissimilation, etc. While we tested various levels of resolution and ways of setting up the taxonomy, a relatively shallow taxonomy of 15 types revealed all structure that we found with other taxonomies as well. Figure 3 shows the taxonomy that entered into the analysis.

INSERT FIGURE 3 ABOUT HERE

Figure 3: Pw-pattern taxonomy

The terms of the taxonomy are defined as follows. (Numbers in brackets indicate how often the pattern is encountered in the database, again only counting lexically general pw-patterns):

- *allomorphy* (4): some constraint is resolved by choosing a specific allomorph, e.g. an anti-hiatus constraint within (but not between) words is satisfied by systematically choosing C-initial allomorphs.
- *allophony* (4): the surface form of a phoneme depends on the position of the segment within the word, e.g. the Limbu [l] ~ [r] alternation discussed above.

- *assimilation* (41): the realization of a segment is dependent on the segmental context within the word domain, e.g. the Limbu nasal /n/ assimilates to a following velar segment and surfaces as [ŋ], as discussed above.
- *vowel harmony* (7): all vowels within a domain share one or more feature specifications.
- *(phonotactic) constraint* (123): the prosodic word imposes restriction on the distribution of segments within the domain, e.g. a ban on velar nasals in word-initial position.
- *deletion* (25): a segment is deleted in a position specified by the prosodic word domain, e.g. a vowel is deleted in word-final position.
- *dissimilation and metathesis* (3): the realization of a segment is dependent on the segmental context within the word domain, such that the co-occurrence of similar segments is avoided; this also includes metathesis.
- *insertion* (28): some constraint is resolved by inserting a vowel or consonant, e.g. an anti-hiatus constraint within (but not between) words is satisfied by inserting an epenthetic consonant.
- *other process* (2): segmental processes which cannot be grouped with the other segmental categories in the taxonomy, e.g. re-syllabification across morpheme boundaries within a word domain.
- *quantity* (9): generalizations related to the suprasegmental feature length, e.g. vowel lengthening in word-final position.
- *rhythm* (1): phonological patterns which result in the rhythmic structuring of a given domain, e.g. alternating stresses within the word yielding trochaic or iambic feet under the word node.

- *size-related* (35): constraints which specify the maximal or minimal word size in counts of syllables or moras, e.g. minimal or maximal disyllabicity required for stems, or for stem-affix combinations, but not, e.g. for prefix-stem or host-clitics combinations.
- *strengthening* (25): processes which result in the phonological strengthening of a segment, e.g. fortition.
- *stress* (36): generalizations related to the suprasegmental feature stress, e.g. word-final stress placement.
- *tone* (10): generalizations related to the suprasegmental feature tone, e.g. generalizations such as only one high-low pitch drop per word.
- *weakening* (36): processes which result in the phonological weakening of a segment, e.g. lenition.

The next methodological step is the construction of a distance matrix that measures the Euclidean distance between the relative coherence *c* of each pair of the 382 pw-patterns in our dataset. These are $\frac{382 \cdot (382-1)}{2} = 72{,}771$ pairs, of which a randomly selected handful is shown for illustration in Table 1.

INSERT TABLE 1 ABOUT HERE

Table 1: Illustration of distance matrix. Numbers behind language and pw-pattern type name are unique identifiers pointing to the specific processes involved (and described in detail in the database).

Like in geographical distance charts, identical pw-patterns have a distance of 0; but unlike in a geographical distance chart, non-identical pw-patterns can also happen to have a distance of 0. This is the case whenever two pw-patterns happen to target domains of the same relative coherence degree. In Table 1, this is illustrated twice, once within a language (Kinnauri), where two distinct pw-patterns (one of final devoicing and one phonotactic constraint) converge on the same domain, and once between languages, where these two patterns happen to target a same-sized domain as a certain assimilation pattern in Burmese.

The general problem of large distance matrices like the one obtained from our database is that it is virtually impossible to detect, by mere ‘hand-inspection’, whether certain kinds of elements are closer to each other than others. Therefore, we applied Multi-Dimensional Scaling, which projects the 382 dimensions of the observed distance matrix into lower-dimensional space, keeping the necessary distortions at a minimum. Kruskal’s non-metrical algorithm (as implemented by Venables and Ripley 2002) converged on a two-dimensional solution, with a relative low stress value of 4.6%.

4.3. Results

Figure 4 plots the results of the two-dimensional solution. Because clusters are very dense where they occur, we add pie chart insets displaying the internal composition of two selected clusters.

There is an impression of clusters at regular intervals from left to right, but these are artifacts of the possible distances between coherence degrees which are not fully continuous. Closer inspection by zooming into any of these smaller clusters does not reveal any trends. For example, the third cluster from the left in Figure 4, highlighted by the pie chart inset on the left shows a 40%-preference for pw-patterns defining phonotactic constraints, but such pw-patterns also show up at any other position. This is different with the dense cluster set off to the far right in Figure 4. This cluster shows a 35%-preference for stress-related patterns that is not repeated anywhere else. (The only other pattern with increased frequency in this cluster is again phonotactic constraints.)

INSERT FIGURE 4 ABOUT HERE

Figure 4: Multidimensional Scaling results

4.4. Discussion

From a probabilistic point of view, the Prosodic Hierarchy Hypothesis leads one to expect that pw-patterns would tend to cluster into groups of similar coherence so that in any given language the pw-patterns form natural hierarchies of increasing size. Our analysis suggests no such trend in the data. And we noted earlier that the Prosodic Hierarchy hypothesis does not find statistical confirmation when interpreted in a categorical way, i.e. as predicting exactly one cluster per language.

The only pattern that emerges from the MDS analysis is the increased proportion of stress-related pw-patterns in one cluster that is clearly set off from all other patterns. Inspection of the actual coherence degrees of these pw-patterns suggests that the reason for their increased cross-linguistic similarity is most probably that they tend to be systematically larger than those of other pw-patterns. Figure 5 plots the density distributions of word coherence degrees for each pw-pattern. Pw-patterns of types “allomorphy”, “allophony”, “dissimilation”, “rhythm” and “other process” were removed from the plots because their respective total frequency in the entire dataset was below 5.

INSERT FIGURE 5 ABOUT HERE

Figure 5: Density of relative word coherence per phonological type (*N* = 368). (The x-axis of each panel shows the relative coherence, the y-axis the density of pw-patterns with that coherence degree.)

The panel for stress-related pw-patterns in Figure 5 shows a density peak at higher coherence levels than in any other pw-pattern type, including tone or vowel-harmony patterns. It seems to be fairly common across languages that stress-defined phonological domains are like the one illustrated by Limbu in (3), with $c = 1$. What seems much less common are stress-related pw-patterns which have a coherence degree $c < 1$ and which thereby show up as similar to non-stress-related patterns in the Multi-Dimensional Scaling analysis.

To be sure, there are cases of stress-related pw-patterns with $c < 1$, but they are far less common. One example comes from Mon (Austro-Asiatic; Bauer 1982). This language exhibits five morpheme types, i.e. proclitics, prefixes, infixes, stems, and enclitics. Word-final stress placement singles out the combination of a stem and its prefix or infix:

(7) a. *ˈlàc* ‘break down’
b. *pə-ˈlac* (CAUS-break.down) ‘tear down’
c. *k-ə-ˈlɒʔ* (cross-CAUS-over) ‘take across’
d. *ˈpaʔ ˈkɒ ˈklɒʔ* (do CAUS cross.over) ‘make cross over’

In (7a) the monosyllabic stem receives stress as a main syllable. The morphological derivations in (7b) and (7c) result in disyllabic words which take one primary stress on the word-final syllable. (7d) contains a proclitic causative marker *kɒ*: it is not an affix because it allows gapping of its host stem (e.g. in question-answer pairs), and it is not a separate grammatical word because it doesn't have the syntactic possibilities like a full word. At the word level, *kɒ* receives independent stress. This translates into higher acoustic prominence whenever the stress mark coincides with an intonational peak in information structure.

5. A probabilistic universal

The findings from the MDS analysis and the inspection of the density distributions of each pw-pattern suggest there might be a universal trend (i.e. a probabilistic or statistical universal):

(8) Stress-related domains tend to be universally larger than other domains.

A genuine universal must hold independently of genealogical stocks and linguistic areas, i.e. it must be observable to a similar degree in any stock and any area (e.g. Bickel in press; Dryer 1989). In order to control for the potential impact of stock and area onto the distribution of relative word coherence, we submit our data to a multiple regression analysis of a stratified sample of 40 Sino-Tibetan, Indo-European and Austroasiatic languages, containing a grand total of 246 pw-patterns. The reason for selecting just these stocks is that their geographical distribution allows areal stratification in the sample since all three stocks overlap in South Asia and two stocks (Sino-Tibetan and Austroasiatic) also overlap in Southeast Asia — both regions with well-established linguistic areality.

Formulated as a multiple regression problem, (8) predicts that coherence degrees are systematically affected by a binary factor distinguishing between domains that are defined by stress and other domains, independent of area and stock. If there is such an effect, stress-related pw-patterns will have, on average, higher coherence degrees than other pw-patterns, i.e. the mean coherence will systematically differ between the two types. The coherence degree may also be affected by factors of stock and area, but to the extent that (8) is a genuine universal, these factors will not interact with the difference between stress and other pw-patterns. This prediction is captured by the following regression model, where PW-PATTERN is a binary factor defined by the contrast between stress-related vs. other pw-patterns:

(9) $\mu(c) \sim \alpha + \beta[\text{PW-PATTERN}] + \gamma[\text{STOCK}] + \delta[\text{AREA}]$

The factors STOCK and AREA both enter the design as three-level factors, as explained in the following.

5.1 Coding of stock and area factors

In order to achieve sufficient sample sizes for the STOCK factor, we included 11 Austroasiatic, 12 Indo-European and 17 Sino-Tibetan languages. In each case, we chose one representative per sub-branch of the major branches of the stock (e.g. one West Germanic and one North Germanic language), assuming the standard genealogical trees available in AUTOTYP (Bickel and Nichols 1996ff). Where we had access to sufficient data and also knew that there are considerable within-subbranch divergences in the historical phonological developments, we included two languages per sub-branch. We did this in four cases: (i) from the Tibetan sub-branch of Bodish (Sino-Tibetan) we included both Kyirong (Nepal, Central Tibetan) and Dege (Tibet, Eastern Kham); (ii) from the Aslian sub-branch of Mon-Khmer (Austroasiatic) we included both Jahai and Semelai; (iii) from the Italic sub-branch of Italo-Celtic, we included both (Colloquial) French and Spanish; and (iv) from the Indo-Aryan sub-branch of Indo-Iranian, we included both Romani and Nepali as representatives. Map 1 shows the sample languages coded for genealogical affiliation.[6]

INSERT MAP 1 ABOUT HERE

Map 1: Genealogical affiliation of sample languages (black circles: Indo-European; grey diamonds: Sino-Tibetan; white squares: Austroasiatic)

For the AREA factor, we coded each language of the sample as belonging to either Europe, South Asia or Southeast Asia. We followed the standard area definitions from AUTOTYP[7], except that we assigned Lithuanian to Europe on the account that we have no separate category for northern Eurasia; Sepecides Romani to Europe on the account of the overall distribution of Romani; and Armenian and Persian to South Asia on the account of the long-standing historical links between the Caspian region and South Asia. Map 2 illustrates the areal coding.

[6] The full dataset with all coding and all references is available as an electronic appendix from our website (see Note 2)

[7] Again, see our web site for exact definitions and maps.

INSERT MAP 2 ABOUT HERE

Map 2: Areal affiliation of sample languages (black circles: Europe; grey diamonds: South and Southwestern Asia; white squares: Southeast Asia)

However, none of these areal coding decisions had any impact on our results: we performed all analyses reported below also on a sample without Romani, Armenian, and Persian, and there were no differences in the results except for a weakening of effects due the reduced sample size.

5.2. Results

For the actual analysis, two languages, Lahu (Sino-Tibetan; Matisoff 1973) and Vietnamese (Austroasiatic; Schiering, Hildebrandt and Bickel 2007) were excluded because we find no evidence for pw-patterns in the sense defined here, i.e. phonological patterns that are lexically general and that reference a morphological domain smaller than a phrase or a compound.

We first tested the additive model in (9) against a model with interactions between all factors (i.e. PW-PATTERN * STOCK * AREA). A Monte-Carlo permutation test[8] suggests that there is no significant interaction between any factor. There are significant main effects of STOCK ($F(2) = 11.40$, $p < .001$) and PW-PATTERN ($F(1) = 21.80$, $p < .001$). AREA, by contrast, has no significant effect ($F(2) = 1.64$, $p = .30$).

[8] As noted earlier, since our data are not randomly sampled, we applied permutation testing throughout. All results are based on 10,000 random permutations (cf. Janssen, Bickel, and Zúñiga 2006).

The number of stress-related pw-patterns (19) is much smaller than the number of other pw-patterns (227), and this makes it important to subject the PW-PATTERN factor to a Reliability Analysis (Janssen, Bickel, and Zúñiga 2006). This analysis tests the significance of PW-PATTERN in 10,000 randomly permutated samples, where the largest, then the second-largest, then the third-largest etc. *c*-value from stress-defined pw-patterns are replaced by the grand mean, and where the smallest, then the second-smallest, then the third-smallest etc. *c*-value from the other pw-patterns are replaced by the grand mean. The results of this suggest that PW-PATTERN is retained as a significant factor at an α-level of .01 up to replacing the 5 stress-defined pw-patterns with the largest *c*-values and up to replacing at least 19 of the other pw-patterns with the smallest *c*-value.[9] Thus, the results obtained here would be retained at a .01 α-level even if up to 5 of the largest stress-defined pw-patterns turned out to be misanalyzed or if they hadn't ended up in our sample; and the results obtained would be retained if our data grossly underestimated the coherence degrees of at least 19 other pw-patterns.

This leaves us with a simple additive model including PW-PATTERN and STOCK as factors. In order to assess the overall model fit, we calculated the coefficients in the model and compared the predictions derived from them to the observed density distributions. For the purposes of coefficient estimations, the STOCK factor was parameterized into two binary factors with Austroasiatic (AA) as the (arbitrary) baseline, i.e. coefficients are estimated as contrasts between Sino-Tibetan (ST) and Austroasiatic and between Indo-European (IE) and Austroasiatic. The resulting model is:

[9] We ran the simulations up to the replacement of 19 pw-patterns, which is the maximum for stress-related patterns. But it is likely that the findings are retained after replacing many more than 19 non-stress pw-patterns.

(10) $\mu(c) = .69 + .27$ [STRESS vs OTHER] - .30[IE vs AA] - .14 [ST vs AA]

Figure 6 plots the predictions of this model, together with 95% confidence intervals, over the observed density distributions for the two types of pw-patterns, separately for each stock.

INSERT FIGURE 6 ABOUT HERE

Figure 6: Predictions of the model in (10) (black dots for the means, lines for the 95% confidence intervals) and the observed densities of coherence degrees (grey-shaded) across stocks and pw-pattern types.

While the overall fit of model to data is fairly good, the fit is slightly reduced in the stress-related pw-patterns. This is not surprising given the fact that the number of datapoints is considerably smaller.

5.3 Discussion

The regression and reliability analyses provide robust evidence for the hypothesized probabilistic universal. The trend for stress-related pw-patterns to target larger domains than other pw-patterns is independent of linguistic area and stock. While the evidence from our dataset is strong, the sample is clearly limited, and before the effect of phonological type can be fully accepted as a genuine universal, the hypothesis needs to be tested against further samples in other parts of the world.

The statistical results also suggest that genealogical affiliation significantly contributes to the distribution of phonological domains as well: to a significant extent, the kinds of domains found in a language are determined by the individual phonological 'signature' that the language has inherited from its family. Since there is no significant interaction between the PW-PATTERN and the STOCK factor, however, this means that at the same time, stress-defined pw-patterns are always larger than others, across families (as is also evident in Figure 6). In other words, the distribution of pw-patterns is best predicted by appeal to both their phonological type and the historical phonology from which they have developed.

Interestingly, the AREA factor has no effect. This is a surprising result because especially Southeast Asia is known to be prosodically fairly homogenous across families. For example, Sino-Tibetan and Austroasiatic have similar tone systems, known to have arisen in many cases through language contact (Matisoff 2001, among others). Apparently, relative word coherence is more faithful here to the individual families and has escaped the areal assimilation pressure. The reasons for this are a question for further research.

6. Conclusions

Prosodic Phonology assumes a finite list of prosodic domain types which are hypothesized to be construed in accordance to the Strict Layer Hypothesis. Prosodic domains are conceived as pre-existing entities which can be discovered by phonological patterns which reference them within or across languages. The theory thus states a number of absolute universals and makes numerous predictions with respect to the manifestation and architecture of prosodic structure. From a cross-linguistic perspective, strict adherence to the methodological procedure provided by the theory yields contradictory results in a number of cases (see Schiering, Hildebrandt and Bickel 2007 for extensive discussion). In the present chapter, we applied quantitative methods in order to establish what, if any, principles guide the distribution of phonological word domains across languages.

In sharp contrast to the predictions of the Prosodic Hierarchy framework, we find a cross-linguistic trend for languages to multiply prosodic domains between the foot and the phrase. But the observable diversity is not without limits. We find tentative statistical support for the following probabilistic universal, which remains to be tested in other samples from other parts of the world:

(11) Stress-defined domains tend to be significantly larger than other domains.

Compared to the huge body of absolute universals enshrined in the Prosodic Hierarchy, this universal seems to have small scope and to be rather local. However, this probabilistic universal is empirically founded and, at least in our sample, stands the test for genealogical and areal bias. No other pw-pattern has a systematic impact on domain size. This is particularly noteworthy with regard to pw-patterns of vowel harmony, which, like stress, are intrinsically 'relational' in the sense that they both involve syntagmatic relationships between smaller units (syllables, morae, vowels). Despite this, vowel harmony pw-patterns do not cluster with stress patterns, nor do they show similar degrees of coherence. Tonal pw-patterns, too, are different from stress with regard to coherence: although both are 'suprasegmental', there is no trend for the two pw-patterns to show similar coherence degrees.

The finding that stress-defined domains behave differently from other word-related domains is compatible with pre-generative conceptions of prosodic structure. Pike (1945), for example, used the hierarchically organized domains phoneme, syllable, stress group, pause group and breath group in his analysis of American English. The last three of these are designated to provide domains for stress and intonation phenomena which are necessarily included in hierarchical structures. It is a task of further research to flesh out a theory of hierarchical structure in strictly prosodic domains.

In this light, the present study illustrates how quantitative methodologies that are standardly used in other disciplines yield testable cross-linguistic generalizations. These generalization in turn form a robust empirical foundation for theory construction that accounts for diversity as much as for universality. Hopefully, future investigations into deviations and universals of prosodic structure will further acknowledge the typological variation to be found in the languages of the world.

Acknowledgments

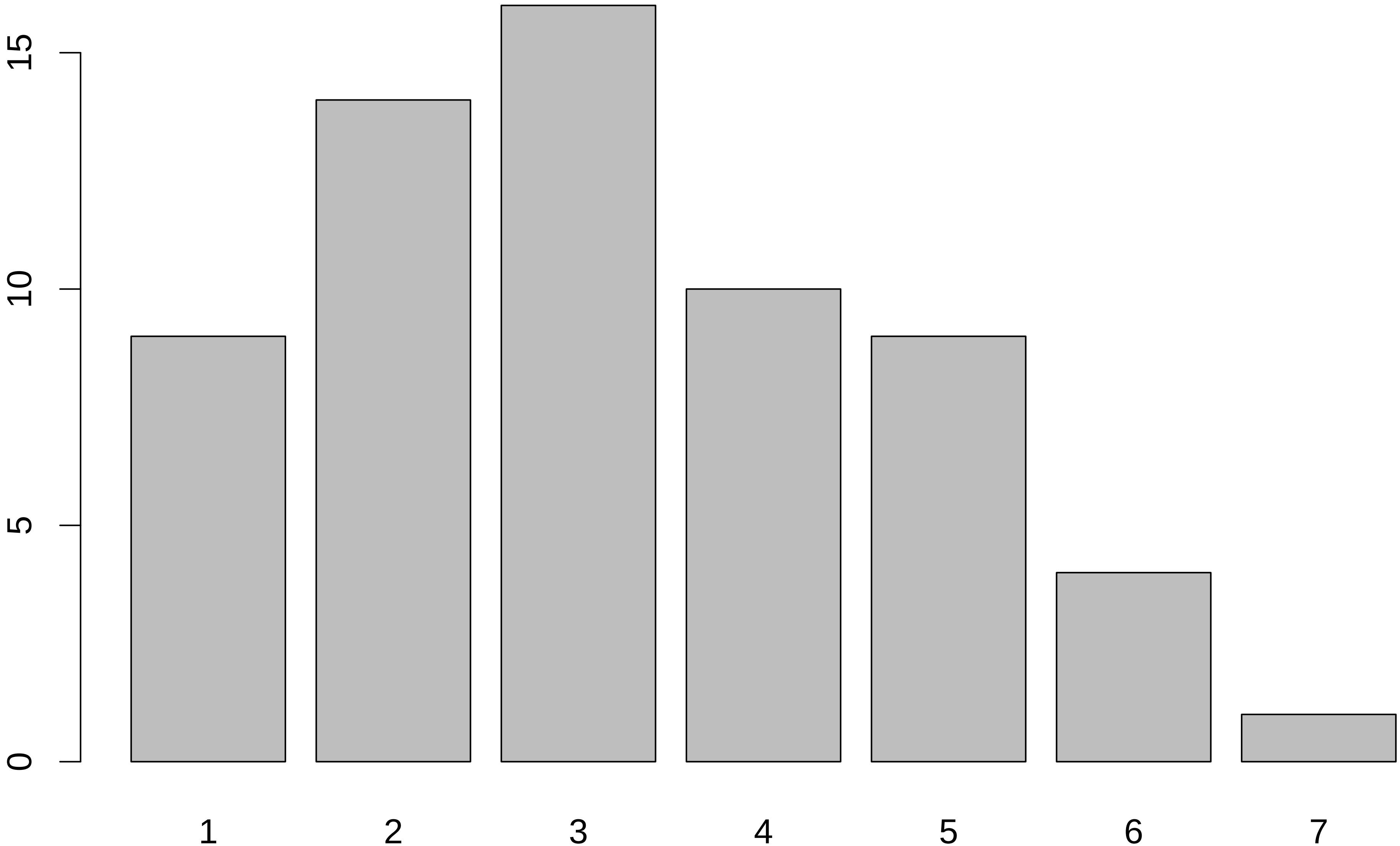
We gratefully acknowledge our student assistants' help in data collection and databasing: Thomas Goldammer, Franziska Crell, Sven Siegmund, Taras Zakharko, Jenny Seeg, Sebastian Hellmann, and Josh Wilbur. Much of our understanding of the Prosodic Hierarchy has profited from discussions with Tracy Hall. We are also grateful to useful comments on an earlier draft by two anonymous reviewers, and to Mathias Jenny for helping us with the Mon data. This research was funded by the DFG under research grants BI 799/2 and 799/3. All statistical analysis and all plots were done in R 2.6.1 (with added packages *cluster, coin, lattice, MASS,* and *trotter;* ff). Maps were created running Hansjörg Bibiko's iAtlas tool on our FileMaker Pro database.
Previous versions of this paper were presented at the Workshop "Prosodic Domains: Universals and Deviations" at the 29th Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft, Siegen, February, 2007, at the 15th Manchester Phonology Meeting, May, 2007 and at the 7th Biannual Conference of the Association for Linguistic Typology, Paris, September, 2007.

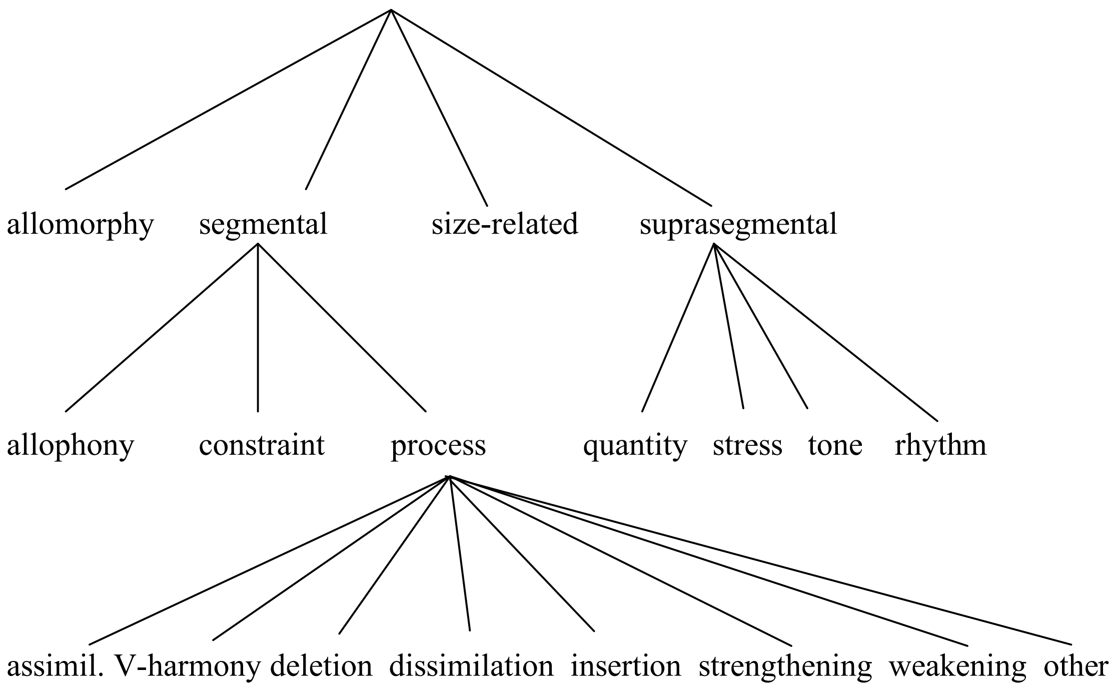
References

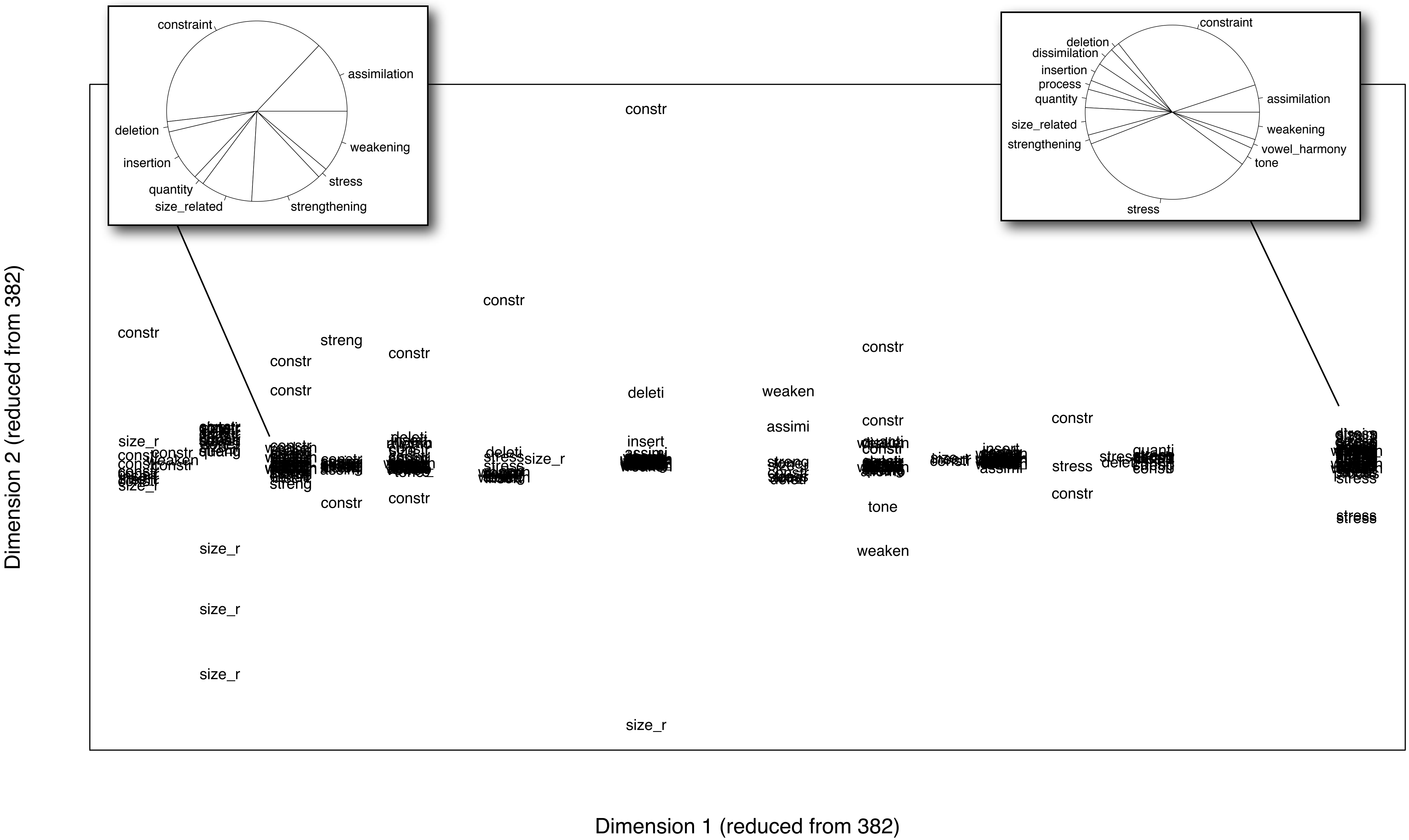
Bauer, Christian
1982 *Morphology and Syntax of Spoken Mon.* London: SOAS.
Bickel, Balthasar
in press A refined sampling procedure for genealogical control. *Sprachtypologie und Universalienforschung* #: ## - ##.
Bickel, Balthasar, and Johanna Nichols
1996ff The AUTOTYP database. Electronic database; http://www.uni-leipzig.de/~autotyp.
Cox, T. F.
1994 *Multidimensional scaling.* New York: Chapman & Hall.
Driem, George van
1987 *A grammar of Limbu.* Berlin: Mouton de Gruyter.
Dryer, Matthew S.
1989 Large linguistic areas and language sampling. *Studies in Language* 13: 257 – 292.
Ebert, Karen H.
1994 *The structure of Kiranti languages.* Zürich: ASAS Press.
Everitt, Brian S., and Torsten Hothorn
2006 *A handbook of statistics using R.* Boca Raton, FL: Chapman & Hall.
Hildebrandt, Kristine
2007 Prosodic and Grammatical Domains in Limbu. *Himalayan Linguistics Journal* 8: 1 - 34.
Hyman, Larry M., Francis Katamba, and Livingstone Walusimbi
1987 Luganda and the Strict Layer Hypothesis. *Phonology Yearbook* 4: 87 – 108.
Inkelas, Sharon, and Draga Zec
1995 Syntax-phonology interface. In: John A. Goldsmith (ed.), *The Handbook of Phonological Theory,* 535 – 549. London: Blackwell.
Janssen, Dirk, Balthasar Bickel, and Fernando Zúñiga
2006 Randomization tests in language typology. *Linguistic Typology* 10: 419 – 440.
Matisoff, James A.
1973 *The Grammar of Lahu.* Berkeley: University of California Press.
2001 Genetic vs. areal linguistics in Southeast Asia: prosodic diffusibility in Southeast Asian languages. In: Aleksandra Y. Aikhenvald, and R. M.W. Dixon (eds.). *Areal diffusion and genetic inheritance,* 291 – 327. Cambridge: Cambridge University Press.
Nespor, Marina, and Irene Vogel

1986 *Prosodic phonology.* Dordrecht: Foris.
Peperkamp, Sharon
1996 On the prosodic representation of clitics. In: Ursula Kleinhenz (ed.), *Interfaces in phonology,* 102 – 127. Berlin*:* Mouton de Gruyter.
1997 *Prosodic words.* Amsterdam: Holland Academic Graphics.
Pike, Kenneth L.
1945 *The intonation of American English.* Ann Arbor: University of Michigan Press.
R Development Core Team
2007 *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing (www.r-project.org).
Schiering, René, Kristine Hildebrandt, and Balthasar Bickel
2007 The prosodic word is not universal. Ms., http://www.uni-leipzig.de/~autotyp/download/SchieringHildebrandtBickel2007.pdf.
Selkirk, Elisabeth O.
1984 *Phonology and syntax: the relation between sound and structure.* Cambridge, Mass.: MIT Press.
1995 The prosodic structure of function words. In: Jill N. Beckman, Laura Walsh_Dickey, and Suzanne Urbanczyk (eds.). *Papers in Optimality Theory,* 439 - 469. Amherst*:* GLSA.
Tumbahang, Govinda Bahadur
2007 A descriptive grammar of Chhatthare Limbu. Ph. D. dissertation, Tribhuvan University.
Venables, W. N., and Brian D. Ripley
2002 *Modern applied statistics with S.* New York: Springer.
Weidert, Alfons, and Bikram Subba
1985 *Concise Limbu grammar and dictionary.* Amsterdam: Lobster Publications.

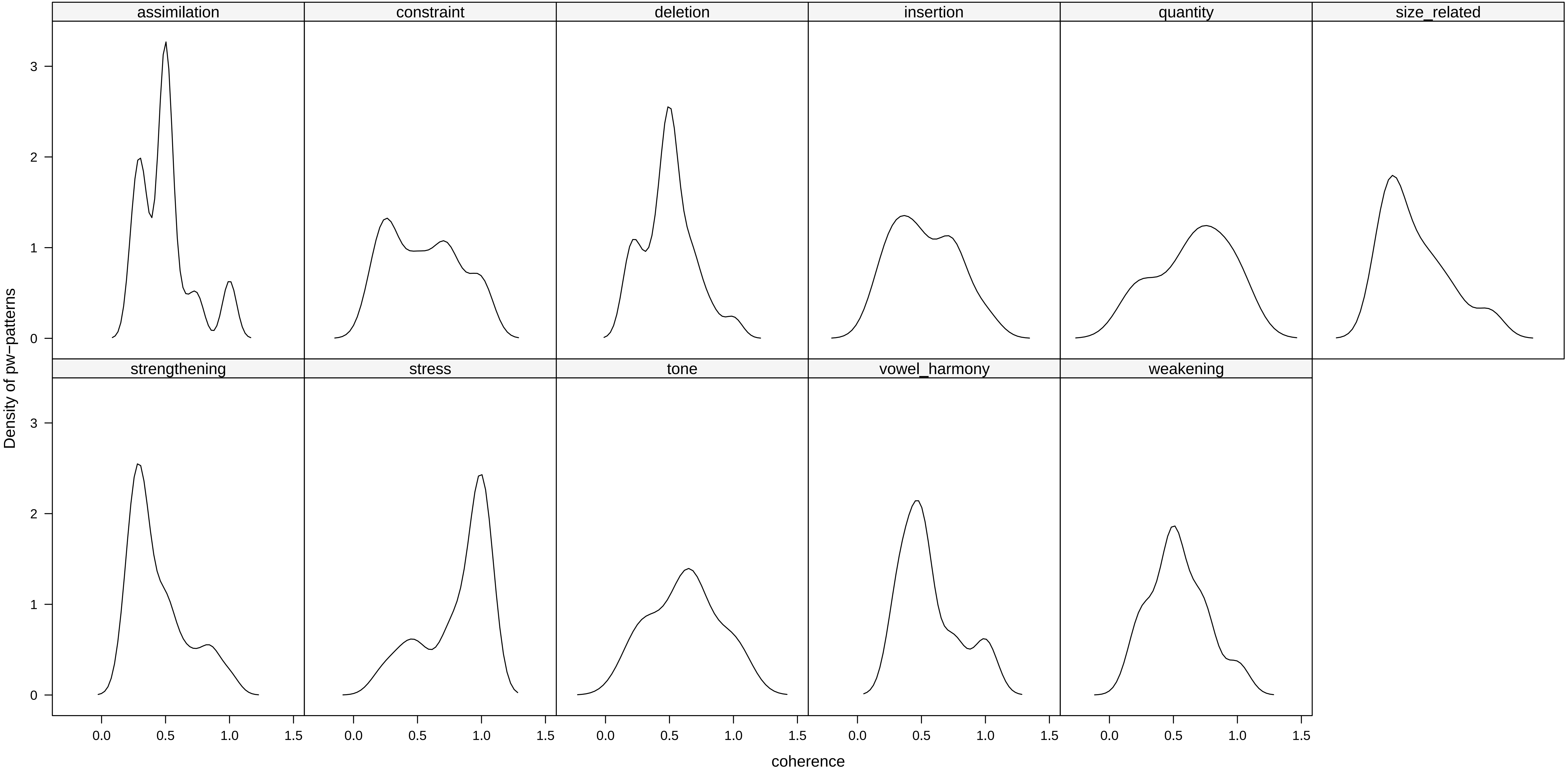
Domain		Phonological pattern
P		Voicing Assimilation, e.g. /p/ → [b]
|		*pe:k-maʔ* ***b****o:ŋ* ‘it’s time to go’ go-INF time
ω_4	prefix-stem-suffix=clitic	One stress per word (among others)
|		(*mɛ-ˈthaŋ-e=aŋ*) ‘they come up and…’ 3ns-come.up-PST=ADD
ω_3	prefix-stem-suffix=clitic	Glottal Stop Insertion (among others)
|		(*ʔa-*)(*ʔi:r-ɛ=aŋ*) ‘we wandered and…’ 1-wander-PST=ADD
ω_2	prefix-stem-suffix=clitic	/m/ → [ŋ] / *ŋ*_ (among other patterns)
|		(*haŋ-****ŋ****ʔna*) ‘sent’ send-PASS.PTPCL
ω_1	prefix-stem-suffix=clitic	/n/ → [ŋ] / _*ŋ* (among other patterns)
|		(*kɛ-****ŋ****-gɦo:s*)*-u-n* ‘you didn’t find it’ 2-NEG-find-3P-NEG
φ		Secondary stress assignment:
|		*ʔaˈʔoŋˌŋe:* ‘my brother in law!’
σ		Canonical syllable template C(G)V(C)


0
5
10
15
1
2
3
4
5
6
7

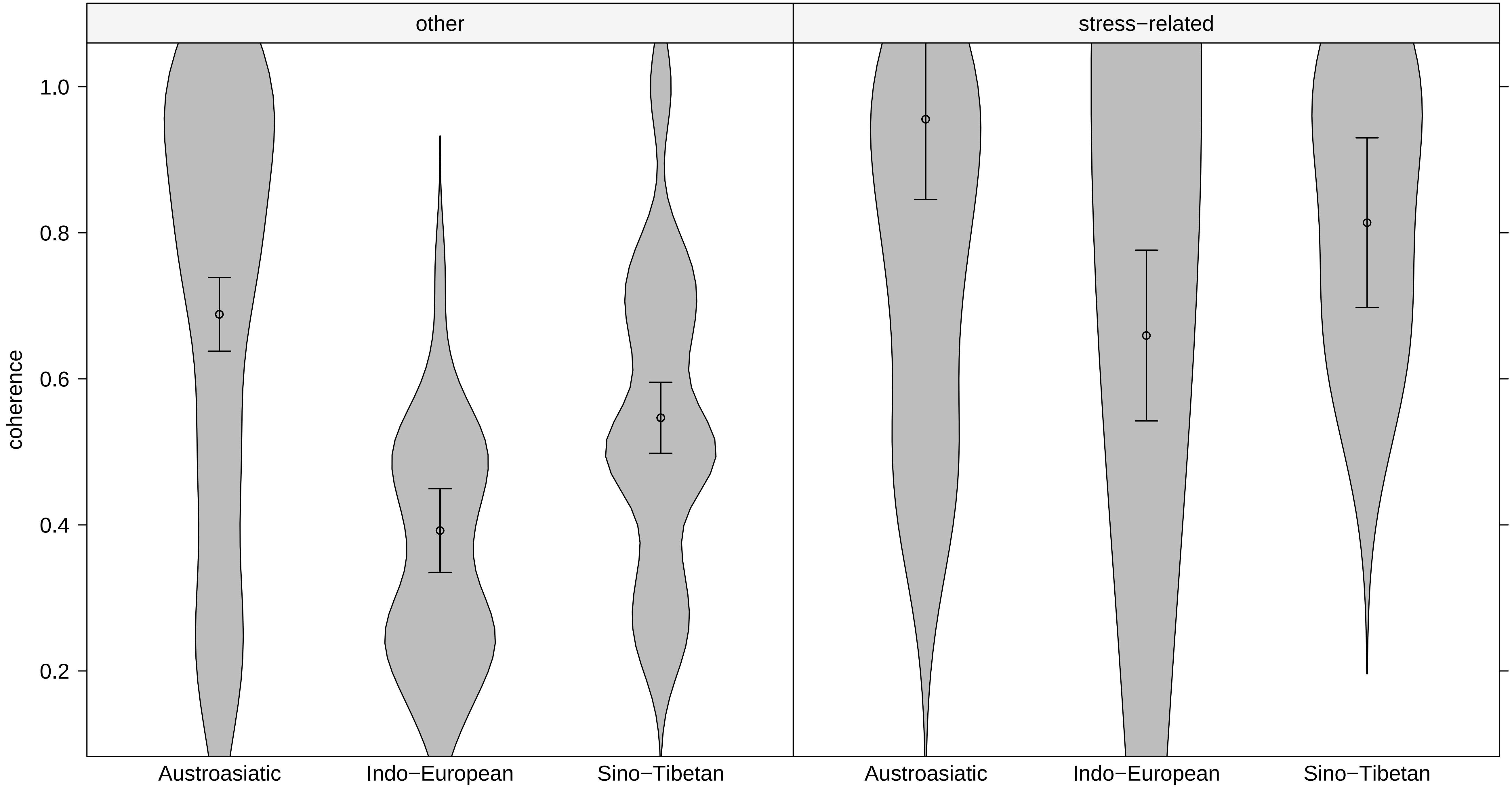

allomorphy
segmental
size-related
suprasegmental
allophony
constraint
process
quantity
stress
tone
rhythm
assimil.
V-harmony
deletion
dissimilation
insertion
strengthening
weakening
other



constraint
assimilation
deletion
weakening
insertion
stress
quantity
size_related
strengthening
constraint
deletion
dissimilation
insertion
process
quantity
assimilation
size_related
weakening
strengthening
vowel_harmony
tone
stress
constr
streng
deleti
weaken
assimi
insert
tone
stress
size_r
Dimension 2 (reduced from 382)
Dimension 1 (reduced from 382)



assimilation
constraint
deletion
insertion
quantity
size_related
strengthening
stress
tone
vowel_harmony
weakening
Density of pw−patterns
coherence
0
1
2
3
0.0
0.5
1.0
1.5


other
stress−related
coherence
1.0
0.8
0.6
0.4
0.2
Austroasiatic
Indo−European
Sino−Tibetan
Austroasiatic
Indo−European
Sino−Tibetan


Swedish
Lithuanian
Irish
Dutch
Polish
French (colloquial)
Spanish
Greek (modern)
Romani (Sepecides)
Armenian (Eastern)
Persian
Kinnauri
Mandarin
Wu
Meithei (Manipuri)
Kharia
Cantonese
Khmu
Kayan Li (Red Karen)
Pacoh
Mon
Cambodian
Car
Jahai
Semelai


Swedish
Lithuanian
Irish
Dutch
Polish
French (colloquial)
Spanish
Greek (modern)
Romani (Sepecides)
Armenian (Eastern)
Persian
Kinnauri
Mandarin
Wu
Meithei (Manipuri)
Cantonese
Kharia
Burmese
Khmu
Kayah Li (Red Karen)
Pacoh
Mon
Cambodian
Car
Jahai
Semelai

	assimilation Burmese 624	stress Burushaski 799	strengthening Kinnauri 294	constraint Kinnauri 299	constraint Kinnauri 300	insertion Persian 705	constraint Semelai 890
assimilation Burmese 624	0						
stress Burushaski 799	0.25	0					
strengthening Kinnauri 294	0	0.25	0				
constraint Kinnauri 299	0	0.25	0	0			
constraint Kinnauri 300	0.25	0.5	0.25	0.25	0		
insertion Persian 705	0.1	0.35	0.1	0.1	0.15	0	
constraint Semelai 890	0.5	0.25	0.5	0.5	0.75	0.6	0