The prosodic word is not universal, but emergent

Schiering, R; Bickel, B; Hildebrandt, K
The prosodic word is not universal, but emergent

RENE´ SCHIERING
Westfälische Wilhelms-Universität Münster

BALTHASAR BICKEL
University of Leipzig

KRISTINE A. HILDEBRANDT
Southern Illinois University Edwardsville

(Received 16 November 2007; revised 31 October 2009)

In Prosodic Phonology, domains for the application of phonological patterns are commonly modeled as a Prosodic Hierarchy. The theory predicts, among other things, that (i) prosodic domains cluster on a single universal set of domains (‘Clustering’), and (ii) no level of prosodic structure is skipped in the building of prosodic structure unless this is required by independently motivated higher ranking principles or constraints (‘Strict Succession’). In this paper, we demonstrate that if, as is standardly done, evidence is limited to lexically general phonological processes, some languages systematically violate the Strict Succession Prediction, evidencing no prosodic word domain, and some languages systematically violate the Clustering Prediction, evidencing more than one domain between the phonological phrase and the foot. We substantiate these claims by in-depth studies of phonological rule domains in Vietnamese (Austroasiatic) and Limbu (Sino-Tibetan). As an alternative to the Prosodic Hierarchy framework, we advocate a heuristic for cross-linguistic comparison in which prosodic domains are conceived of as language-particular, intrinsic and highly specific properties of individual phonological rules or constraints. This allows us to explore empirically the actual degree of variation to be encountered across prosodic systems. It turns out that the ‘word’ has no privileged or universal status in phonology, but only emerges through frequent reference of sound patterns to a given construction type in a given language.

[1] This research was funded by the Deutsche Forschungsgemeinschaft (Grant No. BI 799/2-3). We are indebted to Juliette Blevins, Tracy A. Hall, Larry Hyman, Sharon Inkelas, Jochen Trommer and three anonymous JL referees for helpful comments on an earlier draft of this paper. For ancillary material, including summary reports of individual languages and a downloadable database, see www.uni-leipzig.de/~autotyp.
I. Introduction

A crucial aspect of phonological patterns is the domain in which they apply. The massive body of research in Prosodic Phonology culminates in a prosodic hierarchy that encompasses a finite set of prosodic structures (Selkirk 1980, 1984; Hayes 1989; Nespor & Vogel 2007). Those central to the present study are summarized in (1).

(1) Prosodic domains between the syllable and the phonological phrase

\[ \text{P} \quad \text{Phonological Phrase} \]
\[ \text{ω} \quad \text{Prosodic Word} \]
\[ \varphi \quad \text{Foot} \]
\[ \sigma \quad \text{Syllable} \]

Each of the domains in (1) delimits the application of different phonological patterns. Whereas the syllable (\(\sigma\)) is, among others, the locus of internal phonotactic generalizations like the Sonority Sequencing Principle (e.g. Blevins 1995), the foot (\(\omega\)) forms the domain for rhythm generalizations (e.g. Hayes 1995) and other patterns which reference the combination of syllables. The prosodic word (\(\omega\)) is different from the aforementioned domains in being mapped onto morphological structure, e.g. a stem and its affixes. A number of generalizations, including minimality constraints, phonotactic generalizations, and the application of phonological processes, impinge on this domain cross-linguistically (e.g. Peperkamp 1997, Hall & Kleinhenz 1999, Dixon & Aikhenvald 2002a). The phonological phrase (P) references syntactic phrases and provides the constituent for post-lexical processes. Although the domains in the prosodic hierarchy interact with morphosyntactic structure, the theory does not assume that prosodic and syntactic structures are always isomorphic. Quite to the contrary, mismatches between these modules of grammar justify the formulation of Prosodic Phonology as a component in its own right (Nespor & Vogel 2007: Chapter 2).

A critical methodological assumption of the theory is that it only holds for word-level processes that apply generally across the lexicon and that these processes are distinct from processes at the word level that make direct reference to specific morphological elements (Nespor & Vogel 2007: 18). The

[2] Nespor & Vogel’s (2007: 145–163) original proposal also includes the Clitic Group, which is meant to provide a prosodic domain for phonological patterns restricted to combinations of a host word and a clitic, for example the Latin clitic stress rule. A number of researchers reject the concept and argue for other mechanisms in the prosodization of host–clitic combinations, e.g. prosodic integration, prosodic incorporation, or recursivity (Zec 1988, 1993, 2005; Inkelas 1989; Zec & Inkelas 1991; Booij 1995, 1996; Selkirk 1995; Peperkamp 1997; Inkelas & Orgun 2003).
latter rules would take place in the Lexical Phonology component that precedes the Prosodic Phonology component, although the interaction of these two components might be even more complicated. We refer to this assumption as the ‘Generality Assumption’.

Prosodic Hierarchy theory makes a number of predictions. In this paper, we discuss two of them, labeled the ‘Clustering Prediction’ and the ‘Strict Succession Prediction’, respectively. The Clustering Prediction states that phonological domains cluster on the single universal set of domains enshrined in the prosodic hierarchy. In other words, it predicts that, within and across languages, phonological patterns will be sensitive to these domains and not more (Inkelas & Zec 1995: 547f.). Specifically, the theory predicts that word-related phonological patterns will cross-linguistically converge on one and only one prosodic word domain. This assumption forms the common denominator of most theories of prosodic word structure. Accordingly, more recent Optimality-theoretic treatments of prosodic structure implement the architecture of Prosodic Phonology as representational primitives of the generator (McCarthy & Prince 1986, Prince & Smolensky 1993, see also Selkirk 1995). Lexical Phonology and its successors, although allowing greater variation with respect to morphophonological strata, nevertheless make a principled distinction between the lexical (word-related) and the post-lexical (phrasal) levels (Kiparsky 1982).

The Strict Succession Prediction claims that in the building of prosodic structure none of the levels of the hierarchy can be skipped, i.e. that each level $L_n$ is followed by a level $L_{n-1}$ until the terminal level $L_0$. For example, a prosodic word can only be composed of one or several feet, but not of one or several syllables (Nespor & Vogel 2007: 7). An alternative formulation of the Strict Succession Prediction is in terms of Proper Headedness (Ito & Mester 1992). This principle ensures that every prosodic constituent must have a head, defined as an immediately dominated category of the immediate subordinate prosodic rank. At the word level, this entails that a prosodic word dominates at least one foot.

Whereas most work in Prosodic Phonology accepts these predictions as mostly confirmed, some contributions question their universal validity and empirical basis. As Inkelas & Zec (1995: 548) put it, ‘[w]hile the predictions are clear, in practice the evidence is less so’. In agreement with a growing body of literature challenging many traditionally accepted phonological universals (e.g. Pierrehumbert 2003, Blevins 2004, Mielke 2004), we propose to submit the predictions of the Prosodic Hierarchy to empirical testing. In this paper, we report some of the results from a large-scale typological survey that allows such testing.

In its canonical form, the Prosodic Hierarchy is an absolute universal, an ingredient of our descriptive metalanguage that is deemed necessary in the analysis of any language, on a par with notions such as ‘distinctive feature’. Another way of putting this is to declare the elements of the hierarchy as atoms of Universal Grammar (e.g. Nespor & Vogel 2007). Challenges to such declarations cannot come from individual data or typological surveys but only by demonstrating that the predictions of the hierarchy lead (i) to analyses that are self-contradictory (e.g. claiming and not claiming word status for the same string at the same time) or (ii) to analyses that violate simultaneously held assumptions (Bickel in press). An additional possibility for falsifying such declarations, which we accept in the following, has to do with overgeneration: a declaration fails if it generates more structure than what can be descriptively motivated in at least one language. Some theorists reject this mode of falsification, but even if one does reject it, it seems to us a fruitful enterprise to determine which elements of our descriptive metalanguage indeed are necessary in the analyses of any language, and which elements are not necessary in any language. The latter can be taken to constitute inventories of possible, but not necessary, structures, on a par with such notions as ‘± high tone’. An interpretation of the Prosodic Hierarchy as an inventory, from which languages may chose, has been advocated early on by Selkirk (1980). It seems to us to represent an increasingly widespread view of phonological theory and is the standard view in phonological typology.

In this paper, we attempt to demonstrate that the Generality Assumption either leads to a violation of the Strict Succession Prediction or overgenerates prosodic structure in Vietnamese (Austroasiatic), and that it leads to a violation of the Clustering Prediction in Limbu (Sino-Tibetan). Specifically, we will argue that applying the Generality Assumption in Vietnamese obstructs any evidence for distinguishing between phonological phrases and prosodic words. In the case of Limbu, we argue that more than one domain must be posited between the foot and the phonological phrase, violating the Clustering Prediction and resulting in an analysis under which the same string of morphemes is and is not a prosodic word.

To be sure, deviations from the predictions of the Prosodic Hierarchy have been noted before, but we will argue that the proposed solutions capture only very specific kinds of deviations and may work well only if deviations comprise marginal phenomena. In Sections 2 and 3, we embark on in-depth analyses of Vietnamese and Limbu as representative examples illustrating

[4] When deviations from the Prosodic Hierarchy are discussed in the literature, they are accounted for by including the exceptions in a finite list (Kabak & Vogel 2001, Nespor & Vogel 2007: 109–144), by positing recursive domains (Peperkamp 1996, 1997), and by factoring out prosodic domains to different phonological tiers (Hyman, Katamba & Walusimbi 1987).
that languages may lack or multiply the level of the word in their prosodic structure. Section 4 details an alternative perspective on prosodic domains as conceived of as language-particular, intrinsic and highly specific properties of individual rules or constraints, and lays out the merits of this perspective for cross-linguistic comparison. Section 5 concludes the paper and highlights the consequences of our proposal, namely that the prosodic word emerges through frequent reference of sound patterns to a given construction type in a given language.

2. CHALLENGE 1: VIETNAMESE, OR THE LACK OF PROSODIC WORD DOMAINS

In a number of languages, motivating a prosodic word domain under the Generality Assumption turns out to be a non-trivial matter. For example, in a recent study on the word in Dalabon, a polysynthetic Australian language of Arnhem Land, Evans, Fletcher & Ross (2008) apply a number of phonological tests for the definition of the prosodic word. Neither morphophonemic processes, such as vowel harmony, assimilation or dissimilation, nor phonotactic constraints, nor suprasegmental processes, such as stress assignment, provide conclusive evidence for word boundaries. In their discussion of the relation between the grammatical and the prosodic word, the authors thus have to rely on pause placement and intonation contours. Note that within Prosodic Phonology, pauses and intonation are usually taken to operate on the level of the intonational phrase and not the word.5

Although structurally very different, Vietnamese, a Viet-Muong language of Vietnam, poses a comparable problem to the universality of the prosodic word, since on purely descriptive grounds ‘there is no significant unit in Vietnamese intermediate between the syllable and the phonological phrase’ (Thomas 1962: 521). When limiting the evidence to lexically-general processes, a monosyllabic word in Vietnamese is prosodically indistinct from other syllables and polysyllabic words behave prosodically like other polysyllabic strings, i.e. phrases. If we simply postulate all of the levels in (1) above for this language, we would incorrectly predict the application of word-level processes, i.e. we would overgenerate prosodic structures. If we try to capture the lack of word-level prosodic processes by the exclusion of the ω-node in the prosodic tree, we must conclude that both Clustering and Strict Succession/Proper Headedness are violated. Vietnamese thus constitutes a real challenge to the Prosodic Hierarchy and the Strict Layer Hypothesis.

[5] For a similar case where prosodic word boundaries seem to be equated with intonational phrase boundaries, see Henderson 2002. See also Blevins 2001 for pause placement and the prosodic word in Yurok.
Since prosodic words by definition map onto morphological structure, a discussion of the grammatical word in Vietnamese is mandatory before we examine candidates of word-related prosodic processes motivating a prosodic word domain.

2.1 Morphology and the structure of grammatical words in Vietnamese

Although the definition of the word in Vietnamese has stirred up some discussion (Thomas 1962, Thompson 1963, Noyer 1998), linguists working on the language usually agree on the point that on morphological grounds, one can isolate forms showing different degrees of morphological complexity and phonological length (Emeneau 1951, Thompson 1965, Dinh-Hoa 1997). The examples in (2) summarize the main options for the phonological shape of monomorphemic and polymorphemic grammatical words.

(2) Possible words in Vietnamese (Thompson 1963: 50f.; Nha`n 1984: 181)

<table>
<thead>
<tr>
<th>Monomorphemic</th>
<th>Polymorphemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monosyllabic</td>
<td></td>
</tr>
<tr>
<td>sòm ‘early’, đì ‘go’</td>
<td>d-ây ‘here’, v-ây ‘this way’</td>
</tr>
<tr>
<td>có ‘exist’, ghê ‘chair’</td>
<td>n-ào ‘any’, s-ao ‘however’</td>
</tr>
<tr>
<td>Disyllabic</td>
<td></td>
</tr>
<tr>
<td>Sài-gòn ‘Saigon’</td>
<td>Mỹ-quo´c ‘America’</td>
</tr>
<tr>
<td>va-lì ‘suitcase’</td>
<td>bóí-rói ‘perplexed’</td>
</tr>
<tr>
<td></td>
<td>bàn-ghê ‘furniture’</td>
</tr>
<tr>
<td>Trisyllabic</td>
<td></td>
</tr>
<tr>
<td>Thu-dâu-môt ‘town name’</td>
<td>Liên-hiep guòc ‘United Nations’</td>
</tr>
<tr>
<td>com-mi-nít ‘communist’</td>
<td>ngôn-ngự hòc ‘linguistics’</td>
</tr>
<tr>
<td>Tetrasyllabic</td>
<td></td>
</tr>
<tr>
<td>a-me-ri-ca ‘America’</td>
<td>vö-tuyê´n diênn-thoái ‘radio telephone’</td>
</tr>
<tr>
<td>Hexasyllabic</td>
<td></td>
</tr>
<tr>
<td>—</td>
<td>bóí-rói bóí-rói bóí-rói ‘be very perplexed’</td>
</tr>
</tbody>
</table>

Monomorphemic words are most commonly monosyllabic, the few polysyllabic monomorphemic words are either place names, e.g. Sài-gòn, or loans, e.g. the French borrowing va-lì ‘suitcase’. Polymorphemic words, on the other hand, are generally polysyllabic. The morpheme break-up of the deictic elements d-ây ‘here’ and v-ây ‘this way’ is controversial, but it has no repercussions on prosodic structure Other morphologically complex forms

[6] While we use the rhetoric of morphemic analysis in this paper, none of our claims depends on whether one analyzes morphological structure in terms of morphemes or by some other mechanism of feature realization.

Unless there is need to present the exact segmental composition of Vietnamese forms, we present data from this language in the standard orthography.
are the result of the two productive morphological processes that are found in the language, viz. reduplication and compounding (see below).

The main rationale for considering the forms in (2) as grammatical words is the fact that they can occur as basic free forms and that they have a conventionalized coherence and meaning. However, as we will discuss in Section 2.1.3, their status as grammatical words is problematic with respect to other criteria for grammatical wordhood that have been proposed in the literature, namely the criteria of cohesiveness and fixed ordering (Dixon & Aikhenvald 2002a: 19).  

2.1.1 Reduplication

Reduplication is a morphological process in Vietnamese that expresses a number of meanings, such as distributive, iterative, attenuative, intensive and emphatic (see Thompson 1965: 139ff., Noyer 1998: 80f. for a summary). The various reduplicative patterns with prefixed and suffixed reduplicants are given in (3). In what follows, we underline the bases of the reduplicated forms.

(3) Patterns of reduplication in Vietnamese (Thompson 1965: 139f.)

<table>
<thead>
<tr>
<th></th>
<th>Prefixed</th>
<th>Suffixsed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alliterative</td>
<td>la-lét ‘do with much pain’</td>
<td>rò-rét ‘be very clear’</td>
</tr>
<tr>
<td>Rimming</td>
<td>bôi-roi ‘be uneasy, troubled’</td>
<td>khóc-lóc ‘cry, whimper’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sáu-nhách</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘be absolutely clean’</td>
</tr>
<tr>
<td>Vocalic</td>
<td>léu-láo ‘be ill-mannered’</td>
<td>map-map ‘be fat, chubby’</td>
</tr>
<tr>
<td>Tonal</td>
<td>bô-bû ‘be very big’</td>
<td>dên-dên ‘be rather black’</td>
</tr>
<tr>
<td>Complete</td>
<td>no noi ‘keep talking and talking’ (noi ‘talk’)</td>
<td>sáu sáu ‘be rather clean’ (sáu ‘be clean’)</td>
</tr>
</tbody>
</table>

In alliterative reduplication, only the onset of the base is copied to the prefixed or suffixed reduplicant. Rimming reduplication is characterized by the fact that the rime of both base and reduplicant are identical. In some cases the tone of the reduplicant changes in the process. Comparing the two
constituents in vocalic reduplications, we see that everything except for the vowel is repeated in the reduplicant. In tonal reduplication everything except for the tone is retained in the reduplicant. Finally, since both members of complete reduplications are entirely identical, we cannot decide whether the reduplicant is prefixed or suffixed in disyllabic forms. Apart from the segmental resemblances between the base and its reduplicant, some of these forms also exhibit a special tonal phonology, which we will discuss in some detail in Section 2.2.3.

As noted above, the various processes of reduplication may apply successively and/or recursively on a base, yielding output forms of considerable length. Consider the following examples:


(a) **luc** → **luc-loi** → **luc-luc-loi**
   ‘search’ ‘search inside out’ ‘search over and over’

(b) **ngâm** → **tâm-ngâm** → **tâm-ngâm-tâm-ngâm**
   ‘underground’ ‘very secret’ ‘extremely secretive’

In (4a), alliterative suffixing reduplication applies on the base *luc* in the first step of the derivation. The output is input to a rule of total reduplication applying to both members of the base *luc-loi*, yielding the form *luc-luc-loi-loi*. In (4b), riming prefixing reduplication operates on the base *ngâm* in the first step of the derivation. The output of this process *tâm-ngâm* is then input to tonal prefixing reduplication which results in the form *tâm-ngâm-tâm-ngâm*. With respect to the degree of complexity which reduplications can exhibit, this morphological process behaves like the other productive morphological process in the language, namely compounding.

### 2.1.2 Compounding

Stems can be combined in a number of ways to derive compound words. Thompson (1965) distinguishes two major types of compound structures in Vietnamese, syntactic compounds, which parallel phrases, and non-syntactic compounds, with a structure that is not paralleled by syntactic phrases, see (5).

(5) *Types of compounds in Vietnamese* (Thompson 1965: 126ff.)

<table>
<thead>
<tr>
<th>Syntactic</th>
<th>Non-syntactic</th>
</tr>
</thead>
</table>
| (a) **bàn-ghé** ‘furniture’ (b) **số-hoảng** ‘be terrified’
  (bàn ‘table’+gé ‘chair’)  (số ‘be afraid’+hoảng ‘be panic stricken’)
  ~ **bàn gé** ‘tables and chairs’
| (c) **người-o’** ‘servant’ (d) **học-trờ** ‘schoolchild, pupil’
  (người ‘person’+o’ ‘be located’)  (học ‘to study’+trờ ‘school-age child’)
  ~ **người-o’** ‘person residing’
|
Cross-cutting the distinction between syntactic and non-syntactic compounds, a further distinction can be made between coordinative (i.e. dvandva) formations, in which both elements are weighted equally, and adjunctive formation, which show internal complement–head structure (Nhãń 1984: 270ff.). The generalizing compound in (5a) and the reinforcing compound in (5b) can be said to consist of two coordinated heads. The specializing compound in (5c) and the attributive compound in (5d) consist of a head and a modifier. In the former, the first element is the head and the second element is the modifier, in the latter, the second element is the head and the first element is the modifier. Some compound-like structures, such as quóc-gia ‘nation, country’ (quốc ‘nation’+gia ‘household, establishment’), are exceptional insofar as their components only appear in compound structures. Thompson (1965: 120) calls such elements ‘pseudo-compounds’.

Compounds can recursively undergo further compounding, as shown in (6).

(6) Multiple compounding in Vietnamese (Thompson 1965: 136f.)
   (a) ngọn-ngữ học ‘linguistics’
       (ngọn ‘speech, word’ + ngữ ‘language’ + học ‘to study’)
   (b) vô-điện thoại ‘radio telephone’
       (vô ‘without’ + điện ‘wire’ + thoại ‘speech, conversation’)

In the derivation of the word in (6a), the two bound forms ngọn ‘speech, word’ and ngữ ‘language’ are combined to yield the compound ngọn-ngữ ‘language’, before this latter form is compounded with the stem học ‘to study’. The second derivation, in (6b), involves the combination of vô ‘without’ and điện thoại ‘radio telephone’ yielding the compound vô-điện thoại ‘radio telephone’.

2.1.3 Grammatical wordhood

The outputs of reduplication and compounding often have conventionalized meanings that are no longer compositionally derived. From this point of view, the forms would seem to qualify as unitary grammatical words. But the forms fail on other conventional criteria of grammatical wordhood as terminal nodes in the syntax, such as non-interruptability and ordering constraints. In Vietnamese, polysyllabic strings with a conventional meaning can be interrupted by phrasal elements, regardless of whether the strings consist of one or more morphemes.
Interruptability of Vietnamese words (Nhàn 1984: 6; Noyer 1998: 82)

(a) cà-phê ‘coffee’ vs. cà với phê ‘coffee and the like’
(b) đồ đồ ‘reddish’ vs. đồ không đồ ‘not reddish’
(c) nhà cửa ‘house, home’ vs. Tôi xây nhà xây cửa ‘I build a house.’

In (7a), the monomorphemic French loan cà-phê ‘coffee’ is split by the conjunction với ‘with, and’, which is inserted between the two syllables of the string. The constituents of the reduplication đồ đồ ‘reddish’ in (7b) may also be separated by a phrasal element, in this case the element không ‘not’. The insertion of với and không could prima facie also be analyzed as endoclitis (Harris 2002, Bickel et al. 2007). But such an analysis would not carry over to (7c). Here, the two constituents of the compound nhà cửa ‘house, home’ appear in interlocking order with the adjacent verb xây ‘build’. In such sentential constructions, a reduplicated twin of the adjacent word (xây) in the clause breaks the unity of the compound word. Analyzing xây as endoclitic would entail that most lexical words in Vietnamese can endocliticize, a consequence that challenges the observation that in all other known cases, endoclitis is an extremely limited process. As a result, the data in (7) cast serious doubt on the grammatical integrity of the forms.

In many languages, the constituents of grammatical words are subject to strict ordering constraints not found among phrasal elements. Vietnamese confronts us with variable orders like those presented in (8).

Variable order in Vietnamese words (Nhàn 1984: 6; Thompson 1965: 130)

(a) quần-áó vs. áó-quần ‘clothes’ (quần ‘trousers’+áó ‘tunic’)
(b) chọn lựa vs. lựa chọn ‘to select’ (chọn ‘choose’+lựa ‘choose’)
(c) bôi-rồi bôi-rồi vs. bôi-rồi bôi-rồi ‘be troubled’
(d) com-rom com-rom vs. com-rom com-rom ‘be emaciated’

The first two examples show two conjunctive compounds consisting of two nouns in (8a) and two verbs in (8b), in which the order of the constituents can be varied without a difference in meaning. In (8c) and (8d), we see reduplications with disyllabic reduplicative bases, bôi-rồi and com-rom, respectively. In the output form, the reduplicant of these bases can appear either before the base or after it, without any semantic differentiation. Although it is not clear to us how common such variable orderings are in Vietnamese, the fact that such forms are possible speaks against strict ordering constraints for grammatical words in the language.

There are languages where elements inside a grammatical word can freely permute (e.g. Chintang, Bickel et al. 2007). However, in Vietnamese, it is not even clear whether the domain of free permutation is a grammatical word: the only evidence that they are words is non-compositional semantics, but non-compositional semantics can also be a property of phrases, as witnessed by idioms. Since, as we observed before, the syllables of polysyllabic strings can mix with phrasal elements even if they have a conventional and
non-compositional semantics, it is indeed likely that they are themselves
terminal nodes in the syntax, and not subconstituents of terminal nodes. In
line with this, Nhàn (1984) postulates the syllabeme as the minimal unit of
grammatical analysis in Vietnamese. In the more recent analysis of Noyer
(1998), the distribution of such syllabemes is governed by general rules of
syntax, which operate across various levels of grammatical structure. For
practical purposes, we will use the notion ‘lexical unit’ to refer to syllabeme
strings with conventional and non-compositional semantics.

To conclude, although descriptions of Vietnamese use notions of (gram-
matical) wordhood, there is no evidence that there are terminal nodes in
the syntax longer than one syllable. This characterization forms the basis of
the following discussion of the prosodic properties of monosyllabic and
polysyllabic lexical units in Vietnamese.

2.2 Phonological domains in Vietnamese

In order to determine whether monosyllabic and polysyllabic words of the
kind discussed above constitute prosodic words, we examine all candidates
with respect to word-related phonological properties of the language and
discuss their relation to morphosyntactic structure. It turns out that on
phonological grounds monosyllabic lexical units are indistinguishable from
other syllables and that polysyllabic lexical units are indistinguishable from
other polysyllabic strings.

2.2.1 The syllable and the monosyllabic word

Traditionally, the study of Vietnamese phonology is primarily concerned
with the syllable, which is the domain for phonotactic generalizations, the
distribution of tone and the realization of stress. The table in (9) presents a
schematized representation of syllable structure in Vietnamese.

(9) The syllable in Vietnamese

<table>
<thead>
<tr>
<th>Tone</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Rhyme</td>
</tr>
<tr>
<td>C</td>
<td>(w)V(C)</td>
</tr>
</tbody>
</table>

Each syllable obligatorily starts with an onset consonant, traditionally re-
ferred to as the initial. The initial can be any consonant except for the bilabial
voiceless stop /p/. The rhyme obligatorily consists of a vocalic nucleus, which
is filled by either a short, long or diphthongal vowel. The glide /w/ may
optionally precede the nucleus. The coda position can optionally be filled by
one of the eight segments /p t k m n η y w/. The phonetic realization of the
final velar consonants /k η/ is determined by the quantity and quality of the
preceding vowel. After the short, round vowels /u o ɔ/, the two consonants
surface in their labio-velar allophones, i.e. [kp] and [ɲm], respectively. If the preceding vowel is short and front, they are realized as the palatals [c] and [ɲ], respectively (Pham 2003: 138). This rule establishes the rhyme as a key domain of Vietnamese phonology.

The syllable is traditionally treated as the tone-bearing unit of the language. The tonal distinctions in (10) are contrastive in Vietnamese.

(10) Vietnamese tones (Pham 2003: 10)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Tonal Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>ngang</td>
<td>(high level)</td>
</tr>
<tr>
<td></td>
<td>sác</td>
<td>(high rising)</td>
</tr>
<tr>
<td></td>
<td>ngà</td>
<td>(creaky rising)</td>
</tr>
<tr>
<td>Low</td>
<td>huyê̄n</td>
<td>(low level)</td>
</tr>
<tr>
<td></td>
<td>nâng</td>
<td>(creaky falling)</td>
</tr>
<tr>
<td></td>
<td>hói</td>
<td>(low rising)</td>
</tr>
</tbody>
</table>

The high series comprises the three tones ngang, sác, and ngà. Huyê̄n, nâng, and hói can be grouped in a low series. Every syllable is realized with one of these six tones. Sác and nâng tones have short allotones (called sác2 and nâng2, respectively) whose distribution is restricted to syllables which are closed by a stop consonant. This phonotactic constraint on the distribution of allotones underlines the interdependence of phonotactics and the distribution of tones at the sub-syllabic level of the rhyme. Further evidence for this constituent comes from the phonetic realization of tone, such that the broken part of the ngà tone, for instance, appears in the middle of the rhyme and not in the middle of the vowel (Pham 2003: 141).

According to Thompson (1965: 106–107), each syllable in Vietnamese is accompanied by one stress. At higher levels of prosodic structure, several degrees of stress may be distinguished: weak, medium and heavy stress. The general pattern is that the degree of stress on the syllables within a pause group will be determined by the information load of the different elements, with heavy stress marking contrast, emphasis or the semantically most salient piece of information. In polysyllabic strings, we also encounter a strong preference for iambic stress asymmetries, which will be discussed in more detail in Section 2.2.2 below.

Many languages show a clear distinction between syllable structure and the canonical phonotactic structure of words. Mon, a distantly related

---

[9] Diachronically, the syllable served as the domain for Vietnamese tonogenesis. Whereas the voicing contrast in the onset lay ground for the distinction between the high and the low register, the deletion of final /h/ and /ʔ/, respectively, resulted in additional tone distinctions in both registers (Haudricourt 1954).

We adopt the traditional terminology of Vietnamese linguistics, in referring to the two series of tones as high and low. However, note that the exact phonetic characterization of these registers is more complicated and does not rely on pitch alone (see Pham 2003 for a critical discussion).

[10] There are several obstacles to an unambiguous classification of ngà and hói as high or low tones. We will return to this issue in 2.2.3.
language of the Mon-Khmer family, for instance, allows simple CV syllables with a short vowel like /pa/, i.e. monomoraic syllables. Due to the sesquisyllabic structure of the word, however, such a syllable can only occur as the first syllable of a disyllabic word, e.g. /pa-lac/ ‘tear down’. As soon as a simple CV syllable acquires word status, it surfaces with a lengthened vowel, e.g. /pa/→[?a] ‘go’ (Bauer 1982, Jenny 2005). Whereas the former piece of evidence emphasizes the necessity to distinguish syllable- and word-related phonotactics in Mon, the latter piece of evidence goes to show that the bimoraic minimal word constraint is at work in the language (McCarthy & Prince 1986). Note that establishing the bimoraic word already presupposes a number of levels of prosodic structure, such that the prosodic word dominates a foot domain which consists of two morae. Keeping this mode of reasoning in mind, consider the Vietnamese data in (11).

(11) *Monosyllabic and polysyllabic words in Vietnamese* (Nhàn 1984)

<table>
<thead>
<tr>
<th>Monosyllables</th>
<th>Disyllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) di /di/</td>
<td>lô-mô /lý: mý:/</td>
</tr>
<tr>
<td>‘go’</td>
<td>‘very dim’</td>
</tr>
<tr>
<td>(b) noa/nwa:/</td>
<td>toe-loe/twê:iw:/</td>
</tr>
<tr>
<td>‘lazy’</td>
<td>‘bell-mouthed’</td>
</tr>
<tr>
<td>(c) bay /bay/</td>
<td>bây-nhây/bũy nũy/</td>
</tr>
<tr>
<td>‘fly’</td>
<td>‘very sleazy’</td>
</tr>
<tr>
<td>(d) ngoài /nwa:y/</td>
<td>chuềnh-choâng</td>
</tr>
<tr>
<td>‘outside’</td>
<td>/cwên cwâŋ/</td>
</tr>
</tbody>
</table>

The first column in (11) lists the possible phonotactic shells of monosyllabic words in Vietnamese. The minimal word /di/ ‘go’ in (11a), which can also constitute a minimum pause group, exhibits a simple CV structure and contains a short vowel as its nucleus. The other forms in the first column of (11) add up to demonstrate that the phonotactic shape of monosyllabic words is coextensive with the generally available syllable templates of the language, i.e. CV, CwV, CVC and CwVC. On the basis of this observation, there is no need to posit a bimoraic minimal word. Additionally, the data so far suggest that there is no difference between syllable-related and word-related phonotactics. However, it could be that such differences turn up in larger structures, e.g. the disyllabic word. In order to test whether this expectation is borne out, the second and third columns of (11) give disyllabic reduplicative words with prefixed and suffixed reduplicants, respectively. In each case, the shape of the reduplicants adheres to the syllable template given in (9), i.e. the phonotactic generalizations about the syllable remain valid for each and every syllable in the language, irrespective of their position within larger structures. Note that since the forms given in (11) present the possible inputs for (possibly recursive) processes of compounding and reduplication,
differences between syllable-related and word-related phonotactics cannot be expected in larger words either.

To conclude this section, the phonotactics of the language does not motivate a distinction between syllable-related and word-related patterns, i.e. all syllables exhibit the same structure, irrespective of their grammatical status as affixes, stems, or words. Phonotactically, monosyllabic words are thus indistinguishable from other syllables. Whether suprasegmental features, such as stress and tone, are sensitive to a prosodic word domain will be explored in the next sections.

2.2.2 Stress and the polysyllabic word

Although the analysis of stress in Vietnamese still offers rich research opportunities, the descriptive consensus at this point is that (i) weak and strong stresses alternate in sequences of up to three syllables and (ii) heavy stress is usually placed on the last syllable of the sequence (Thomas 1962: 521). In other words, the basic rhythmic pattern of Vietnamese is one of iambic (or anapestic) phrasing, (syllables with main prominence are marked by '), see (12).

(12) Phonological phrasing in Vietnamese (Thompson 1965: 107)

(‘Nói) (phải 'cé nguôi) (nói 'đi) (nói 'lại) (chổ

speak ought exist person speak go speak come prohibit

'bát) (người ta 'nói) (một mình 'hoài)!

constrain someone speak alone continually

‘For a conversation, [you] ought to have people talking back and forth, not make somebody talk alone all the time!’

The utterance in (12) consists of seven stress units that vary in phonological length from monosyllabic to trisyllabic. The first unit consists of a monosyllabic lexical unit only, whereas the second one is made up of three syllables. Although final stress is the default, the placement of heavy stress within a stress unit is ultimately governed by the semantic saliency of the elements within the structure. This principle accounts for the non-final stress in the second stress unit of the utterance, in which the head có ‘exist’ receives heavy stress in penultimate position of the phrase. As we will see below, such cases of non-final stress placement are also evidenced in polysyllabic lexical units.

In Section 2.2.1, we already noted that each syllable is realized with one stress. Accordingly, the monosyllabic units in (11) will be realized with heavy stress if they constitute a stress unit or even an utterance when uttered in isolation. The crucial question is how stress assignment operates in the polysyllabic strings discussed above. If there is a separate layer of word-level stress in Vietnamese, stress assignment in such polysyllabic strings should differ from the general principles of phonological phrasing illustrated above. Consider (13), listing polysyllabic strings of different morphosyntactic types and exhibiting varying degrees of complexity.
Stress in Vietnamese


(a) va-li ‘suitcase’
(b) nói ‘nói ‘keep talking and talking’
(c) mò ‘mò măng ‘măng ‘deep in the state of dreaming’
(d) người ‘ta ‘somebody’
(e) môt ‘minh ‘alone’
(f) Liên-hợp quốc ‘United Nations’
(g) hoa ‘hồng ‘rose’ (lexicalized, ‘single-word’ reading)
(h) hoa ‘hồng ‘pink flower’ (‘phrasal’ reading)
(i) Tôi không ‘biết. ‘I don’t know.’

Stress placement in lexical units is exemplified in (13a–g), with monomorphemic, reduplicative and compound strings. Stress assignment in monomorphemic units such as (13a) does not differ from that in reduplication (13b–c) or in compounding, as in (13d–g). In each case, the final syllable constitutes the locus for main prominence. Note that the tetrasyllabic reduplication in (13c) is parsed as two stress units in iambic phrasing.\(^{11}\) Since the stress assignment in these structures is identical to that in the phrases (13h) and (13i), we must conclude that stress units in Vietnamese are construed on a number of adjacent syllables irrespective of their internal morphosyntactic composition. Polysyllabic strings with monomorphemic, reduplicative, or compounding structure are thus all subject to exactly the same principles of stress assignment as bona fide phrases. Example (13f) illustrates the only deviation from the default final stress. Here, the pseudo-compound has its origin in structural borrowing from Chinese. In such Sino-Vietnamese compounds, stress is assigned to the left branch of the construction, in which stress is realized on the final syllable.

The examples in (13) all involve strings uttered in isolation. Comparing the two isolated realizations of the morphologically complex forms người ta ‘somebody’ and môt mình ‘alone’ in (13d) and (13e) with their counterparts in actual phrasing in (12), we notice that within larger phrases, the last element is regularly stressed, whereas the compounds do not take heavy stress: người ta ‘nói and môt mình ‘hoài, respectively. This observation provides additional evidence for the analysis of a stress unit in which stress is assigned irrespective of morphosyntactic constituency.\(^{12}\)

---

[11] Iambic phrasing of tetrasyllabic expressions has also been described as the preferred rhythm of Vietnamese poetry, e.g. Làm’troi mưa ‘xuông ‘I pray for the rain to fall’ (see Liêm 1970).

[12] Thompson 1965 reports that the initial syllable takes weak stress in some compounds. For môt mình ‘alone’ in (13e), this analysis would entail that the word is characterized by an iambic pattern of a weak stress immediately followed by a medium stress. However, it is unclear if this stress assignment differs from other phrases in which the semantically less salient element receives weak stress. Another question concerns the status of weak stress assignment. Thompson (1965) describes it more as a tendency than as an obligatory rule.
The phonetic evidence relating to stress is often taken with skepticism, especially when it is based on qualitative auditory judgments of individual analysts. In the case just discussed, one objection would be that although the stress placement in lexical units and phrases is both final, there might be differences in the degree of stress which entails that the phonetic realization of stress at the phrasal level should be more salient. Minimal pairs like the one in (13g–h) constitute perfect test cases to clarify the issue: in terms of stress placement the examples are identical, but with respect to the phonetics of stress we would expect differences if we assume that stress assignment differs between the word and the phrase level. This question has been the focus of a number of recent production and perception experiments by Ingram & Nguyen (2006), Nguyen & Ingram (2006) and Nguyen & Ingram (2007). With respect to the phonetic cues of syllable duration, F0 range, intensity, spectral slope and formant frequencies, no significant difference is detectable between compound-phrase minimal pairs of the type hoa hông ‘rose’ vs. hoa hông ‘pink flower’ under a normal speech production condition of a picture naming task. Under a maximal contrast condition, where native speakers are encouraged to bring out the difference in meaning between the two forms, the disambiguation strategy that speakers chose most often consists in inserting a phrasal juncture between the two constituents of the phrase. This prosodic juncture has a number of phonetic effects such as a pause after the first syllable and concomitant lengthening, widening of the pitch range and lowering of intensity in the pre-pausal syllable. Crucially, however, the chosen disambiguation strategy does not rely on word-level stress assignments.

The acoustic evidence from the production experiments does not support the idea of a separate layer of word stress in Vietnamese. This finding is further backed up by the results of a perception experiment by Nguyen & Ingram (2007). Native speakers showed a response bias towards compound forms when asked to distinguish compounds from phrasal structures on the basis of the stimuli obtained from the normal speech production condition of a picture naming task. When confronted with noun phrases lacking a phrasal juncture between the two constituents, listeners performed at chance level when asked to identify the phrasal structures. Performance was slightly better with stimuli from the maximal contrast condition where the two syllables of the noun phrase were separated by a juncture. This shows that listeners do not rely on any phonetic cues for distinguishing compounds and phrases beyond phrasal juncture cues, which were most pronounced under the maximal contrast condition. In line with their findings from the production experiments, the authors conclude that there is no word-level phonological mechanism for distinguishing disyllabic lexical units from phrases. However, there is evidence for phrase-level phonology, which may be recruited to disambiguate the structures when necessary.
To summarize, Vietnamese provides ample evidence for a genuine stress domain that is preferably disyllabic and maximally trisyllabic. Within this domain, stress is realized on the final syllable in the default case. Crucially, this domain is computed irrespective of the morphosyntactic status of its constituent syllables, i.e. stress phonology does not distinguish between a word-level and a phrasal-level of prosodic structure. Metrically, polysyllabic words are thus indistinguishable from other combinations of syllables. Since the most complex structures which are referenced by the rules for iambic rhythm are phrasal, stress may most adequately be attributed to the prosodic domain of the Phono logical Phrase.

2.2.3 The tonal phonology of reduplication

In Section 2.2.1, we discussed the syllable as the domain of tone and the relevance of the rhyme in the phonetic realization of tone with respect to timing. In running speech, the realization of tones is remarkably stable, i.e. there are no tonal sandhi rules which operate on phrasal or sentential levels (see Pham 2003: 3). However, there are several patterns of reduplication in Vietnamese which result in interesting tone patterns. Recall that the different reduplication patterns illustrated in (3) exhibit different correspondences between the base and the reduplicant, ranging from complete identity to partial segmental identity with respect to initial or rhyme or partial identity in tone. The following data illustrate alliterative suffixing reduplications with bases covering all eight phonetic tones.

(14) Tone in Vietnamese reduplication (Pham 2000: 228)

(a) mau ‘fast’ → mau măn ‘very fast’ (ngang – sác)
(b) laú ‘clever’ → laú lình ‘very clever’ (sác – hỏi)
(c) vát ‘laborious’ → vát va ‘very hard’ (sác2 – hỏi)
(d) mò ‘greasy’ → mò mang ‘very greasy’ (ngâ – hyên)
(e) tàn ‘worn out’ → tàn tà ‘very worn out’ (hyên – náng)
(f) lạnh ‘cold’ → lành lê ‘very cold’ (náng – ngâ)
(g) ngạt ‘severe’ → ngạt nghèo ‘very hard’ (náng2 – hyên)
(h) dơ ‘red’ → dơ đanken ‘very red’ (hội – ngang)

The examples in (14a–d) show reduplications of bases with high tones, namely ngang, sác, sác2 and ngâ, respectively. The bases in the reduplications in (14e–h) carry low tones, namely hyên, náng, náng2 and hỏi, respectively. The tone of the suffixed reduplicant is never identical to the one of the base. However, in the first example, the tone of the base and the reduplicant belong to the same tone series, such that both ngang and sác are high tones. In (14e) and (14g) we find parallel cases with low tones: in the former example we find the combination hyên – náng, in the latter the combination náng2 – hyên. Based on such observations, one may generalize that within
reduplications the tone of base and reduplicant harmonize with respect to tone series (or tone ‘register’).

The formalization of tone series harmony faces a number of problems, however. Recall that ngã is traditionally grouped with the high tones. But in examples (14d) and (14f) this tone patterns with low tones, yielding the combinations ngã – huyê̄n and năng – ngã, respectively. A similar complication arises with hoi̋, traditionally grouped with the low tones, which patterns with high tones in the combinations sâc – hoi̋, in (14b), sâc2 – hoi̋, in (14c), and hoi̋ – ngang, in (14h), respectively. In order to rescue the tone harmony analysis, an additional mechanism must be stipulated that reverses the series membership of the ngã and hoi̋ tones for the sake of the tonal phonology of reduplication, e.g. a rule of Concave Tone Reversal (Nhân 1984: 78) or a Flip-Flop-Rule (Pham 2001). Another option lies in the re-categorization of the tone series, such that ngã becomes a ‘low’ tone and hoi̋ becomes a ‘high’ tone. This approach, however, produces a mismatch between the phonetics of tone, which is the basis of the traditional classification, and the phonological patterning of tone, which is evidenced in tone harmony. Although the mechanisms listed above can account for the deviant behavior of the two tones ngã and hoi̋, they offer no explanation for other deviations from tone harmony. For instance, in the form cuè̄ng-cuè̄ng ‘very hard’ the high tone sâc is combined with the low tone năng. In this case of tone register dissimilation (Thompson 1965: 155–157), tone register harmony is violated, but neither ngã nor hoi̋ is involved.

In order to evaluate the relevance of the distribution of tone within reduplication for the motivation of a disyllabic word domain, the forms discussed above must be situated within the broader picture of Vietnamese reduplication. In Section 2.1.1, the major structural distinctions between the various types of reduplication have already been summarized. Complete reduplication, in which the entire segmental and suprasegmental content of the base is copied in the reduplicant, is a common and fully productive process in the language. Since the reduplicant is an identical copy of the base in these formations, the fact that the two syllables of the reduplication agree in their tone series is not due to an explicit rule of tone harmony, but rather a side effect of the process itself. In a second productive process of reduplication, the reduplicant differs from the base with respect to tone and coda consonants. For instance, when the form biêt ‘disappear’ constitutes the base, its prefixed reduplicant takes the even tone of the same tone series. In this case, the base carries the tone năng of the low series and, accordingly, the reduplicant will be assigned the even low tone, i.e. huyê̄n. Concomitant with

the tone alternation, the coda of the reduplicant is subject to a consonant
alternation in which the final stop is replaced by homorganic nasal. The
output of the reduplication process is therefore *biên-biệt*. It is evident that in
this second reduplication process less specifications of the base are copied
into the reduplicant. At the segmental level, everything except for the manner
feature of the coda consonant is copied. The change from a stop to a nasal
consonant is inherently associated with the reduplication process. At the
suprasegmental level, the tone series is copied, but all other tone features
are provided by a requirement of the specific process, which demands an
even tone.14 Apart from these two productive processes, there are innumer-
able unproductive patterns differing with respect to how much features of
the base are copied and how much features are assigned by the specific
rules underlying the derivation. The shape of the reduplicant in an alliter-
avative reduplication, for instance, cannot be predicted by general phonologi-
cal rules. Although the onset consonant will be copied from the base,
the rhyme can take one of a number of lexically pre-specified shapes.15 A
further obstacle in this context stems from the fact that a given form can be
the input for several reduplication processes. For example, given the base
cúng ‘be hard’, at least the three reduplicated forms cúng-cáp, cúng-cát and
cúng-coi are attested. Finally, note that tone harmony does not in general
apply to monomorphemic and compound lexical units: in such units, tones
of the two series can combine without limit, e.g. bàn-ghé ‘furniture’ (huyền-
sắc).

This brief survey of the nature of Vietnamese reduplication illustrates how
diverse the processes and patterns found in this morphological domain are.
The process and its accompanying phonological concomitants cannot be
captured in a unified way. Quite to the contrary, a number of sub-processes
must be distinguished, each of them exhibiting its own degree of productivity
and phonological predictability. In the extreme case, the idiosyncratic phonol-
ogical relationship between a base and its reduplicant is more efficiently
handled by a lexicon entry than by the formulation of a phonological rule.
As a result, no general rule of tone harmony can be said to characterize
Vietnamese reduplication in its entirety. In line with this, a more adequate
description must distinguish the different reduplicative processes and must
specify in which of these a rule of tone register harmony is at work. We will
elaborate on this point in 4, where we situate these cases in a broader
theoretical context.

[14] Note again that this rule is not without exceptions. Thompson (1965: 172) describes, for
instance, that the ngã tone can pattern with either ngang or huyền, the opposing even tones
of the two registers.

lists approximately 100 of such unpredictable, pre-specified rhymes.
Since the tone patterns under review operate on the output of the morphological process of reduplication, the domain of these patterns could potentially provide evidence for the prosodic word domain (see DiCanio 2005 for register harmony within the prosodic word in Mon and Khmer). However, as just noted above, the tone patterns are subject to lexical specification and only hold for specific subtypes of one morphological process (reduplication). The Generality Assumption that underlies the Prosodic Hierarchy does not admit such patterns as evidence for prosodic structure.

2.3 Discussion

In the previous sections, we surveyed the morphological and phonological properties of ‘words’ in Vietnamese. The grammatical word as terminal node in the syntax has a dubious status since the individual syllables of monomorphic, reduplicative or compounding lexical strings allow variable ordering and can mix with phrasal elements. This suggests that Vietnamese syllables, or ‘syllabemes’, are direct targets of syntactic operations, not shielded off from the syntax by a more complex grammatical word node (cf. already Emeneau 1951: 2–4, 44–45).

In phonology, the structure of monosyllabic lexical units is identical to that of other syllables, whereas the metrical constituency of polysyllabic lexical units is indistinguishable from that of other polysyllabic strings. The tonal phonology of reduplication reveals a lexically specified morphophonological rule of tone harmony that does not count as evidence for a prosodic domain by standard criteria. On the basis of this evidence, we must conclude that there is no lexically general phonological rule which exclusively references the prosodic word in Vietnamese.

2.3.1 A prosodic hierarchy for Vietnamese

The representation of prosodic structure with appeal to the prosodic hierarchy introduced in (1) is far from trivial for Vietnamese. A stubborn application of all prosodic layers defined a priori yields the following prosodic tree:

---

[16] This characterization is in line with Nhàn’s (1984: 56) equation of the syllable boundary and the word boundary in his analysis of Vietnamese.
A prosodic tree for Vietnamese

This representation adheres to the architecture for prosodic structure expressed in the Prosodic Hierarchy and the Strict Layer Hypothesis. Each level between the syllable and the phonological phrase is employed in such a way that neither Clustering nor Strict Succession, nor Proper Bracketing is violated. However, the prosodic tree in (15) suffers from severe discrepancies with the actual empirical findings.

With respect to the first level of prosodic structure, the construction of syllable nodes is uncontroversial and follows the general phonotactic and suprasegmental generalizations concerning this domain as exemplified in Section 2.2.1 above. The evidence for the foot nodes is less straightforward. In the context of the minimum pause group /di/ ‘go’, we noted that the bimoraic minimal word constraint is not at work in Vietnamese. Accordingly, there is no evidence for a bimoraic foot dominated by the word node to provide a domain for the application of rules to fulfill minimal word requirements, such as compensatory lengthening. In (15), the foot node in some cases merely captures the fact that an element is bimoraic, e.g. /tôi/ ‘I’. In the other cases, however, it would inadequately predict bimoraicity where it is not manifested or it would predict the application of foot-based processes, such as compensatory lengthening, where in fact none apply. This problem of overgeneration continues at the level of the prosodic word. As soon as this level of structure is present in the representation, it predicts the application of word-related phonological processes. The survey in Section 2.2 demonstrates that there is no positive evidence for such a domain. Moreover, experimental findings (Ingram & Nguyen 2006, Nguyen & Ingram 2006, 2007) suggest that this domain lacks psycholinguistic reality. The last node of the prosodic tree discussed here, the phonological phrase, seems to be as uncontroversial as the syllable node. The insights from stress phonology motivate this domain as the locus for the iambic rhythm of the language.
Overgeneration is thus the major problem of strict adherence to the Prosodic Hierarchy and the Strict Layer Hypothesis in the context of Vietnamese. Most of the predictions made by the prosodic representation are not manifested by phonological patterns evidenced in the language. Paring the tree down to those elements for which there is positive evidence in Vietnamese leaves us with the structure in (16), which violates the Strict Succession Prediction.

\[ (16) \text{ A revised prosodic tree for Vietnamese} \]

\[ \text{Tới 'dèn nha, mà 'tói mở cửa 'ra, tôi 'vô. I arrive house mother I open door exit I enter} \]
\[ \text{‘I arrived at the house, my mother opened the door, and I went in.’} \]

2.3.2 Skipping levels of prosodic structure

The discussion so far suggests that the syllable is the direct daughter of the phonological phrase in the prosodic representation of Vietnamese and that the building of prosodic structure in this language must consistently skip the level of the word (and the foot), i.e. Strict Succession must be violated in every prosodic tree.

In the literature on prosodic domains, a number of prosodic structures which violate Strict Succession have been proposed and defended, and it has been shown that only such structures can correctly predict the application of the domain-specific rules involved. For instance, a number of prosodic structures violating Strict Succession can be motivated for procliticization in Dutch (in (17a)) and German (dialectal, in (17b) and Middle Frankish in (17c)).

\[ (17) \text{Violations of Strict Succession (Kabak & Schiering 2006: 82)} \]

\[ \text{(a) k vind \quad \sigma \quad \varphi \quad \omega} \quad \text{‘I find’} \]
\[ \text{(b) đan man \quad \sigma \quad \varphi \quad \omega} \quad \text{‘the man’} \]
\[ \text{(c) durks dorf \quad \varphi \quad \omega} \quad \text{‘through the village’} \]

The examples above show that proclitics can constitute single segments, syllables or feet which are adjoined to the prosodic word at the level of the phonological phrase. In each case, the left branch of the tree violates Strict Succession.
Succession in skipping the level of the word, (17c), the level of the word and the level of the foot, (17b), and the level of the word, the level of the foot, and the level of the syllable, (17a). The prosodic trees nevertheless correctly predict the facts encountered in the various languages. For instance, the structure in (17c) correctly predicts the application of foot-based processes such as trochaic rhythm and syllabification within function word contractions but rules out the application of processes applying across word boundaries within a phonological phrase in such constructions (see Kabak & Schiering 2006 for details). Note that it is still true that a node \( L_n \) dominates one \( L_{n-1} \) constituent.

Data like these, where only one element is dominated by \( L_{n-1} \), motivate Itô & Mester’s 1992 version of Weak Layering. In order to account for word structures such as \( ((\text{tere})_{Q} hi)_{\omega} \) ‘television’ in Japanese, in which the final syllable is included in the prosodic word without constituting a prosodic foot, the authors present a restatement of Prosodic Layering which critically rests on the requirement (Proper Headedness) that every prosodic constituent must have a head, defined as an immediately dominated category of the immediate subordinate prosodic rank. Accordingly, the prosodic trees in (17) are acceptable because they still respect Proper Headedness, such that every phonological phrase contains at least one prosodic word.

But the Vietnamese facts challenge even this weakened interpretation of Layering since the phonological phrase would not contain any single prosodic word and thus violate Proper Headedness. In this respect, the structures which have been discussed in the literature differ significantly from the constellation which we would have to assume for Vietnamese if we do not want to postulate overgenerating structures.

Would Vietnamese fit theoretical expectations better if the prosodic hierarchy were modeled in terms of violable constraints, as in Optimality Theory? In Selkirk’s (1995) approach, for example, the Strict Layer Hypothesis is decomposed into four constraints on prosodic domination: Layeredness (no \( C^i \) dominates a \( C^j \), \( j > i \)), Headedness (any \( C^i \) must dominate a \( C^{i-1} \) (except if \( C^i = \sigma \)), Exhaustivity (no \( C^i \) immediately dominates a constituent \( C^j \), \( j < i–1 \)), and Nonrecursivity (no \( C^i \) dominates \( C^j \), \( j = i \)). Whereas the first two constraints are presented as holding universally, the last two are conceived of as violable. As already said, Vietnamese seriously questions this assumption with respect to (Proper) Headedness, since its prosodic structures contain P-nodes which do not dominate \( \omega \). But if we accept that Headedness is also a violable constraint, we have to ask which constraint overrides what is required by Strict Succession. Selkirk (1995: 453) alludes to Prince & Smolensky’s (1993: 25) *Struc constraint, which ensures that structure is constructed minimally, essentially saying ‘less is best’. The tableau in (18) shows that the empirically adequate tree is chosen as the optimal candidate if *Struc is ranked higher than Headedness.
Candidate (18a), in which every level of the prosodic hierarchy is projected, yields two violations of *Struc, i.e. the unmotivated foot and word nodes. In this constraint ranking, it is discarded even though it does not violate Headedness. Candidate (18b) embodies a violation of *Struc at the word level but also a violation of Headedness at the foot level. Finally, although the prosodic structure in (18c) disobeys Headedness at the foot and the word levels, it wins as the optimal candidate since it respects *Struc completely.

This brief analysis illustrates the fact that recent approaches may well account for the Vietnamese data, if one accepts the basic assumptions of Optimality Theory. However, by allowing Headedness to be violable, the ‘essence of the Strict Layer Hypothesis’ (Selkirk 1995: 443) has to be given up. As a consequence, the Prosodic Hierarchy loses its predictive power in stating what a prosodic structure can look like, what is universal about it, and which prosodic systems are possible in the languages of the world.

2.3.3 Summary

Under the Generality Assumption, two distinct levels of prosodic structure can be established for Vietnamese, namely the syllable as the locus for generalizations related to phonotactics, stress and tone, and the phonological phrase, which provides the basis for the iambic (or anapestic) rhythm of the language. Neither the prosodic foot nor the prosodic word can be motivated for the language. If we formalize these findings in a structure that encompasses all levels of the Prosodic Hierarchy, we face the problem of overgeneration since several levels are represented for the individual language although they are not actually manifested. In order to avoid this inadequacy, we opt for a structure that dispenses with the unmotivated levels of the foot and the word.\(^{17}\) In the context of the Strict Layer Hypothesis, this

\[^{17}\text{One JL referee suggests an even more radical interpretation of the Vietnamese data. If the phonological phrase we posit for stress assignment is indeed coextensive with syntactic}\]
solution forces us to postulate a prosodic tree which skips the levels of the foot and the prosodic word, and thus violates Strict Succession as well as (Proper) Headedness.

The case study from Vietnamese suggests that the parsing of the prosodic word domain as well as Strict Succession or (Proper) Headedness are essentially violable in individual languages. If this observation also translates to other levels of the Prosodic Hierarchy and other components of the Strict Layer Hypothesis, the overall predictive power of the theory is seriously challenged.

3. Challenge 2: Limbu, or Multiple Prosodic Word Domains

Vietnamese and similar languages challenge the theory of the Prosodic Hierarchy by evidencing fewer domains than expected. Another true challenge, we claim, comes from Limbu, an Eastern Kiranti (Sino-Tibetan) language of Nepal (Weidert & Subba 1985; Michailovsky 1986, 2002; van Driem 1987, 1997; Ebert 1994; Bickel 1998, 1999, 2003; Hildebrandt 2007). For this language, at least two distinct prosodic word domains must be recognized when standard tests for motivating prosodic domains are applied, arranged in such a way that the multitude cannot be reduced by positing additional domains, recursive structures or prosodic subcategorization. We first delimit the grammatical word, and then turn to prosodic domains.

3.1 Morphological domains in Limbu

The grammatical word in Limbu is characterized by the morphological properties of different morphemes, such as stems, affixes and enclitics as well as the morphological processes that govern their combination. Affixes, which are attached to the stem by prefixation, suffixation or circumfixation, are realized in various positions within nominal and verbal word forms. Stems themselves can be combined to form complex compound verbs. Limbu also has a class of enclitics that syntactically have phrasal distribution and scope, but nevertheless are integrated into a phonological domain at the word level.

---

phrases, the postulation of that domain would also be superfluous. In that case, a direct reference mechanism which maps the syntactic tree onto prosodic structure would be a more elegant way to account for the data. If the foot, the prosodic word and the phonological phrase are not identifiable, Vietnamese may in fact be providing evidence for the absence of any prosodic constituency above the level of the syllable.

The same referee points at a language-internal explanation for the lack of the prosodic word in Vietnamese. If our characterization of lexical units as interruptible and permutable is correct, one could conclude that there are no morphological words in the language. If the prosodic word is dependent on morphosyntax in a morphology-prosody mapping, then the absence of a morphosyntactic category results in the absence of a corresponding prosodic category.
3.1.1 Affixation

In both nominal and verbal morphology, the grammatical word may be defined as consisting of grammatical elements that must occur together following a strictly ordered template. Nouns inflect for possessor agreement (in person and number), number, definiteness and case. Definiteness is marked only in the singular (by -pin, which in turn assimilates to -πil before the ergative/instrumental in -le) and only in the absolutive, genitive, and ergative/instrumental case. Since singular number is zero-marked, the two categories appear in a single position: POSS–stem–DEF/NUMBER–CASE.

Nominal inflection is obligatory, except for non-relational nouns where possessor agreement is optional, and for inanimate nouns where number marking is optional. All affixes realizing these categories exclusively subcategorize for nominal stems, which indicates that we are dealing with morphological elements and not independent phrasal elements. The only exception is the ergative/instrumental case, which is also found on finite, nominalized verb forms. However, when the ergative/instrumental appears on such forms (as well when following derivational suffixes in -a), definiteness can no longer be marked, and the sequence definite-ergative (-pin-le > -πille) is reanalyzed as the post-consonantal allomorph of the plain ergative -le which appears only and always after vowels. Derivational morphology of nominal stems shows regular affixation, but also circumfixation. One case of circumfixation is the formation of free from bound color roots such as mak ‘black’. In this derivation, the discontinuous affix ku-root-la is circumfixed to the color morph, e.g. ku-mak-la ‘black’.

Verb morphology makes use of the same affixation types, but it involves many more affixes and categories. The Limbu verb agrees with the intransitive subject, and among transitives, with both the subject and the primary object (i.e. the sole object of monotransitives and the most goal-like argument of ditransitives). Agreement categories are person and number. Verbs also inflect for tense, aspect and mood. The marking of all of these categories is distributed across four prefix and thirteen suffix positions. Evidence that that entire string of prefixes, a stem (or several) and suffixes form a single grammatical word comes from the following: (i) all inflectional categories are obligatory in finite verbs; (ii) all affixes involved are exclusively subcategorized for verb stems and do not appear on other stems or word forms; (iii) some categories are simultaneously expressed by a prefix and one or more suffixes, e.g. negation; (iv) no phrasal element can ever intrude, including clitics. That all inflectional elements are contained within the same grammatical word, and that none of them behaves like a clitic is furthermore

[18] Note that, although he discusses the category and its marking, van Driem (1987) does not gloss definiteness systematically in his reference grammar; also note that Limbu has in addition a phrase-level enclitic article (=n), which appears on either attributes or heads in NPs (see Bickel 1999 for discussion on the syntax of the article).
shown by the fact that agreement affixes (prefixes and suffixes in verbs, prefixes in nouns) regularly co-occur with agreement-triggering NPs in argument position, and that agreement affixes cannot be gapped under identity.

3.1.2 Complex stems

Both verbal and nominal stems can form compounds, but they differ in their morphological structure (and as we will see, in their phonological structure as well). The following are nominal compounds:

(19) Nominal compounds in Limbu (van Driem 1987: 27, 54)\(^{19}\)

(a) ləŋ+yo:ŋ vs. ku-ləŋ+yo:ŋ
    leg+imprint 3poss-leg+imprint
    ‘footprint’ ‘his footprint’

(b) te:ʔl+phuŋ vs. ku-de:ʔl+ku-bhun
    clothes+flower 3poss-clothes+3poss-flower
    ‘garments, clothing’ ‘his clothing’

(c) cum+de:n vs. a-njum+a-nde:n-hai?
    friend+comrade 1poss-friend+1poss-comrade-pl
    ‘buddy’ ‘my buddies’

All these are comprised of two nominal stems and are presented in their bare forms in the first column. The second column illustrates these compounds in affixation. In (19a), one possessive prefix is attached at the left edge of the compound, i.e. in this case, the compound behaves like a simple stem. In (19b), on the other hand, prefixation of the possessive marker applies twice, both to the first and the second member of the compound. For the sake of prefixation, the compound in (19b) thus constitutes two morphological stem domains. With respect to suffixation, example (19c) shows that even such compounds that form two stem domains in prefixation behave like single stems when it comes to plural marking.

Verbal compounds differ from nominal compounds with regard to prefixation possibilities. Consider the data in (20).

(20) Verbal compounds in Limbu (van Driem 1987)

(a) cuŋ+ji:k-maʔ vs. cuŋ+ge-ji:k-ʔ=ʔiʔ?
    cold+cool-off-inf cold+2-cool-off-pst=ʔq
    ‘be cold’ ‘Are you cold?’

[19] The following abbreviations are used in the word-for-word glosses for Limbu: 1 = first person; 2 = second person; 3 = third person; A = transitive subject; ASS = assertive; COND = conditional; CTR = contrastive; DEPR = deprehensative; EXIG = exigency; GEN = genitive; INF = infinitive; IPFV = imperfective; LOC = locative; NEG = negation, negative; NOM = nominalizer; NS = non-singular; P = primary object; PL = plural; POSS = possessive; PST = past; Q = question marker; REFL = reflexive; S = sole argument of intransitives; SG = singular; SUB = subordinator; VOC = vocative; x > y = x is transitive subject, y is primary object; ‘equals’ (=) sign = clitic boundary; ‘plus’ (+) sign = stem boundary.
The first column presents complex stems in infinitival (citation) form. The second column shows affixation to these stems. While in nominal compounds, agreement prefixes are found on one or both stems, in verbal compounds prefixes only ever occur on the rightmost stem.20

3.2 Prosodic words in Limbu

Sino-Tibetan linguistics has traditionally focused on the syllable as the key domain of both phonological and morphological regularities in the languages of the family (e.g. Matisoff 1991a, b), with occasional extensions to the foot (e.g. Bickel 2003). However, a number of studies have revealed the relevance of additional prosodic domains in the application of phonological processes. For instance, the tone domain in many Bodish languages includes roots and various bound morphemes and can, accordingly, best be analyzed with appeal to the prosodic word (Mazaudon 1973, Denwood 1999, Hildebrandt 2003, Noonan 2003). In what follows, we present an analysis of the prosodic word domains in Limbu. We discuss how multiple prosodic word domains are referenced by phonological generalizations and processes.21 As a byproduct it turns out that the syllable and the bimoraic trochaic foot have salient status in the prosodization of words. In order to distinguish the various prosodic words from higher levels of prosodic structure, we also discuss processes that apply at the level of the phonological phrase. The data discussed in the following make our second challenge to the Prosodic Hierarchy obvious enough: phonological processes do not always cluster on one prosodic word domain intermediate between the foot and the phonological phrase.

[20] The reason for this difference is that verbal compounds historically derive from complement–verb structures, where the complement is another verb, as in (20a), a verbal loanword, as in (20b), or a nominal, as in (20c). It is likely that in none of these cases the complement was ever treated as an inflectable verb stem.

[21] In fact, our survey of word-related phonological processes in Limbu is not exhaustive. In keeping with the Generality Assumption, we concentrate on phonological rules which are demonstrably general across the postulated domains. Accordingly, we do not consider rules whose exact domain of application cannot be decided because of their optional application or lack of evidence. But we will return to the discarded word-related processes in Section 4.
3.2.1 The major prosodic word

In Limbu, a number of phonological patterns converge on a prosodic domain referencing the stem, its prefixes, its suffixes, and enclitics. Since this domain includes a larger portion of the grammatical word than the domain to be discussed in Section 3.2.2, we refer to it as the major prosodic word. This domain can be motivated with appeal to at least two processes: the suprasegmental rules of word-level stress placement and the segmental rule of coronal to labial assimilation.

The first piece of evidence for the major prosodic word comes from stress phonology. Since stress in Limbu ‘is not very pronounced and ... non-distinctive’ (van Driem 1987: 15), grammatical descriptions have so far only rudimentarily treated the distribution and placement of primary stress. In the majority of cases, stress placement is predicable in a straightforward manner from the morphological composition of the word. To substantiate the claim that the stress domain includes the stem, all its affixes and enclitics, we present a more thorough analysis of stress in Limbu, based on a detailed phonetic study.\(^{22}\)

(21) Stress in Limbu

(a) /ku-laːp/ [kulaːp]
   3POSS-wing
   ‘its wing’
(b) /peːɡ-i/ [peːɡi]
   go-IPL.S
   ‘We go.’
(c) /a-oŋ-eː/ [aʔoŋeː]
   1POSS-brother.in.law-VOC
   ‘My brother in law!’
(d) /ku-taŋ=me/ [ku'taŋme]
   3POSS-horn=CTR
   ‘its horn, on the contrary’
(e) /me-thaŋ-e=aŋ/ [me’tʰaŋjaŋ]
   3NS-come.up-PST=and
   ‘they come up and ...’

With respect to the distribution of stress, the forms in the examples above all take only one primary stress. This generalization holds true for combinations of prefix and stem, as in (21a); stem and suffix, (21b); prefix, stem and suffix, (21c); prefix, stem and enclitic, (21d); and finally prefix, stem, suffix and enclitic (21e). On the basis of this evidence we can postulate a prosodic word

\[^{22}\] For the purposes of these analyses, the texts ‘An Untimely Death’ and ‘Father-in-Law’ available at http://lacito.vjf.cnrs.fr/archivage/index.html have been analyzed using the PRAAT software for phonetic analysis (Boersma & Weenink 2007).
domain defined by the presence of one and only one primary stress and referencing a morphosyntactic structure which contains the stem, its prefixes, its suffixes and enclitics, e.g. \((me^\prime than\_e\_an)^o\).

The data in (21) also allow us to analyze rules of stress placement within the stress-defined word. (21b–e) show the default pattern, in which primary stress is realized on the stem-initial syllable. Depending on the morphological structure, this entails that stress will be word-initial if there is no prefix present, as in (21b), and on the second syllable if the word is prefixed. A deviation from this default stress placement is illustrated in (21a). Stem-initial stress would result in an illicit iambic foot here and stress shifts to the prefix.

The stress facts exemplified above also suggest an additional level of prosodic organization, namely the foot. In this context, the examples (21c) and (21e) illustrate that within the prosodic word several bimoraic stress domains are constructed in which primary and secondary stresses are assigned. In (21c), the bimoraic stem syllable carries primary stress and the following bimoraic syllable is parsed as a foot for the sake of secondary stress, i.e. \((?a(?o\eta)^o,(\eta e)^o)^o\). The same principle of stress assignment is also evidenced in (21e), such that the string will be prosodized as \((me^\prime(?a\eta)^o,(\eta\eta)^o)^o\). Note that monomoraic syllables within the word remain unfooted. Accordingly, no secondary stress is assigned in (21d) where the final syllable does not constitute a bimoraic foot, i.e. \((ku(?a\eta)me)^o\).

The second piece of evidence which motivates the major prosodic word domain comes from a productive segmental rule of regressive coronal to labial assimilation:

\[
(22) \text{Regressive coronal to labial assimilation (van Driem 1987: 17)} \]
\[
\left\{ \begin{array}{c}
/t/ \rightarrow [p] \\
/n/ \rightarrow [m]
\end{array} \right\} / \left\{ \begin{array}{c}
/m/ \\
/p/
\end{array} \right\}
\]

The rule states that the coronal phonemes /t, n/ regressively assimilate for place of articulation to the bilabial phonemes /m, p/, respectively. This rule is sensitive to the prosodic word domain as defined above and its application includes stems, prefixes, suffixes and enclitics, as illustrated by the following data.

\[
(23) \text{Coronal to labial assimilation in Limbu (van Driem 1987: 17, 136, 230)} \]
\[
(a) /\text{m}^\prime \text{m}^\prime \text{m}^\prime \text{m}^\prime / [?\text{m}^\prime \text{m}^\prime \text{m}^\prime ]
\]
look.at-INF
‘to look at’

\[
(b) /\text{m}^\prime \text{m}^\prime \text{m}^\prime \text{m}^\prime / [?\text{m}^\prime \text{m}^\prime \text{m}^\prime ]
\]
NS.A-NEG-tell-1SG > 3.PST
‘I did not tell him’

\[
(c) /\text{m}^\prime \text{m}^\prime \text{m}^\prime \text{m}^\prime / [?\text{m}^\prime \text{m}^\prime \text{m}^\prime ]
\]
die-1SG.S.PST-COND-IPFV
‘I might die’

686
In (23a), the rule of regressive coronal to labial assimilation applies across the morpheme boundary of the stem cmmet ‘look at’ and the nominalizer suffix -ma?i. In the second example, the application of the rule is demonstrated for a morphologically complex word consisting of a stem, two prefixes and a suffix. The coronal of the second prefix -n- assimilates to the following labial of the stem met and the stem-final coronal assimilates to the labial of the suffix -pa?i. (23c) shows that the rule also applies across the morpheme boundary of two adjacent suffixes. When the conditional suffix -men is followed by the imperfective marker -pa, its final coronal assimilates to the labial of the following suffix. Note that in this example the labial of the second suffix also undergoes voicing assimilation from /p/ to [bh] (discussed in Section 3.2.3). (23d), finally, shows that clitics are integrated into the prosodic word domain for coronal to labial assimilation. The element -phelle is a subordinator attaching to the right margin of a clause. As a result of cliticization, the initial labial of the clitic triggers regressive assimilation of the final coronal in the host word hen ‘what’. Additionally, the initial labial is subject to voicing assimilation from /ph/ to [bh]. In terms of prosodic status, the phrasal subordinator behaves exactly like the imperfective suffix -pa in example (23c).

We can sum up the discussion of coronal to labial assimilation by concluding that this morphophonological process applies across morpheme boundaries in the major prosodic word domain which references the stem, its affixes and enclitics.

Coronal to labial assimilation does not cross word boundaries, as is evidenced by a separate but related allophonic alternation applying to the coronal plosive /t/. Of relevance here are two variants of /t/: an unreleased, glottalized variant [tʰ], which occurs in word-medial coda position, and a variant comprised of a glottal-plus-plosive cluster with simultaneous lateralization [ʔtʰ], occurring in word-final position (van Driem 1987: 3). Only the word-medial, glottalized variant is subject to the coronal to labial assimilation rule, while the word-final, lateralized variant does not undergo this assimilation:

(24) Allophonic alternation of /t/ in Limbu (van Driem 1987: 377, 485, 279)

(a) [ɔɔmmtʰ-n-i-n]
look.at-1 > 2-PL.P-1SG.A
‘I look at you (plural).’

(b) [ɔɔmɔʔtʰ]
look.at
‘look at’

(c) [ɔɔmɔp-maʔ]
look.at-INF
‘to look at’
(d) nirəŋ-sumləŋ [cwaʔt¹ maʔt¹]  bəʔ pikt-maʔ poʔŋ
twice-thrice  water  be.finished  only  suck-INF  EXIG
‘If it [beer mixture] has sat for many days, it will ripen…’

(24a) shows the word-medial allophone of /t/, while (24b) shows the word-
final allophone. Word-medial /t/ undergoes coronal to labial assimilation in
(24c). Crucially, example (24d) demonstrates that across word boundaries
(within a clause), the lateralized glottal-plosive variant occurs, and assimila-
tion does NOT apply.

3.2.2 The minor prosodic word

Although the motivation of the major prosodic word in Limbu is unprob-
lematic with appeal to the standard methodology of analyzing prosodic
structure, a complication arises when other salient word-related phonologi-
cal processes of the language are also considered. These phonological pat-
terns converge on a prosodic domain which is smaller than the major
prosodic word in that it systematically excludes the prefix. We refer to this
domain as the minor prosodic word. This domain can be motivated with
appeal to two generalizations: the [l] ~ [r] alternation and the process of [ʔ]-
insertion.

One regularity referencing the minor word is the alternation between the
two allophones [l] and [r] of the phoneme /l/. The distribution of these can
be stated with exclusive reference to syllable structure, i.e. with reference to
the syllable as a prosodic domain, only for two contexts. First, in loans, /l/ is
realized as [l] in syllable-final position, e.g. beːl ‘plant name’. Second, in
native words, the phoneme surfaces as [r] when it is the second member in
initial clusters, for instance in the second syllable of cək.krək.mə ‘uvula’.
Apart from this, the distribution of the allophones is governed by word
structure. 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 (which, accordingly, could be ana-
lyzed as a vocalic feature: Michailovsky 1986). The realization remains [l]
after closed syllables, i.e. after consonants, see (25).

(25) Distribution of syllable-initial, word-medial /l/ in Limbu

\[ /l/ \rightarrow [r] / \{V, ?\} \]

[23] Although this synchronic pattern resembles rhoticization, such a characterization would be
inadequate from a diachronic point of view. The allophonic variation shown above is the
result of an older merger of */r/ and */y/, leaving */l/ and */r/ in complementary distri-
bution, see van Driem 1990.
The domain to which this distributional rule applies is smaller than the major prosodic word because it always excludes prefixes. Consider the following data, which illustrate the application of the rule in morphologically simple and complex words. The forms on the left show /l/-onsets after open syllables, the forms on the right after closed syllables.

(26) *The [l] ~ [r] alternation in Limbu* (van Driem 1987: 4ff.)

(a) nerkel vs. lupli
   ‘heart’ ‘earthquake’
(b) pha-re siŋ vs. mik-le raŋ
   bamboo-GEN wood eye-GEN color
   ‘the wood of bamboo’ ‘the color of the eyes’
(c) ke-le? vs. me·l-ke·baŋ (< me-n-l-e·baŋ)
   2-say NEG-NEG-know-1SG > 3.PST
   ‘you say’ ‘I didn’t know [it].’

In line with (25), /l/-onsets are realized as [r] in post-vocalic position and as [l] in post-consonantal position (26a–b). However, as shown by the data in (26c), when /l/ is located at a prefix–stem boundary, it does not exhibit the predicted alternation and surfaces as [l] both after open and after closed syllables. This suggests that prefixes are outside the domain in which the /l/-alternation applies. This can be accounted for by positing a prosodic word structure like *ke·(l-e)*. In this structure, where stem-initial /l/ now appears at the left edge of the prosodic word, it always surfaces as [l].

While at the beginning, the domain of the /l/-alternation is delimited by the stem boundary, at the end, it includes everything that is part of the same grammatical word (suffixes, other stems in compounding) and even clitics. The following data show the predicted alternations with enclitics and complex stems.

(27) *The [l] ~ [r] alternation with enclitics and complex stems*

(a) pe·g-i=ro: vs. pe·g-an=lo:
   go-PL=ASS go-1SG.PST=ASS
   ‘Come on, let’s go!’ ‘I’m on my way!’
(b) kəŋ le·s-u=ra ça vs. maŋgha ke·n-nis-u-n=la ça
   this know-3P=DEPR far 2-NEG-see-3P-NEG=DEPR
   ‘He appears to know.’ ‘You seem to be myopic.’
(c) yaʔ+rak-maʔ vs. lə:k-maʔ
   paddy+trample-INF trample-INF
   ‘to perform the rice dance’ ‘to trample, stamp, walk underfoot, kick’

[24] Limbu does not have prefixes which start in /l/. There is thus no way of telling whether prefixes or prefix clusters form their own prosodic domain with respect to the [l] ~ [r] alternation.
In all cases, /l/-onsets surface as [l] after consonants and as /r/ after vowels and the glottal stop within the minor prosodic word.

If, as we claim, the domain includes everything following the initial stem, this predicts that prefix–stem boundaries inside this domain would show the /l/-alternation, unlike boundaries between prefixes and the initial stem. Domain-internal prefix–stem boundaries are regularly found as a result of verb compounding, because, as noted in Section 3.1.2 above, only the rightmost stem can ever host prefixal morphology in such compounds. The prediction is borne out by the following data.

\[(28) \text{ya?} + \text{ge-ra:kt-u} \]
\[\text{paddy} + \text{2-trample-3P} \]
\[\text{‘You performed the rice dance.’}\]

In sharp contrast to the non-cohering status of the prefix with simple stems (see (26c)), the prefix which is attached to the rightmost syllable of the first part of a complex stem is prosodically integrated into the prosodic word domain. The initial segment of \(\text{lak-ma?}\), accordingly appears in word-medial position in such a construction, i.e. \((\text{ya?} + \text{ge-ra:kt-u})_\omega\), and therefore surfaces as [r].\[25\]

The same domain as the one referenced by the /l/-alternation is also relevant for the distribution of non-contrastive glottal stops in Limbu.\[26\] One variant involves glottal stop insertion.\[27\] This is found before vowel-initial syllables at the left edge of initial stems.

\[(29) \text{Glottal stop insertion in Limbu (van Driem 1987: 15)}\]
\[(a) \text{/in?gh?n/ [?in?gh?n]} \]
\[\text{‘message’}\]
\[(b) \text{/ku-in?gh?n/ [ku?in?gh?n]} \]
\[\text{3poss-message}\]
\[\text{‘his news’}\]
\[(c) \text{/a?-ir-\varepsilon/ [Pa?ir?re]} \]
\[\text{i-wander-pst}\]
\[\text{‘We (plural, inclusive) wandered.’}\]

(29a) shows a monomorphemic word \text{in?gh?n} ‘message’ which begins with a vowel underlyingly. In the surface form, the empty onset position is filled by the prothetic glottal stop. In (29b) the same word forms the base for the

\[\text{[25]} \text{Note that in a limited number of cases, where stem combinations have not yet been re-}\]
\[\text{analyzed as complex stems, the second stem initiates its own prosodic word and thus allows}\]
\[\text{the realization of [l] (see Hildebrandt 2007 for discussion).}\]

\[\text{[26]} \text{Another regularity that is likely to be restricted by the same domain is a ban on velar nasals}\]
\[\text{from word-initial position. However, there are no prefixes with a velar nasal onset, and so}\]
\[\text{we cannot know whether the ban extends to all prefixes or only to word-initial prefixes.}\]

\[\text{[27]} \text{Apart from the position discussed here, the glottal stop also appears as a contrastive}\]
\[\text{(lexical) segment in word-medial and word-final position (van Driem 1987: 7).}\]
prefixation of the possessive prefix *ku-. The glottal stop is still inserted to provide an onset for the vowel-initial stem. This shows that the prefix is not included in the domain referenced by the word-based rule of glottal stop insertion, see also (29c). Unlike prefixes, suffixes and enclitics are integrated into the word domain referenced by glottal stop insertion. Therefore, if a vowel-initial suffix or clitic follows a stem or another suffix, no glottal stop is inserted. In these positions, hiatus is not resolved by glottal insertion, but by diphthongization (or, trivially, by resyllabification if a vowel-initial suffix follows a consonant or by glide formation if the suffix follows a vowel that has a glide alternant). The following data show this with minimal pairs of /u–e/, /a–e/, /e–i/, and /a–i/ sequences; the forms on left involve prefix–stem, the forms on the right stem–suffix and host–clitic boundaries, respectively.

(30) **Glottal stop insertion/diphthongization across different morpheme boundaries**

(a) /ku-e:k/ [kuʔe:k] vs. /a-mphu-e:/ [ʔamphuːɛː]

3POSS-back 1POSS-brother-VOC

‘its/his/her back’ ‘Brother!’

(b) /a-e:k/ [ʔape:k] vs. /yuma-e:/ [yumaɛː]

1POSS-back grandmother-VOC

‘my back’ ‘Grandma!’

(c) /ke-im/ [keʔim] vs. /naks-e=i:/ [nakʃeiː]

2-sleep go.crazy-PST=Q

‘You sleep.’ ‘Has he gone crazy?’

(d) /a-iR-e:/ [ʔarɪʔeː] vs. /nu-ba=i:/ [nuˈbaːiː]

1-wander-PST be.alright-NOM=Q

‘We wandered.’ ‘Is this good?’

The only difference between the forms with prefixes and those with suffixes is that prefixes are outside of the prosodic domain in which glottal insertion applies, while suffixes are inside this domain. Thus, the domain is the same as the one observed above for the /l/-alternation, i.e. prefix–(stem–suffix=clitic) or.

To sum up, /l/-alternation and glottal prothesis evidence a prosodic word including stems and suffixes (and enclitics) but crucially excluding prefixes. This is in conflict with other phonological processes of Limbu referencing the major prosodic word domain that includes stems, prefixes, suffixes and enclitics.

[28] There is one context in which glottal hiatus applies with the interrogative clitic =iː, namely when it is attached to a host ending in /i/. The application of the process seems to be conditioned by the impossibility of diphthongizing two identical vowels. Note, however, that this is still different from glottal stop insertion at prefix–stem boundaries. In the latter case, the rule applies irrespective of whether the vowels are identical or different.
3.2.3 The phonological phrase

In order to rule out that the major prosodic word is really a phrasal domain and not a word domain, we turn to higher levels of prosodic structure, i.e. the phonological phrase. There is at least one phonological process that applies across word boundaries within this domain. (31) illustrates that word-initial plosives optionally assimilate to the preceding word with respect to the feature voice (which is not distinctive in Limbu).

(31) Voicing assimilation in Limbu (van Driem 1987: 18f.)
   A: anige hen jo:kmaʔ?  cf. /co:kmaʔ/  
      ‘What shall we do?’
   B: pe:kmaʔ bo:ŋ.  cf. /po:ŋ/  
      ‘It’s time [for you] to go.’

In the first turn of this conversational exchange, the word co:kmaʔ? ‘to do, make, build’ is realized following a word ending in a voiced segment /n/. According to the phrasal rule of voice assimilation, the word-initial unaspirated voiceless lamino-postalveolar affricate /c/ [ʈʃ] is replaced by its voiced counterpart [dʒ] (written as ɲm) in this context. The auxiliary word po:ŋ in the second sentence of (31) assimilates to the preceding word in the same fashion. Note that the glottal stop does not block voice assimilation in this example, cf. its vocalic status noted in (25).

Nominal compounds constitute two word domains for the sake of the /l/-alternation. For instance, in compounds such as haʔ+luŋ (fire+stone) ‘fireplace’ and makhí+lam (blood+road) ‘artery’ the initial segment of the second compound member always surfaces as [l], although it appears after a glottal stop in the first compound and intervocally in the second example. This suggests that /l/ in these examples does not appear in word-medial position, but in word-initial position, so that both stems appear to have prosodic word status, i.e. (haʔ)₀+(luŋ)₀ and (makhí)₀+(lam)₀. But the two prosodic words are conjoined into a single phonological phrase. Let us reconsider the nominal compound cum+de:ŋ ‘buddy’, which is composed of the two elements cum ‘friend’ and te:ŋ ‘comrade’. Here, the initial plosive of the second member surfaces in its voiced counterpart in the compound. In an analysis that treats the two stems as two separate prosodic words which are joined in a phonological phrase, this fact is predicted by the application of the phrase-level rule of voicing assimilation, i.e. [(cum)₀+(de:ŋ)₀]₀. In summary, nominal compounds fit into the higher domain structures in constituting two prosodic words, which are in turn subsumed under a single phrase node.

Our knowledge of Limbu phonology is still too limited to propose an analysis of intonation and other phonological processes which might allow the postulation of further prosodic domains for the language, such as the intonational phrase and the utterance. However, our point is that by
concentrating on the phonological processes which operate between the levels of the syllable and the phonological phrase we find highly specific and overlapping domains whose structure is not predicted by the Prosodic Hierarchy.

3.3 Discussion

In the previous section, we discussed the word-level prosodic constituency in Limbu, while making reference to other domains of prosodic structure, namely the syllable, the foot and the phonological phrase. At first sight, the domains which need to be distinguished in Limbu adhere to the prosodic layers predicted by the Prosodic Hierarchy. However, the crucial deviation lies in the fact that the word-related processes themselves do not converge on a single domain, but constitute two domains which reference different portions of the grammatical word. The first one, which references the entire grammatical word + enclitics, is evidenced by stress assignment and coronal to labial assimilation. The second one, which references the initial stem, subsequent stems, suffixes and enclitics, but which excludes the prefix, is instantiated by the [l]~[r] alternation and the [r]-insertion.

3.3.1 A prosodic hierarchy for Limbu

If we follow the Prosodic Hierarchy blindly, we would have to postulate the following prosodic tree for Limbu.

(32) A prosodic tree for Limbu

```
P
    ω
    |
  φ   φ
  |
σ σ σ
```

This representation makes a number of correct predictions concerning low level constituency in Limbu. The syllable captures the syllabification of the given string, which includes the resyllabification of the stem-final consonant to fill the onset position of the last syllable. The second and third syllables constitute bimoraic feet, which qualifies them as the locus of stress placement at the higher level of the word. Here, the stem syllable is assigned primary
stress and the final syllable is assigned secondary stress. The initial, monomoraic syllable is adjoined to the prosodic word without being parsed as a foot. Although this violates Strict Succession, it is still in line with the principle of Proper Headedness. The word node in this example provides the domain for stress placement and for the application of the coronal to labial assimilation. Finally, the inclusion of the prosodic word in the phonological phrase captures the possibility that the final voiced element of the word will trigger voicing of the initial segment of the following prosodic word within the phonological phrase.

However, the prosodic representation in (32) suffers from some inconsistencies. These relate to the fact that only one word domain can be recruited from the universal repertoire of domain types provided by the Prosodic Hierarchy. With appeal to the prosodic tree given above it is impossible to formalize the fact the [l] ~ [r] alternation and the [ʔ]-insertion reference a domain which is smaller than that delimited by the word node in the tree, i.e. the prosodic tree undergenerates. An obvious solution would be to posit a second word domain intermediate the foot and the major prosodic word to optimize the descriptive adequacy of the representation.

(33) A revised prosodic tree for Limbu

In (33), the minor word node provides the domain for [l] ~ [r] alternation and the [ʔ]-insertion, whereas the major word node provides the domain for stress assignment and coronal and labial assimilation. Although the revised prosodic tree correctly predicts the application of all relevant prosodic patterns, it violates the Clustering Prediction, which specifies that word-related processes should converge on one and only one prosodic word domain. Second, since the major word node immediately dominates a node of the same prosodic category, i.e. the minor word, the prosodic tree also violates Strict Succession or Proper Headedness.
It is noteworthy that a picture similar to that of Limbu is offered by other Eastern Kiranti languages and is likely to be a diachronically robust pattern of this branch of the family (perhaps even beyond that). In Belhare, for example, there is an intervocalic voicing rule applying within stem–suffix strings but not across prefix–stem sequences; and a rule of deleting velar stops at the end of a domain including prefixes along with the stems and suffixes (Bickel 1996, Bickel & Hildebrandt 2005). In Chintang, a rule of glottal insertion (similar to the one in Limbu) and the host definition of endoclitics identifies a domain excluding prefixes; rules of stress assignment and the host definition of enclitics identifies a domain that includes prefixes (along with stems and suffixes); a third domain is referenced by intervocalic voicing which optionally applies within sequences of prefixes, or within sequences of stems and suffixes, but not across prefix–stem boundaries (Bickel et al. 2007).

3.3.2 Multiplying prosodic word domains

Limbu is not the only language which deviates from the predictions of the Prosodic Hierarchy because its prosodic system distinguishes multiple non-isomorphic domains for word-related processes. Better-known examples come from Bantu (Hyman 1987, 1998, 2008; Odden 1987; 1995, 1996; Myers 1998; Downing 2001; Bickmore 2007) or Athabaskan (McDonough 2000). What makes the challenge from Limbu extraordinary is the fact that even the additional mechanisms of prosodization which have been proposed in the literature cannot resolve the problems raised by the data.

One answer to deviations from the Prosodic Hierarchy at the level of the word simply lists them by specifying which morpheme strings (e.g. stem–suffix vs. stem–clitic vs. stem–suffix=clitic strings) support which prosodization. This is obviously not a theoretically appealing solution to begin with, but, worse, any solution along these lines fails to account for the fact that in Limbu, one and the same morpheme string can simultaneously support two non-isomorphic prosodic structures. In a string like ku-la’p (3POSS-wing) ‘its wing’, which surfaces as [’ku-la’p], both major and minor words are present simultaneously. The minor word domain parses the string as ’ku(la’p)o, for the sake of the [l] ~ [r] alternation, which accounts for the fact that the stem-initial /l/ surfaces with its word-initial allophone [l]. The major word domain parses the same string as (’ku-la’p)o, for the sake of stress placement and this accounts for initial stress. Crucially, unlike languages like Dutch where two classes of suffixes can be distinguished which differ in their prosodic status, the prefix in Limbu is at the same time integrated into a word domain and excluded from another. As a result, one would have to simultaneously claim that ku- is part of the prosodic word and that it is not part of the prosodic word.

To make this point even clearer, the Limbu situation is crucially different from well-known cases of non-cohering prefixes as, for instance, in Bantu. In
Kimatuumbi, for example, the object prefixes and the inflectional stem constitute the ‘superstem’ domain. The next higher level domain of the ‘simple verb’ encompasses the superstem plus the preprefixed subject and tense markers (Odden 1996: 71). In such a system, there are distinct types of prefixes, differentiated with respect to which domain they attach to. But it is not the case that a given prefix is at the same time within and outside a domain. It is not the case in Kimatuumbi that the object prefix is part of the inflectional stem for the sake of some process and at the same time outside this domain for another process.  

Yet this is exactly what one finds in Limbu.

A second solution to multiple word domains offered in the literature is to posit an additional prosodic domain. An argument along those lines would re-categorize the major word domain in the prosodic tree (33) as a Clitic Group node (C), yielding the bracketing \((\text{kula}(p)\)\)C for the form discussed above. In Nespor & Vogel’s (2007) original proposal, the C domain references clitic elements which can be identified by morphosyntactic criteria and binds them to their phonological host. The crucial evidence for the motivation of such an intermediate domain comes from regular phonological processes which specifically apply in host–clitic combinations, such as enclitic stress in Latin. The Limbu word domains cannot be resolved with appeal to such a domain because the prefix in \(\text{kula}(p)\) is not a clitic in either phonological or morphological terms: it is not a clitic in phonological terms because all phonological processes that link it to the stem also apply between stems and suffixes (cf. the coronal to labial assimilation and the stress placement rules), i.e. it does not trigger a special host–clitic phonology. The prefix is not a clitic in morphological terms, either: as noted in Section 3.1, Limbu possessive agreement prefixes, like all other nominal affixes, exclusively subcategorize for nominal stems and are therefore not phrase-level affixes; they are regular elements of noun inflection; they cannot be gapped under identity; they systematically co-occur with agreement-triggering argument NPs; and, as agreement markers, one would conventionally analyze their scope as being over the head, not the phrase. Moreover, those elements in Limbu, which are true phrase-level clitics, are regularly included in both word domains under discussion, as demonstrated in Sections 3.2.1 and 3.2.2. Analyzing (despite all evidence) the prefixes as clitics\(^{30}\) would therefore necessitate two distinct Clitic Group nodes, one node that includes enclitics

\[\text{[29] In Bantu, some prefixes are subject to ‘prefix–stem fusion’, such that a phonologically defined set of object prefixes is integrated into the stem domain (Marlo 2008). This is again different from the Limbu case: these specific object prefixes are then integrated into the stem domain and are not simultaneously parsed as superstem prefixes.}\]

\[\text{[30] A similar paradox emerges in the analysis of prosodic domains in Kukuya. Hyman (1987: 332) raises the question whether lexically unfooted prefixes are in fact post-lexical clitics in this language. To the best of our knowledge, this question has, however, not been answered in a theoretically appealing way.}\]

696
and prefixes analyzed as proclitics and one node that includes enclitics, but no proclitics. As a result, the Clustering Prediction would again be violated, only now on the Clitic Group instead of the word level.³¹

A similar problem arises if one tries to address the Limbu challenge by equating the major word domain with the phonological phrase, e.g. \((\text{ku}(\text{la:p})_\omega)_p\). This would simply shift the problem of multiple domains from the word to the phrase level because there already is clear evidence from intervocalic voicing (reviewed in Section 3.2.3) for a phonological phrase that encompasses longer sequences than any of the two word domains.

A third formal mechanism which has been proposed in the literature to account for multiple domains is recursion. One such analysis is Peperkamp’s (1996) analysis of stress assignment in Neapolitan Italian. Applied to Limbu, one would then posit a stacked structure like \((\text{ku}(\text{la:p})_\omega)_\omega\), in which both word domains are of the same rank. In such an approach the smaller domain would first be built, and in a second step, the larger domain would be constructed by the recursive application of the domain-defining processes. But, in contrast to the Italian stress data, the evidence for two domains in \text{kula:p} does not come from recursive application of one and the same phonological process, but from very distinct processes: \([l] \sim [r]\) alternation and glottal stop insertion for the smaller and stress assignment and coronal to labial assimilation for the larger domain. Put differently, a recursive structure entails that the stacked domains are one and the same prosodic domain. This would then predict that they have the same phonological properties. For instance, applied to Limbu, it would predict that the same principles of stress assignment will recursively apply at all levels. In the stacked structure \((\text{ku}(\text{la:p})_\omega)_\omega\), stress should accordingly be assigned to the smaller domain first and then again within the larger domain. The actual form \text{ku:la:p} proves this prediction patently false.³²

A fourth possible solution lies in assigning different domains to different phonological tiers, such as the tone and quantity domains in Luganda (Hyman, Katamba & Walusimbi 1987). However, this solution cannot be carried over to Limbu because the different domains are associated with phonological patterns from different phonological tiers: for example, the segmental process of coronal to labial assimilation and the suprasegmental process of stress assignment converge on the major word domain. In Limbu,

---

³¹ Note that this problem of multiplying domains at higher levels also arises if the ‘Clitic Group’ is redefined as a ‘Composite Group’ as proposed by Nespor & Vogel (2007: xv–xx).

³² One JL referee suggests that domain recursion can be defined independently of recursive rule application. If that is the case, domain recursion becomes void of empirical predictions, and would be equivalent of positing two arbitrary prosodic domains between phrase and foot. The only difference from our conclusion would presumably be that we do not expect domains to be properly nested in each other, whereas a recursion mechanism would predict this. Crucially, however, both analyses result in violating the expectations from the Prosodic Hierarchy theory.
the different prosodic word domains are not exclusively confined to one specific phonological tier but cross-cut several components of the phonology.

A fifth solution is to assume a much more fine-grained theory of prosodic structure (Downing 1999, 2001). Even under this approach, however, there remain crucial problems for the analysis of Limbu. Following the proposed distinction between the prosodic stem and the prosodic word has no consequences for the analysis of Limbu since all the phonological patterns relevant for motivating prosodic domains referencing morphological structure target larger portions of the grammatical word and not exclusively the stem. It might be objected that the prosodic stem might show language-particular mappings to morphological structures extending the stem, just like the prosodic word might include or exclude prefixes across language-particular grammars. From that point of view, one probably could state mapping rules for Limbu which incorporate the stem–suffix–enclitic sequence into a prosodic stem domain. As we saw above with the Clitic Group approach, the crucial problem in this proposal is the fact that the lexical and post-lexical level are mixed in non-trivial ways. Already at the level of the prosodic stem, post-lexical clitics are integrated and the derivation would need to step back to include lexical prefixes at a later stage. Note that this problem seems not to arise with the Bantu languages for which the prosodic stem domain has been motivated. In languages like Cilungu (Bickmore 2007), there is evidence for a principled prosodic distinction between the stem and the macrostem integrating lexical affixes at different levels. Clitics, however, are prosodized in a higher prosodic domain, presumably the Clitic Group.33

Finally, one might opt for a framework in which the Prosodic Hierarchy is replaced by morpheme-bound prosodic subcategorization frames (Inkelas 1989). The varying status of the prefix across the different word domains in Limbu would be an equal challenge for analyses in such terms. One and the same Limbu prefix is phonologically both cohering and non-cohering; the choice is not lexical but depends on the rule one looks at. The prefix would therefore have to come with two simultaneous subcategorization frames, specified for the relevant phonological processes: one frame selecting prosodic words would have to be specified for the [l]~[ɾ] alternation, while another frame selecting some smaller prosodic constituents (e.g. bimoraic feet) would have to be specified for stress assignment. Although we are sympathetic to construction-based co-phonologies (Itoh & Mester 1995, Orgun 1996, Antilla

[33] One JL referee reports evidence from Kuria that locative enclitics outside the grammatical word count as part of the word-internal prosodic domain, see also Odden (1987). S/he objects that there thus must be ‘some mechanism’ available for allowing a domain to be defined that is significantly mismatched from the grammatical word. However, until this very mechanism has been formalized, we still consider such cases a crucial challenge to the available theory.
2002, Booij 2005, Inkelas & Zoll 2005), this problem also extends to this more recent theory of the morphology–phonology interface. This model assigns a specific co-phonology to individual morphological constructions, such as, for example, an affixation construction. In the Limbu case, the prefixing construction /ku/-[laip], would have to be associated with one co-phonology which parses the prefix as part of the prosodic word and at the same time a second one which parses it as outside the prosodic word.

3.3.3 Summary

Having presented an in-depth analysis of the phonological and morphological properties of words in Limbu and situating the evidence in the broader picture of recent achievements in the theory of prosodic structure, we are now able to precisely formulate the language-particular challenges to the theory of the Prosodic Hierarchy.

Accepting the standard analytic tools for motivating a prosodic domain forces us to distinguish two distinct levels of prosodic word structure for Limbu. The major word provides a domain for stress assignment and the segmental process of coronal to labial assimilation. The prosodic domain references a stem, its prefixes, suffixes, and enclitics. The minor word is the locus of the [l]~[r] alternation and initial glottal stop insertion. The morphological string that this domain targets consists of the stem, its suffixes and enclitics, to the exclusion of the prefix. Furthermore, phonotactic restrictions and resyllabification evidence the syllable domain, secondary stress assignment within the word motivates the bimoraic foot, and a rule of voicing assimilation across words suggests the relevance of the phonological phrase.

If we attempt to represent these findings in a structure that only employs the levels of the Prosodic Hierarchy, we face the problem of undergeneration. In such a prosodic tree, the fact that the domain of some word-related processes includes the prefix and the domain of some other processes excludes the prefix cannot be captured. To avoid this inadequacy, we opt for the postulation of two distinct prosodic word domains, the major and the minor word. In the context of the Strict Layer Hypothesis, this solution forces us to accept violations of Clustering and Strict Succession.

In this context, we point to Itoh & Mester’s 2009 version of the prosodic hierarchy, which allows prosodic adjunction to multiply levels of prosodic structure. In their framework, a minimal word can in various steps be enlarged by prosodic adjunction to constitute a larger word domain, up to the maximal word domain. In face of the Limbu data, this move seems to be justified, since it would allow us to refer to the minor word for the sake of some processes and to the major word for the sake of others. However, this approach has itself serious drawbacks, the most severe being that it is too powerful and imposes too little restrictions on prosodic systems. This objection has to be emphasized given our claim in Section 2 that the standard
version of the Prosodic Hierarchy already overgenerates.\footnote{One JL referee notes that Itô & Mester’s (2009) proposal also violates the ‘A-over-A principle’ (Chomsky 1964, Ross 1967, cited in Odden to appear).} Ultimately, revisions along these lines again prune the predictive power of the theory.

The case study from Limbu suggests that there is no reason to assume that phonological processes should universally cluster on one prosodic word domain intermediate between the foot and the phonological phrase. It also demonstrates that deviations from the theoretical predictions of the Prosodic Hierarchy may lie well outside the range of phenomena which are accountable by the additional prosodization mechanisms proposed in the literature.

4. THE PROSODIC WORD AS AN EMERGENT CATEGORY

Up to this point, our discussion has been pessimistic towards the standard descriptive framework for the study of word-related prosodic patterns and the architecture of prosodic structure as enshrined in the Prosodic Hierarchy. In cross-linguistic comparison, this approach forces us, on the one hand, to make languages fit the theory by systematically excluding variation, for instance, in the form of non-general processes referencing morphological structure; on the other hand, the study of cross-linguistic variation forces us to revise the theoretical apparatus to such an extent that it loses its predictive power.

As a reaction to this interim conclusion, we would like to advocate a different view on prosodic domains. We will argue that constellations such as those evidenced in Vietnamese and Limbu can only be handled if we accept that prosodic domains are language-particular, intrinsic and highly specific properties of individual phonological rules or constraints. Our heuristics for the cross-linguistic study of prosodic domains allows us to inductively investigate how much prosodic systems can vary and what the limits of this variation are. Instead of positing absolute universals a priori, we seek statistical universals which help us understand how a probable prosodic system might look like (see Newmeyer 2005, Bickel 2007).

4.1 Decomposing the ‘prosodic word’

The evidence from our two cases suggests that the prosodic word should not be conceptualized as a holistic entity with a fixed set of properties. Even within one language, the prosodic word may be defined by different sets of processes or rules, cf. the minor and major prosodic word in Limbu with their distinct properties. The desire to distinguish ‘different kinds’ of the
prosodic word has already been expressed by Hyman (2008: 335–336) who proposes the following list:

(34) *Different kinds of PW*

(a) The demarcative word: a property marks the beginning or end of the word
(b) The culminative word: a feature occurs only once per word
(c) The harmonic word: a feature is realized throughout the word
(d) The metrical word: a word consists of hierarchically arrayed moras or syllables
(e) The minimal word: a word must consist of a minimum of moras or syllables
(f) The maximal word: a word can consist of a maximum of moras or syllables
(g) The phonotactic word: a word permits only certain output segments/sequences
(h) The morphophonotactic word: a word permits only certain input segments/sequences

Facing the fact that even such finer-grained word notions differ across languages, we would go one step further in positing only language-particular prosodic words. Consider the following prosodic domains for some word-related processes in Vietnamese and Limbu.

(35) (a) The Vietnamese Tonal Dissimilation Word\((\text{BASE } cúng–\text{RED } cúng)\)
    Replace the high tone of the base with a low tone in the reduplicant.
(b) The Vietnamese Tonal Harmony Word\((\text{BASE–RED})\)
    Be faithful to the tone series of the base in alliterative reduplication; note that \(ngã\) and \(họ\) change their series membership.
(c) The Vietnamese Sino-Vietnamese Word\((\text{STEM–STEM})\)
    Stress the left constituent of the compound.
(d) The Limbu Velar Assimilation Word\((\text{PREFIX–STEM})\)
    \(/n/\) assimilates to the following velar place feature and surfaces as \([n]\)
(e) The Limbu Glottal Stop Insertion Word\((\text{PREFIX–STEM–SUFFIX–ENCLITIC})\)
    Insert an initial glottal stop to avoid a vowel-initial syllable.
(f) The Limbu Primary Stress Word\((\text{PREFIX–STEM–SUFFIX–ENCLITIC})\)
    Stress the stem-initial syllable; if an iamb would arise in prefix–stem combinations, stress the initial syllable.

For each language-particular domain, we specify the morphological structure relevant for a given process and also the details of its application. As a matter of fact, the domain specifications vary with respect to their scope. In

[35] This is in fact how we enter the data in our typological database, cf. fn. 1 above.
specific instances of reduplication in Vietnamese are singled out, whereas (35b) refers to an entire subtype of reduplication. In some cases, word-related processes are only active in a certain area of the lexicon, as in (35c), an information also included in the domain specification. Note that the three domains for Vietnamese do not count as evidence for a domain in the Prosodic Hierarchy, since they do not apply generally within the lexicon. In contrast, we consider generality across the lexicon a matter of degree and leave the empirical test of whether this property has an effect on the organization of prosodic structure for future research. In any case, the Limbu domains (35d–f) differ from the Vietnamese ones in that their domain specifications generalize over morphological constructions, such as e.g. prefixation or the entire grammatical word.36

Equipped with lists like (35), we are now in the position to explore empirically, e.g. using standard statistical tools of data-mining such as multidimensional scaling, what if any patterns there are in the distribution of prosodic domains across languages. In Bickel, Hildebrandt & Schiering 2009, we test whether there are probabilistic clusters of similarly-sized domains relative to the kind of sound pattern involved, e.g. tonal vs. segmental processes. In sharp contrast to the predictions of the Prosodic Hierarchy, we find a cross-linguistic trend for languages to multiply prosodic word domains. However, we also find statistical support for the following probabilistic universal:

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

In comparison to the huge body of absolute universals stated in the Prosodic Hierarchy, this universal seems to have rather small scope and to be rather local. However, it is empirically founded and, at least in our sample of 70 languages, stands the test for genealogical and areal bias.

4.2 The emergent prosodic word

The findings just reported suggest that like in many other areas of grammar, the absence of absolute universals does not mean that the distribution of structure is random. This is also evident from the observation that certain domain specifications are referenced more frequently by phonological

[36] Here we see strong parallels with theories of construction-based co-phonologies, the main difference being that we start from the phonological patterns and from there look at their domain specifications instead of starting with the morphological construction to assign a co-phonology. For our purposes, this procedure seems to be more adequate, since we can easily describe the fact that Limbu kulazp exhibits two prosodic structures at the same time: The Limbu Primary Stress Word would treat this as one domain, whereas the The Limbu [l]~[r] Alternation Word excludes the prefix. The construction-based co-phonology approach would need to assign two co-phonologies to one construction type.
processes than others. Consider the frequency of reference to some of the prosodic domains described for Limbu (Hildebrandt 2007).

(37) **Prosodic words in Limbu**

The Limbu Primary Stress Word

The Limbu Labial Assimilation Word

The Limbu [l] ~ [r] Alternation Word

The Limbu Glottal Stop Insertion Word

The Limbu *_ao(ŋ) Word

The Limbu Velar Assimilation Word

The Limbu Velar Assimilation Word

The Limbu Glottal Stop Insertion Word

We note that some domains, such as the prefix domain, or the smaller prefix–stem and stem–suffix domains, are referenced only by one process each. In contrast, several processes converge on the prefix–stem–suffix–enclitic domain and even more cluster on the stem–suffix–enclitic domain. It is not a coincidence that these domains figure most prominently in our discussion of the Prosodic Hierarchy in Limbu. In this line of thought, one could attribute the concept of the prosodic word simply to the domain that is most frequently referenced by phonological patterns in a given language. The prosodic word can thus be redefined as a language-particular category which emerges through frequent reference within the prosodic system. In Vietnamese, on the other hand, no domain seems to be frequent enough to postulate a prosodic word domain in the traditional sense.

The concept of the emergent prosodic word again offers rich possibilities for cross-linguistic comparison. We observe that languages tend to show bipolar distributions in the frequency of reference to prosodic domains. From that point of view Limbu, in which two domains are more frequent than the other attested ones, mirrors a common property of prosodic systems. But we can also ask whether the evidence for the prosodic word is equally strong across languages. We already used the related measure of ‘word dominance’, i.e. the frequency of reference to the most frequently referenced domain, to test whether languages formerly classified as ‘stress-timed’ indeed show more evidence for prosodic words than languages of other rhythm classes as hypothesized by Auer 1993 (Schiering 2009). It turns out that there is a trend for stress-timed languages in Indo-European and Sino-Tibetan to have more patterns reference their most frequent domains than syllable- or mora-timed languages. This finding, among others, suggests that prosodic word domains are stable in time, such that languages tend to retain the prosodic profile they inherited from the proto-language (see also Bickel et al. 2009). Presumably, frequent domains have a gravitating effect within the system, attracting phonological patterns which evolve in the course of sound change. However, at this point we can only speculate about the wider relevance and the psycholinguistic reality of the emergent prosodic
word, but we consider the concept interesting enough to lead future research in the area of prosodic structure.

5. CONCLUSION

In this paper, we discussed the prosodic structure of two languages which challenge the theoretical assumptions of the Prosodic Hierarchy. Vietnamese is a language in which a prosodic word domain cannot be motivated on the basis of purely phonological rules. The phonology of the language thus exhibits less structure than would be predicted by Prosodic Phonology. For the architecture of prosodic structure this entails violations of Clustering and a consistent violation of Strict Succession. Limbu is a language in which multiple word domains can be motivated which overlap in non-trivial ways by referencing different portions of the grammatical word. The phonology of this language thus exhibits more structure than would be predicted by Prosodic Phonology. For the architecture of prosodic structure this entails violations of Clustering. The problems this evidence poses can neither be solved by the standard theory nor by more recent versions which allow more flexibility by weakening the Strict Layer Hypothesis or by providing more fine-grained inventories of prosodic domains or prosodic subcategorization.

The challenge from Vietnamese can be responded to by rejecting the idea that structures posited as universal must be descriptively necessary in any language. This can mean either of two things. (i) It can mean that Vietnamese speakers have prosodic words of the standard kind in their minds, but the words do not manifest themselves in any known phonological effect. (ii) We reinterpret Prosodic Hierarchy theory as constituting a universal inventory of choices rather than as an absolute universal; then, Vietnamese can simply be said to lack prosodic words.

The challenge from Limbu is stronger: whether one interprets the Prosodic Hierarchy as an absolute universal or as an inventory of choices, the theory leads to systematic violations of the Clustering Prediction, or, if one seeks to rescue these predictions, to self-contradictory analyses (‘the same prefix is and is not part of a prosodic word at the same time’). A radical alternative approach is typological in spirit: prosodic domains are conceived of as language-particular, intrinsic and highly specific properties of individual phonological rules or constraints, and all we need is tools for comparing domains across languages and for measuring their similarities and differences in precise ways. We sketched how we explored this in companion papers (Bickel et al. 2009, Schiering 2009). Note that under such a conception, prosodic domains are by no means superfluous or that the phonological properties of morphological constructions can be accounted for by a direct reference mechanism. The point is rather that a more reliable methodology should start from the individual processes at work in the language and
should construct prosodic structure on that basis rather than starting with a
limited number of domain types defined a priori. 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 morphologi-
cal construction type. We hope that this reassessment inspires future research
on prosodic structures.

REFERENCES

Nguyen, Thi Anh Thu & John Ingram. 2007. Word stress and compounding in Vietnamese. Ms., School of English, Media Studies, and Art History, University of Queensland, St. Lucia, Australia.


Authors' addresses:
(Schiering)
Westfälische Wilhelms-Universität Münster, Institut für Allgemeine Sprachwissenschaft, Aegidiistr. 5, 48143 Münster, Germany rene@punksinscience.org

(Bickel)
Institut für Linguistik, Universität Leipzig, Beethovenstrasse 15, 04107 Leipzig, Germany bickel@uni-leipzig.de

(Hildebrandt)
English Language and Literature, Box 1431, Southern Illinois University Edwardsville, Edwardsville, IL 62025, USA khildeb@siue.edu