

## Cross-linguistic Challenges for the Prosodic Hierarchy: Evidence from Word Domains<sup>1</sup>

Kristine Hildebrandt<sup>#</sup>, René Schiering<sup>\*</sup>, Balthasar Bickel<sup>\*</sup>

<sup>\*</sup>University of Leipzig and <sup>#</sup>University of Manchester

### 1. Introduction

#### (1) The Prosodic Hierarchy

PROSODIC DOMAIN	PROPERTIES	MAPPING
U	e.g. English flapping	} syntactic information
I	intonation contours	
Φ	postlexical processes	} morphosyntactic information
C	clitic-specific processes	
ω	minimality, phonotactics, processes	} morphological information
Σ	stress, rhythm, segmental rules	} phonological information
σ	} stress, sonority	
μ		

- (2) Predictions made by the standard theory (see also Selkirk 1984, Itô & Mester 1992)
1. Clustering: phonological domains cluster on a single universal set of domains (discussed by Inkelas & Zec 1995).
  2. Strict Succession: all languages have at least the domains listed in (1) (= Nespors & Vogel's (1986: 7) Principle 1).
  3. Proper Bracketing: no language has non-stacking domains (=Nespors & Vogel's (1986: 7) Principle 2).
- (3) “While there is no *a priori* reason that the phonology of a given must include all [...] units, we will make the assumption here that this is the case, an assumption that can be motivated on both general and theory-specific grounds.” (Nespors & Vogel 1986: 11)

---

<sup>1</sup> We would like to acknowledge the assistance of Joshua Wilbur and Franziska Crell in collecting and evaluating the data presented here. We are also grateful to Michael Dunn who shared his expertise on Chukchi with us.

- (4) “While the predictions are clear, in practice the evidence is less so. [...] More data is clearly needed before any conclusion can be drawn as to the verity of the domain clustering and strict layering predictions.” (Inkelas & Zec 1995: 548)
- In this talk, we report on findings of a typological database project on word domains ([www.uni-leipzig.de/~autotyp/projects/wd\\_dom/wd\\_dom.html](http://www.uni-leipzig.de/~autotyp/projects/wd_dom/wd_dom.html)). The cross-linguistic evidence collected so far challenges the predictions of the prosodic hierarchy in crucial aspects:
    - i) In languages like Vietnamese, a phonological domain  $\omega$  cannot be motivated. This violates Strict Succession which prohibits the skipping of levels.
    - ii) In languages like Chukchi, multiple nonisomorphic word domains can be motivated. This violates Proper Bracketing which prohibits non-stacking domains.

## **2. Word Domains in the Prosodic Hierarchy**

- (5) Deviations from the clustering hypothesis allowed by the theory
1. According to a finite list exceptions defined by Nespor & Vogel (1986: 109-144), e.g. [(prefix) $\omega$ (stem-suffix) $\omega$ ].
  2. Recursive word domains defined by the same process, e.g. Peperkamp’s (1996: 113) analysis of recursive stress domains in Neapolitan Italian, cf. [ $\omega$ [ $\omega$ [ $\Sigma$ cónta]] [ $\Sigma$ ténna]], [ $\omega$ [ $\omega$ [ $\varphi$ cónta]] [ $\sigma$ tə]], [ $\omega$ [ $\varphi$ təna]] [ $\omega$ [ $\varphi$ cóntə]], [ $\omega$ [ $\sigma$ təna]] [ $\omega$ [ $\varphi$ cóntə]].
  3. Non-isomorphic domains defined on different phonological tiers, e.g. Hyman, Katamba & Walusimbi’s (1987) Clitic Groups in Luganda, one on the tonal and one on the quantity tier.
- (6) Deviations from the clustering hypothesis evidenced in our sample
1. Phonological domains cluster less often on a single set of domains than would be expected. Only 7 of the 50 languages in our present sample exhibit such a clustering at the word level.
  2. The lack of phonological domains at the word level is documented for 3 languages in our present sample.
  3. Non-isomorphic domains which do not cohere to proper bracketing form a substantial group of languages showing multiple word domains, i.e. 13 out of 40.

### 3. Challenge 1: Vietnamese, or the lack of phonological word domains

- Despite the assumed universality of phonological words (Dixon & Aikhenvald 2002: 32), motivating such a domain on phonological grounds turns out to be non-trivial in a number of languages (see e.g. Evans, Fletcher & Ross (to appear) on the polysynthetic language Dalabon). In what follows we will discuss the case of Vietnamese which is said to lack phonological word domains (Thomas 1962).<sup>2</sup>

#### 3.1. Morphological Domains in Vietnamese

- Although Vietnamese is characterized by the lack of inflectional morphology, a number of morphological processes in word formation have been distinguished.

(7) Monomorphemic words

- a. *đi* ‘go’      b. *va-li* ‘suitcase’      c. *com-mi-nít* ‘communist’

(8) Prefixation, or better Eidemic Resonance (Bickel & Nichols 2006)

- |                       |                   |                             |                       |
|-----------------------|-------------------|-----------------------------|-----------------------|
| <i>đâu</i> ‘anywhere’ | <i>nào</i> ‘any’  | <i>bao</i> ‘to any extent’  | <i>sao</i> ‘however’  |
| <i>đây</i> ‘here’     | <i>này</i> ‘this’ | <i>bây</i> ‘to this extent’ | <i>vậy</i> ‘this way’ |
| <i>đấy</i> ‘there’    | <i>ấy</i> ‘that’  | <i>bấy</i> ‘to that extent’ | <i>vấy</i> ‘that way’ |

(9) Reduplication, partial and total

- a. *lết* ‘drag about painfully’      b. *la-lết* ‘do with much pain or difficulty’  
 c. *rõ* ‘be clear’      d. *rõ-rệt* ‘be very clear’  
 e. *nói* ‘talk’      f. *nói nói* ‘keep talking and talking’

(10) Compounding

- a. *bà* ‘grandmother’ + *con* ‘child’ → *bà con* ‘be related’  
 b. *ngôn* ‘word’ + *ngữ* ‘language’ + *học* ‘study’ → *ngôn-ngữ học* ‘linguistics’  
 c. *vô* ‘without’ + *tuyến* ‘wire’ → *vô-tuyến* ‘wireless’  
*điện* ‘telephone’ + *thoại* ‘speech’ → *điện-thoại* ‘telephone’  
*vô-tuyến* ‘wireless’ + *điện-thoại* ‘telephone’ → *vô-tuyến điện-thoại* ‘radio telephone’

(11) Discontinuous words

- a. *cà phê* ‘coffee’ →      b. *cà với phê* ‘coffee and the like’ (với = ‘with, and’)  
 c. *mặt-mày* ‘face, honor’      d. *dạn-dày* ‘weather-beaten, shameless’  
 → e. *mặt dạn mày dày* ‘utterly shameless person’

<sup>2</sup> The data in this section was drawn from reference grammars and detail studies of Vietnamese, i.e. Thompson (1963, 1965), Liêm (1970a, b), Nhan (1984), Đình-Hoà (1997), Pham (2003).

- Overall, there is very little uncontroversial evidence for morphologically complex words in Vietnamese, with reduplications constituting the most subtle evidence of grammatical words.

### 3.2. Prosodic Domains in Vietnamese

- Traditionally, the phonological analysis of Vietnamese centers on the syllable. (12) summarizes the syllable structure which integrates the distribution of onset consonants (initials), vocalic nuclei and coda consonants (rhymes).

(12) The syllable in Vietnamese

Tone	
Initial	Rhyme
C	(w)V(V)(C)

- Whereas the onset position can be filled with any consonant except /p/, only eight consonants may appear as codas: /p, t, k, m, n, ŋ, y, w/.
- The syllable may also be considered the tone-bearing unit. All six tones may appear on the syllable shell schematized in (12). However, the tones *sac1* (high/contour) and *nang1* (low/contour) have short allotones, *sac2* and *nang2*, respectively, whose distribution is restricted to stop coda syllables.
- Above the syllable level a *pause group* consists of a sequence of syllables occurring between two pauses and is characterized by at least one heavy stress <^>, whose position is usually final, but is ultimately determined by semantic prominence, and one intonation contour <.>.

(13) *°Tôi °không 'biết.*

I not know  
'I don't know.'

- When two or more such pause groups are combined, the non-final pause groups are realized with decreasing intonation <.>, signaling that the speaker intends to continue.

- (14) *Tôi 'đến nhà, má tôi mở cửa 'ra, tôi 'vô.*  
 I arrive house, mother I open door exit, I enter.  
 'I arrived at the house, my mother opened the door, and I went in.'

- Apart from heavy stress in pause groups as exemplified in (13) and (14), there is arguably another level of stress, in which alternating syllables show a higher degree of stress.

- (15) *Lạy 'Trời mưa "xuống,* 'I pray for the rain to fall'

- What is crucial is that polysyllabic forms, irrespective of their morphological complexity, show the same stress pattern as do phrases, cf. (16a-d).

- (16) a. *°va-'li* 'suitcase'  
 b. *°nói 'nói* 'keep talking and talking'<sup>3</sup>  
 c. *°bà 'con* 'be related'  
 d. *°Tôi °không 'biết.* 'I don't know'

### 3.3. Summary

- The Vietnamese data discussed in the previous sections can be summarized in a prosodic hierarchy which maps prosodic, morphological and syntactic domains.

#### (17) A Prosodic Hierarchy for Vietnamese

PROSODIC DOMAIN	PROPERTIES	MAPPING
U	terminal intonations	sentences
I	pauses, heavy stress, intonation	clauses
Φ	final stress	{ phrases, compounds, reduplications, words
Σ	iambic rhythm	
σ	phonotactics, tone	

- There is no evidence for a phonological word domain. Note that the prosodic hierarchy violates the principle of Strict Succession in that (at least) the level of the

<sup>3</sup> Reduplications are also characterized by a rule of tone harmony, in which the tone of the reduplicant must agree with the tone of the base in register, cf. *trang (sac1, high)* 'white' > *trang trang (ngang, high - sac1, high)* (Pham 2003: 28).

word is skipped. Two languages in our present sample are like Vietnamese: Lahu and Car.

#### 4. Challenge 2: Chukchi, or the multiple improper phonological word domains

- Apart from languages resisting the motivation of a phonological word domain, a number of languages allow the motivation of multiple word domains. In what follows, we will discuss the case of Chukchi, a language where word domains can be motivated which exhibit improper bracketing.<sup>4</sup>

##### 4.1. Morphological Domains in Chukchi

- Chukchi is a morphologically extremely rich language, making use of several morphological processes in derivation and inflection, including prefixation, suffixation, circumfixation and incorporation.

(18) Morpheme order of Chukchi derivational affixes (Dunn 2001: 8).

<i>teŋ-</i>	<i>lyi-</i>	<i>re-</i>	<i>ine-</i>	<i>n-</i>	---	<i>-et</i> <i>-ew</i>	<i>-tku</i>	<i>-ŋ</i>	<i>-tku</i>	<i>-r<sup>2</sup>u</i> <i>-jw</i>	<i>-l<sup>2</sup>et</i>	<i>-m yo</i> <i>-ŋŋo</i> <i>-plətku</i>
INTS	INTS	DESID	AP	CS	root	CS/ VDER	AP	DESID	ITER	COLL	DUR	INCH, COMPL

(19) [cough]      *tə-lge-ta yjaŋ-əŋŋo-ɣ<sup>2</sup>a-k*  
 1sg-INTS-desire-INCH-TH-1sg  
 ‘... oh, I started to want (to cough)’

##### 4.2. Prosodic Domains in Chukchi

- The phonology of Chukchi is very rich in prosodically delimited rules. The smallest prosodic domain is arguably the syllable, which maximally takes CVC structure and which functions as the skeleton for syllabification and processes like epenthesis.

<sup>4</sup> The data in this section was drawn from Dunn (1999, 2001).

(20) a. underlying form: C-V-C-C-C-V-C  
 | | | | | | |  
 ŋ e w c q e t 'woman'

b. syllabification:  
 σ            σ            σ  
 / \        / \        / \  
 C-V-C C-V-C C-V-C  
 | | | | | | | | |  
 ŋ e    w    c q e t

c. C-deletion  
 epenthesis  
 σ            σ            σ  
 / \        / \        / \  
 C-V C-V-C C-V-C  
 | | | | | | | | |  
 ŋ e    w ə c q e t

- Secondary and primary stress assignment provide evidence for the prosodic domains of the (trochaic) foot and the (one main stress) word, respectively. In a phonological word, primary stress is assigned to the first syllable with a consonant onset and a full (non-schwa) vowel and secondary stress is placed on every second syllable before and after that. The domain for stress assignment references stems, prefixes, circumfixes, suffixes, and enclitics.

(21) a. /<sup>1</sup>nù.tec.<sub>1</sub>qə.cə.<sub>1</sub>ku.kin/ 'smth. from the surface of the ground'  
 b. /qə.<sup>1</sup>jet.ɣ<sup>2</sup>i/ 'come'  
 c. /<sub>1</sub>kər.ɣə.<sup>1</sup>re.c<sup>2</sup>ə.<sub>1</sub>kin./ 'smth. made of dry stumps'  
 d. /a.<sup>1</sup>tok.tor.<sub>1</sub>ka/ 'without a doctor'  
 e. /a.<sup>1</sup>mo.le.<sub>1</sub>qaj/ 'bark (DIM)'

- For the sake of Vowel Harmony the Chukchi vowels group in two series: /i ~ e ~ u/ and /e ~ a ~ o/. Morphemes that trigger the vowel harmony prosody make all vowels in a word, i.e. stems, prefixes, circumfixes and suffixes but not enclitics, to agree in openness.

(22) a. /təle-n/        [tə.lan]  
 place-NMLZR  
 'path' (stem + derivational suffix)

b. /luŋ-t<sup>1</sup>l-ŋŋo-tel/        [lŋ.t<sup>2</sup>ə.lə.ŋŋo.ta]  
 NEG-ill-INCHO-NEG  
 'They didn't fall ill.' (stem + circumfixes + suffix)

c. /ɣe-lyi-nm-tku-jw-linat/ [ɣal.ɣen.mət.koj.wə.le.nat]  
 PF-INTS-kill-ITER-INTS-3PL  
 ‘They killed them all.’ (prefixes + stem + suffixes)

d. ʔemi=tlon?  
 where=INTER  
 ‘Where?’ (stem + enclitic)

- Since V-initial syllables are generally dispreferred within words, hiatus which results from the concatenation of morphemes is repaired by the deletion of V1, a rule which applies to the morphological structures stem + suffix, prefix + stem, and circumfix + stem (see 23), or by a certain affix allomorphy in stem + circumfix/suffix combinations (see 24). Although enclitics are subject to the same constraint, hiatus is resolved there by the deletion of the V2 (see 25).

(23) a. /qora-ine-t/ [qorenat]  
 reindeer-POSS-3plABS  
 ‘(calves) belonging to the reindeer’ (stem + suffix)

b. /ɣe-iw-lin/ [ɣiwlin]  
 PF-say-3SG  
 ‘(he) said’ (prefix + stem)

c. /ɣe-uŋel-(t)e/ [ɣuŋele]  
 CONV-collect.firewood-CONV  
 ‘collecting firewood ...’ (circumfix + stem)

(24) a. /ɣe-pplu-ra-(t)e/ [ɣappəlorata]  
 COMMIT-little-house-COMIT  
 ‘with little houses’ (stem + circumfix)

b. /ənqine-epə/ [ən.qe.naj.pə]  
 DEM-ABL  
 ‘from that...’ (stem + suffix)

c. /telenjep-kine-epə=<sup>2</sup>m/ [ta.lan.jap.ke.naj.p<sup>2</sup>əm]  
 long.ago-REL-ABL=EMPH  
 ‘from long ago...’ (suffix + suffix)

(25) a. ətlon meŋqorə?  
 INTER whence  
 ‘Where are they from?’

b. ʔemi=tlon?  
where=INTER  
'Where?'

- Another phonological rule which references morphological structures is the one of Glottalisation, which spans the stem, its suffixes and enclitics.

(26) a. /tinjero=<sup>ʔ</sup>m/ [tinjer<sup>ʔ</sup>om]  
Personal.Name=EMPH  
'Tinjero!' (stem + enclitic)

b. /tajɲat-ɣəpə=<sup>ʔ</sup>m/ [tajɲat-ɣəp<sup>ʔ</sup>əm]  
food-ABL=EMPH  
'from the food' (stem + suffix + enclitic)

- The right edge of the stem is target of a rule of final vowel reduction which reduces /e/ to [ə].

(27) a. /wopqa/ [wopqə] 'moose'  
b. /wala/ [walə] 'knife'  
c. /maka/ [makə] 'nappy'

- Most consonant alternations (assimilations & other alternations) do not appear to be sensitive to morphological structure or morpheme boundaries, or are entirely morpheme-internal. However, there is one type of assimilation (a stop ~ approximant manner assimilation) with sensitivity to a particular type of morphological environment:

(28) a. /k/ → [ɣ]/ \_C (where domain is medial of stem + case suffix + enclitic postposition)  
b. /ətl<sup>ʔ</sup>a-k=reen/ [ətl<sup>ʔ</sup>ayre:n]  
mother-LOC=together  
'together with mother'

- Although the phonotactic structure of the language strongly disprefers vowel-initial syllables, vowel-initial words/stems are allowed.

(29) a. /a.tok.tor.ka/ 'without a doctor'      b. /a.mo.ʔe.qaj/ 'bark (DIM)'  
c. /ekək/ 'son'                                      d. /iw/ 'say'

(30) a. əŋje            ik-we            “cam<sup>?</sup>am”  
 NEG.HORT    say-NEG        unable.MOD  
 ‘Don’t say ‘I can’t.’

b. <sup>?</sup>etki qejwe əŋjw-zej-ə-ne  
 badly truly    uncle-DIM-E-AN.ERG  
 ‘Uncle (will) trully really (beat me).’

- Additional evidence for prosodic domains comes from intonation patterns with a rise-fall contour and a final pitch drop. This domain clearly references the main units of syntactic analysis, i.e. ‘sentences’, which may motivate the prosodic domain of the Intonation Phrase.

### 4.3. Summary

- The Chukchi data discussed in the previous sections can be summarized in a prosodic hierarchy which maps prosodic, morphological and syntactic domains.

#### (31) A Prosodic Hierarchy for Chukchi

PROSODIC DOMAIN	PROPERTIES	MAPPING
I	rise-fall contour	sentences
ω1	Stress	prefix/circumfix-stem-suffix/circumfix=encl
		_____
ω2	Vowel Harmony V1 Deletion	_____
		_____
ω3	V2 Deletion	_____
ω4	Glottalization C assimilation	_____
		_____
ω5	C epenthesis	_____
ω6	Vowel Reduction V-initial syllable	_____
		_____
Σ	trochaic rhythm	
σ	syllabification, epenthesis	

- Between the levels of the foot and the intonation phrase, Chukchi shows evidence for multiple word domains. Unlike the cases of multiple word domains discussed in the literature, the phonological word domains in Chukchi overlap since they all include the stem.
- Since some domains ( $\omega_2$  vs.  $\omega_3$ -  $\omega_5$ ) violate proper bracketing they cannot be analyzed as recursive occurrences of the same domain. This is true for 12 other languages in our present sample. See also the Belhare data on Intersonorant Voicing and Nasal Assimilation in (32) and (33), respectively.

(32) ka-teĩ-ʔ-ni-kak → ka( $\omega$ teĩʔniga), \*( $\omega$ kareĩʔniga)  
 1sP-hit-NPST-NEG-2  
 ‘You won’t hit me’ (stem + suffix + enclitic)

(33) N-kai-chi-N-ka → ( $\omega$ ŋkai)chinga, \*( $\omega$ ŋkaichinga)  
 NEG-come.up-d-NEG-2  
 ‘You won’t come up’ (prefix + stem)

## 5. Conclusions

- Word domains cannot be postulated as universal, but we can compare (typologize) languages by considering a) word domains in particular languages and b) similarities among these language-particular word domains (cf. Dryer 1997). Similarities can be measured by analyzing repeated clustering of domains across processes and languages (in progress).
- Prosodic domains should not be considered as something that exists independently of given rules or constraints. Instead, prosodic domains are intrinsic and highly specified properties of individual rules or constraints. It is an empirical task to find out whether rules or constraints share domains in a language or across languages.
- In principle, prosodic domains may be organized in various ways, hierarchical structure being only one of a number of options. Even if prosodic domains show hierarchical organization it is far from evident that they should converge on a universal list like the one enshrined in the prosodic hierarchy.

## 6. References

- Bickel, Balthasar & Johanna Nichols 2006. "Inflectional Morphology." In: Timothy Shopen (ed.). *Language Typology and Syntactic Description. Second Edition*. Cambridge: Cambridge University Press.
- Đinh-Hoà, Nguyễn 1997. *Vietnamese. Tiếng viet khong son phan*. (London Oriental and African Language Library, 9). Amsterdam/Philadelphia: John Benjamins.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald 2002. "Word: a typological framework." In: Dixon & Aikhenvald (eds.). *Word. A Cross-linguistic Typology*. Cambridge: Cambridge University Press: 1-41.
- Dryer, Matthew S. 1997. "Are Grammatical Relations Universal?" In: Bybee, Joan; John Haiman & Sandra A. Thompson (eds.). *Essays on Language Function and Language Type. Dedicated to T. Givón*. Amsterdam/Philadelphia: John Benjamins: 115-143.
- Dunn, Michael 1999. *A Grammar of Chukchi*. Ph.D.-Dissertation, Australian National University.
- Dunn, Michael 2001. *The Word in Chukchi*. Unpublished Manuscript. Nijmegen: Max-Planck-Institut für Psycholinguistik.
- Evans, Nicholas, Janet Fletcher & Belinda Ross (to appear). "Big words, small phrases: mismatches between pause units and the polysynthetic word in Dalabon." *Linguistics*.
- Hyman, Larry M. Francis Katamba & Livingstone Walusimbi 1987. "Luganda and the strict layer hypothesis." *Phonology Yearbook* 4: 87-108.
- Inkelas, Sharon & Draga Zec 1995. "Syntax-phonology Interface." In: Goldsmith, John A. (ed.). *Handbook of Phonological Theory*. (Blackwell Handbooks in Linguistics, 1). Oxford: Blackwell: 535-549.
- Itô, Junko & R. Armin Mester 1992. *Weak Layering and Word Binariness*. Santa Cruz: Linguistic Research Center, Cowell College.
- Liên, Nguyễn Đăng 1970a. *A Contrastive Phonological Analysis of English and Vietnamese. A Contrastive Analysis of English and Vietnamese. Vol. 4*. (Pacific Linguistics; Series C - Books, 8). Canberra: Australian National University.
- Liên, Nguyễn Đăng 1970b. *Four-Syllable Idiomatic Expressions in Vietnamese*. (Pacific Linguistics; Series D - Special Publications, 6). Canberra: Australian National University.
- Nespor, Marina & Irene Vogel 1986. *Prosodic Phonology*. (Studies in Generative Grammar, 28). Dordrecht: Foris.
- Nhan, Ngo Thanh 1984. *The Syllabeme and Patterns of Word Formation in Vietnamese*. Ph.D.-Dissertation, New York University.
- Selkirk Elisabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. (Current Studies in Linguistics, 10). Cambridge/London: MIT Press.
- Thomas, David D. 1962. "On defining the 'word' in Vietnamese." *Van-Hoa Nguyet-San* 11: 519-523.
- Thompson, Laurence C. 1963. "The Problem of the Word in Vietnamese." *Word* 19: 39-52.
- Thompson, Laurence C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.
- Peperkamp, Sharon 1996. "On the prosodic representation of clitics." In Kleinhenz, Ursula (ed.). *Interfaces in Phonology*. Berlin: Mouton de Gruyter: 102-127.
- Pham, Andrea H. 2003. *Vietnamese Tone: A New Analysis*. (Outstanding Dissertations in Linguistics). New York & London: Routledge.

## Appendix

Language	Stock	Number of non-isomorphic, lexically-general p-domains	
		all	only structure-preserving
Lahu	Sino-Tibetan	0	0
Car	Austroasiatic	0	0
Vietnamese	Austroasiatic	0	0
Cayuvava	Cayuvava (isolate)	1	1
Kewa	Engan	1	1
Kayah Li (Red Karen)	Sino-Tibetan	1	1
Burmese	Sino-Tibetan	1	0
French (colloquial)	Indo-European	1	1
Cantonese	Sino-Tibetan	1	0
Qiang	Sino-Tibetan	1	1
Diegueño	Yuman	2	1
Newar (Dolakha)	Sino-Tibetan	2	2
Yidiny	Pama-Nyungan	2	2
Tibetan (Dege)	Sino-Tibetan	2	2
Martuthunira	Pama-Nyungan	2	2
Wu (Changzhou)	Sino-Tibetan	2	2
Irish	Indo-European	2	2
Kharia	Austroasiatic	2	2
Santali	Austroasiatic	2	1
Belhare	Sino-Tibetan	3	3
Diyari	Pama-Nyungan	3	3
Finnish	Uralic	3	3
Hixkaryana	Cariban	3	2
Nepali	Indo-European	3	3
Turkish	Turkic	3	2
Manange	Sino-Tibetan	3	1
Tibetan (Kyirong)	Sino-Tibetan	3	1
Jahai	Austroasiatic	3	1
Yimas	Lower Sepik	3	2
Mandarin	Sino-Tibetan	3	3
Garo	Sino-Tibetan	3	3
Luganda	Benue-Congo	4	4
Kinnauri	Sino-Tibetan	4	0
Egyptian Arabic	Semitic	4	3
Kham	Sino-Tibetan	4	3
Limbu	Sino-Tibetan	4	4
Greek (modern)	Indo-European	4	2
Chukchi	Chukchi-Kamchatkan	4	4
Sko	Macro-Skou	4	4
Kusunda	Kusunda (isolate)	4	4
Polish	Indo-European	4	1
Lithuanian	Indo-European	4	3
Tibetan (Lhasa)	Sino-Tibetan	5	3
Dutch	Indo-European	5	5
Persian	Indo-European	5	4

Spanish	Indo-European	5	2
Burushaski	Burushaski (isolate)	5	5
Meithei (Manipuri)	Sino-Tibetan	6	4
Swedish	Indo-European	6	3
German	Indo-European	6	3

3 out of 50 languages are like Vietnamese and lack any evidence for a word domain (unless one admits domains that are specified only for a single lexical affix, and are not referenced by general rules). 40 out of 50 languages are like Chukchi and have more than one word domain referenced by lexically-general phonological rules. This number goes down to 33 if we only admit phonological rules that are “lexical” in the sense that they are structure-preserving, i.e. do not introduce elements beyond what is already in the language's inventory. Thus, even on the narrowest definition of what could possibly count as evidence for word-domains, more than half of the languages surveyed here violate the Clustering Hypothesis.