

## Segmental Effects of Cliticization:

### Towards a Typology

René Schiering

University of Konstanz

Rene.Schiering@uni-konstanz.de

*Get rhythm!* (Johnny Cash)

**Research question:** The loss of phonological substance is generally considered an intrinsic concomitant of grammaticalization, cf. the concepts *attrition* (Lehmann [1982] 1995), *erosion* (Heine & Reh 1984) and the formulation of the *Parallel Reduction Hypothesis* (Bybee, Perkins & Pagliuca 1994). Is erosion a universal process in grammaticalization, or is it constrained by language(-type) specific properties in phonology?

**Claim:** Drawing data from a variety sample of cliticization phenomena, this talk will demonstrate the cross-linguistic diversity in the phonology of cliticization. Several phonological clines in prosodic and segmental phonology will be established and a rhythm-based typology will be proposed which predicts the absence, presence and degree of erosion to be found in grammaticalization.

**In a nutshell:** Evidence from cliticization casts doubt on the intrinsic status of erosion in grammaticalization.

## 1. Introduction

### 1.1. Clitics from a diachronic perspective

#### (1) Grammaticalization & associated processes

	content item	function word	clitic	affix	zero
phonetic processes	adaptation, erosion, fusion, and loss				
morphosyntactic processes	permutation, cliticization, affixation, and fossilization compounding				
functional processes	desemanticization, expansion, simplification, and merger				

(Hopper & Traugott 1993; Heine & Reh 1984; Heine, Claudi & Hünnemeyer 1991; Croft 2003: 253ff.; Lehmann [1982] 1995)

(2) Processes associated with cliticization

- *Phonological processes*: dependency on other (phonological) words  
Prosodic Phonology: deaccentuation, prosodic integration, tonal Sandhi, none  
Segmental Phonology: adaptation, reduction, none
- *Morphosyntactic processes*:  
Specialization for a phrasal position: initial, final, 2P, etc.  
Specialization for a morphological domain: pre-Head, post-Head position
- *Functional processes*:  
Scope narrowing: scope over entire phrases, scope over word domains

(3) “Erosion is a process by which the phonological substance of a morpheme is reduced, usually in accordance with its new evolutionary status. Thus, a bisyllabic word may be reduced to a monosyllabic morpheme once it has undergone affixation.” (Heine & Reh 1984: 21)

- With respect to phonology, weak words are said to undergo deaccentuation (prosodic) and phonetic erosion (segmental) when they grammaticalize towards affixes.

## 1.2. The data

- Using the sampling method proposed in Rijkhoff, Bakker, Hengeveld & Kahrel (1993) and Rijkhoff & Bakker (1998) a variety sample of 19 languages has been constructed. For each of Ruhlen’s (1991) language phyla one representative was chosen (see Appendix A).
- For each language, clitic elements have been identified on phonological and distributional grounds, following Zwicky’s (1994: 572) definition of clitics as ‘elements that are words from the syntactic point of view but are phonologically dependent [...] upon a neighboring word.’<sup>1</sup>

---

<sup>1</sup> Parameters for a typological study of clitics were taken from Zwicky 1977, 1985, 1994; Zwicky & Pullum 1983; Klavans 1985; Anderson 1992; Aikhenvald 2002; Haspelmath 2002.

- The Clitics Database, compiled with this sampling method, contains around 400 clitic elements including clitic particles, adpositions, pronouns, verbs, articles, as well as some miscellaneous cases exhibiting different degrees of cliticity encountered on the *clitic > affix* cline; from purely phonological cliticization (simple clitics) to syntactic cliticization (special clitics) and phrasal affixation (see Appendix B for a sample entry).

### **1.3. Outline of the talk**

- The objective of this talk is to flesh out the hypothesis that a rhythm-based typology can make predictions about possible phonological effects of grammaticalization and to establish reasonable scenarios for the evolution of morphology.
- Concentrating on cliticization, I intend to demonstrate that segmental effects of grammaticalization vary in different prosodic types and that erosion is characteristic for a certain language type.
- I will first present the various clines in prosodic phonology ranging from stress reduction and prosodic integration to tone neutralization and tonal Sandhi. In segmental phonology we not only find erosion but also structure preservation, assimilation and strengthening.
- Second, I will try to establish possible correlations between prosodic and segmental processes within the framework of a rhythm-based typology of language. In particular, segmental effects of stress and phonotactic restrictions constrain the possibility and degree of erosion.
- Finally, I will summarize the findings and discuss their significance for morphologization and morphological typology from a diachronic perspective. In the various phonological climates the evolution of bound morphology of different shapes can be expected in diachronic change.

## 2. The phonology of cliticization

### 2.1. Clitics in prosodic phonology

#### 2.1.1. Clitics in stress phonology

- When a word carrying primary stress is encliticized to a host word, its primary stress may be reduced to secondary stress, cf. *thepi* in the Tariana sentences (4) and (5).

(4) **thepí** di-rúku di-á Tariana  
 INTO.WATER 3sgnf-go.down 3sgnf-go  
 ‘Into water he went (contrary to all expectations).’

(5) du-hwá=**thepì** du-á du-aphuá=pidanà  
 3sgf-fall=INTO.WATER 3sgf-go 3sgf-dive=REM.P.REP  
 ‘She (the girl transformed into a snake-woman) fell into water diving.’  
 (Aikhenvald 2002: 64)

- Stress reduction may also lead to the complete loss of stress in the cliticized element. When the Georgian negative particle *ar(a)* in (6) procliticizes, it loses its stress completely and appears unstressed and unstressable (7).

(6) merab-i ekim-i **ár** áris Georgian  
 Merab-NOM doctor-NOM NEG he.be  
 ‘Merab is not a doctor.’

(7) puli **ara**=m-ákvs  
 money NEG=1sg.OBL-have  
 ‘I don’t have any money.’

(8) puli **ára**=m-akvs (Harris 2002: 236)  
 money NEG=1sg.OBL-have  
 ‘I don’t have any money.’

- As example (8) illustrates for Georgian, a clitic element may finally be integrated into the phonological word initiated by its host. In this case, host and proclitic form one unit for the sake of primary stress placement, and the proclitic regularly receives word initial primary stress.

- In the Standard Catalan example in (9), pronominal enclitics are extrametrical and remain unstressed and unstressable. In the Majorcan Catalan example (10), they are integrated into the word stress domain and regularly take final primary stress.

(9) donár=**nos**=**en** [dunár**nuzən**] Catalan  
 give=us=partit  
 ‘to give us some’ (Wheeler 1979: 34)

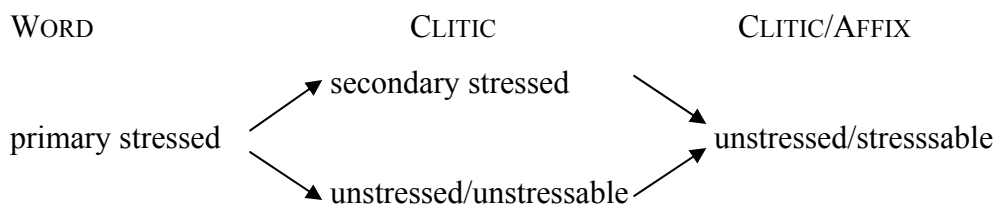
(10) Canta=**m**=**hó** Majorcan Catalan  
 sing=**me**=it  
 ‘Sing it to me!’ (Mascaró & Rigau 2002: 11)

- (11) illustrates the integration of clitic elements into the word domain for secondary stress in Finnish. Primary stress is realized on the initial syllable of the word, secondary stress on every odd-numbered syllable within the word, including clitics.

(11) Lúnta=**kò**=**han** Finnish  
 snow=INT=SOFTENER  
 ‘Some snow?’ (Nevis 1986: 28)

- Such cross-linguistic evidence is highly suggestive of a gradual process of cliticization combining the prosodic sub-processes of stress reduction (from primary stress to secondary stress to unstressed) and prosodic integration (from stressed to unstressed/unstressable to unstressed/stressable).

Figure 1: A prosodic cline for stress phonologies



### 2.1.2. Clitics in tonal phonology

- In combined stress/tone systems, stress reduction may be accompanied by the neutralization of tonal contrasts.

- There are two sets of pronouns in Nigerian Pidgin. Whereas the emphatic free pronouns are stressed and high toned, their enclitic counterparts are unstressable and low toned or as in (12) even toneless, i.e. undergoing tone spreading.

(12) A folo=**am**. [à 'fólò àm] Nigerian Pidgin  
 1sP followF=**3oP**  
 'I followed her/him (Faraclas 1989: 559)

- Unstressed syllables in Mandarin Chinese surface in so-called neutral tone, i.e. a pattern of tonal Sandhi according to which they will surface as 3 after 55, 3 after 35, 4 after 11, and 1 after 51. All enclitics, such as the clitic locative particle *shang* in (13), cf. *shàng*<sub>51</sub> 'to go / above', are unstressed and undergo this tonal Sandhi.

(13) tái<sub>35</sub> **shang**<sub>3</sub> qu le de rén Mandarin  
 carry **up** go PERF DE(a) person  
 'person carried up' (Lin 2001: 50)

- In pure tone languages, clitics come with a lexical specification for tone and are subject to regular rules of tonal Sandhi.
- In Yoruba, high toned pronominal clitics (14) are affected by certain tonal OCP effects, i.e. their high tone changes to a mid tone after high toned verbs or an epenthetic mid vowel is inserted between host and clitic (14c, d).

(14) Object clitic *yín* '2pl' in Yoruba (Akinlabi & Liberman 2000: 39f.)

- ó kọ **yín** 'he/she/it divorced you-all'
- ó pa **yín** 'he/she/it killed you-all'
- ó kọ **yín** 'he/she/it taught you-all'
- ó kọ ọ **yín** 'he/she/it taught you-all'

- In !Xóǎ clitic-like grammatical particles either carry inherent high or low tone and retain them in pre-head position. In some cases, such as the verb /i/ 'be' which has been grammaticalized into a possessive marker /V/, a mid-falling tone changes into a high tone.

(15) The clitic-like particles *ka* ‘to’ and *ka* ‘copulative’ in !Xóǝ (Traill 1985: 48)

- |  |  |
|--|--|
| a. (saa) <sub>H</sub> <b>ka<sub>H</sub></b> /qhuũ <sub>H</sub> | b. //hoa <sub>H</sub> <b>ka<sub>L</sub></b> /qhuũ <sub>H</sub> |
| ‘(go) to the white man’  | ‘not be a white man’   |

- In summary, clitics in tone languages show reduction of tonal contrasts in combined stress/tone systems, undergo tonal Sandhi in pure tone phonologies and may thus change their tone specification.

### 2.1.3. Clitics in intonation phonology

- In pure intonation phonologies, such as instanced in West Greenlandic, cliticization is associated with the loss of an intonation peak and the gradual integration into a neighboring intonation phrase, cf. (16) and (17).

(16) a.	—      —      — —      —      — —      —      —	Aata-mik ati-qar-puq <b>aat</b> Aata-instr. name-has-3s.indic. <b>req.conf.</b> ‘His name is Aata, isn’t it?’	West Greenlandic
---------	---	---	------------------

b.	— — —	isir-nia-rin= <b>aat</b> come.in-try-2s.imp.= <b>req.conf.</b> ‘Come in, won’t you?’	(Fortescue 1984: 9)
----	-------------	--	---------------------

(17) a.	—      —      — —      —      — —      —      —	aallar-pa= <b>guuq</b> leave-3s.inter.= <b>quote</b> ‘Did he <sub>i</sub> say that he <sub>j</sub> left?’	West Greenlandic
---------	---	---	------------------

b.	— — —	aallar-pa= <b>guuq</b> leave-3s.inter.= <b>quote</b> ‘Did he <sub>i</sub> say that he <sub>j</sub> left?’	(Fortescue 1984: 6)
----	-------------	---	---------------------



- Epenthesis is a process which strengthens the morpheme boundary by the insertion of additional segments. Such a process is exemplified by the rule of /j/-insertion in West Greenlandic (22).

(22) Clitic particle *aasiit* ‘again as usual’ in West Greenlandic (Fortescue 1984: 354)  
 urnippaa=**jaasiit**  
 ‘he came up to him as usual’

### 2.2.2. Syllabic phenomena

- As was the case for junctural phenomena, the syllabic composition of cliticized elements may not be effected at all. Compare again the free and clitic variants of *thepi* in Tariana (23) and (24).

(23) **thepi** di-rúku di-á Tariana  
 INTO.WATER 3sgnf-go.down 3sgnf-go  
 ‘Into water he went (contrary to all expectations).’

(24) du-hwá=**thepi** du-á du-aphuá=pidanà  
 3sgf-fall=INTO.WATER 3sgf-go 3sgf-dive=REM.P.REP  
 ‘She (the girl transformed into a snake-woman) fell into water diving.’  
(Aikhenvald 2002: 64)

- Assimilation at the syllabic level is attested in vowel harmony as exemplified by the Udihe examples in (25).

(25) Additive focus clitic =*dA* in Udihe (Nikolaeva & Tolskaya 2001: 87)  
 a. a’ka=**da** ‘and the back’  
 b. te’ge=**de** ‘and the gown’  
 c. mo’xo=**do** ‘and the cup’

- *Syllabic erosion* is a gradual process by which vowels in clitic syllables are reduced and finally deleted. In Standard Catalan, vowels in unstressed syllables undergo vowel reduction in which the seven phonemic vowels are reduced to three in unstressed syllables: /u, o, ə/ reduce to [u], /e, ε, a/ reduce to [ə], and /i/ reduces to [i].

(28) donár=**nos=en** [dunárnuzən] Catalan  
 give=us=partit  
 ‘to give us some’ (Wheeler 1979: 34)

- In Maltese, short vowels in unstressed open syllables are deleted. The application of this rule on clitics is shown in (19).

(29) Clitic preposition *bi* ‘with’ in Maltese (Borg & Azzopardi-Alexander 1997: 310)  
**bi** travu → /bɪ.tɾɐ:.vʊ/ or /pɾɐ:.vʊ/ ‘with a beam’

- In Maori, one of the most constant correlates of primary stress is lengthening of the vocalic element. This phonetic lengthening of stressed vowels is also present in the proclitic article *te* in combination with the deictic *nei* in (21). Here the proclitic gains additional phonetic substance and is thus strengthened.

(30) Clitic article *te* in Maori (Bauer 1993: 507)

'tee=nei  
 the=proxI  
 ‘this’

### 3. Towards a rhythm-based typology of erosion

#### 3.1. Phonological properties of rhythmic types

- Within a rhythm-based typology of language, the following phonological properties are associated with mora-, syllable- and stress-based languages (see Pike 1945: 34ff; Roach 1982; Donegan & Stampe 1983; Dauer 1983, 1987; Gil 1986; Auer & Uhmman 1988; Auer 1993, 2001; Dufter 2003).

(31) Rhythmic profiles (adapted from Auer 1993: 14)

<b>(mora-based) syllable-based</b>	<b>stress-based</b>
no accent-dependent reduction	reduction of non-accented syllables in quality and/or duration
[±long] in consonants and vowels of all syllables possible	no [±long] distinction in non-accented syllables
tone possible	no tone (or non-accented syllables are “neutral”)
simple syllable structure open syllables	complex syllable structure, sonority scale disobeyed

few assimilations	frequent assimilations, dissimilations
syllable division unambiguous	syllable division ambiguous and variable
no word-related phonological processes	word-related phonological processes
external = internal Sandhi	external $\neq$ internal Sandhi
vowel harmony possible	no vowel harmony
phonetically weak word accent or none at all	phonetically strong word accent, realized by pitch (and other prosodic features)
word accent (if any) fixed, no grammatical functions	word accent assigned by complex rules referring to syllable structure, partly morphologized, or free, may have grammatical functions
geminate possible	no geminates
no central (“reduced”) vowel phonemes	central vowel phonemes possible

- Assuming that this basic typological distinction is true, we can suspect that the differences in segmental effects of cliticization that we observed follow from general phonological patterns in the language and are language-type specific.

### **3.2. Correlations**

#### **3.2.1. Syllable structure**

- Restriction in syllable types is the most robust phonological dimension in which languages of the different rhythmic types differ (Ramus, Nespore & Mehler 1999). Whereas mora-based languages favor CV syllables, onset and coda clusters are more complex in syllable-based languages and even more so in stress-based languages.

(32) Canonical syllable structures

CV	CVC	CVCC	CCVC	CCVCC	CCCVCC	CCCCCVCCCC
Maori						
Catalan						
Yoruba						
Nigerian Pidgin						
Udihe						
Tariana						
Chukchi						
West Greenlandic						
Amele						
Slave						
Koyra Chiini						
Mandarin						
		Kayardild				
		Finnish				
			!Xóõ			
				Tamil		
				Basque		
					Maltese	
						Georgian

- Phonotactic restrictions on syllable structure also constrain phonological processes at the host-clitic boundary.
- i) In order to retain an overall CV structure, consonants in front of vowel-initial clitics are *preserved* at the morpheme boundary, cf. Nigerian Pidgin (18).
- ii) If a language allows CVC structures, both onset and coda consonants at clitic boundaries are preserved and may *assimilate*, cf. Koyra Chiini (19).
- iii) *Cluster simplification* prevents consonant clusters prohibited in the language, cf. West Greenlandic (20).
- iv) *Vowel coalescence* prevents illicit vowel sequences, cf. Tamil (21).
- v) *Epenthesis* provides an otherwise lacking onset consonant for CVC syllables and resolves illicit hiatus, cf. West Greenlandic (22).

3.2.2. Segmental effects of stress

- According to a hypothesis about the evolution of stress (Hyman 1977, Bybee et al. 1998), stress originates from the reinterpretation of phrasal intonation patterns as accents on words. The innovation of duration as a phonetic correlate of prosodic

dominance may cause the phonology of a language to develop segmental effects of stress, most notably vowel reduction in unstressed syllables.

(33) “Implicational universal I: The presence of vowel lengthening in stressed syllables implies presence of vowel reduction in unstressed syllables.” (Bybee et al. 1998: 286)

(34) “Implicational universal II: The presence of consonant changes conditioned by stress in a language implies the presence of vowel reduction in unstressed syllables.” (Bybee et al. 1998: 288)

- Following the ‘stress-accent hypothesis’, we can distinguish at least two types, namely ‘non-stress accent’, with pitch as the main phonetic correlate, and ‘stress accent’, which uses material other than pitch to a greater extent (Beckman 1986). Additionally, languages may exhibit stress which is only isolatable in intonation contours and which is often referred to as weak stress, or may completely lack stress.

(35) Prosodic types

no stress	weak stress	non-stress accent	stress accent
West Greenlandic (intonation)			
Slave (tone)			
!Xóõ (tone)			
Yoruba (tone)			
Basque/Gernika (pitch-accent)			
Nigerian Pidgin (pitch-accent/tone)			
	Tamil		
	Chukchi		
	Georgian		
		Udihe	
		Finnish	
			Maori
			Mandarin (tone)
			Amele
			Tariana
			Koyra Chiini
			Kayardild
			Catalan
			Maltese

- The languages in the sample exhibit a diverse behavior with respect to segmental effects of stress. Only languages of the stress accent type show reduction in unstressed syllables. Whereas Maori only shows lengthening of vowels in stressed syllables, Mandarin shows marginal centralization of vowels in unstressed syllables, Amele only of those in word-final position, Tariana only in syllables adjacent to stressed ones, while Koyra Chiini, Kayardild, and Catalan centralize all unstressed vowels. Maltese deletes unstressed vowels completely.
- Note also that tones and vowel harmony are almost exclusively reported for languages which have not developed segmental effects of stress.

### **3.3. Generalizations**

- Junctural erosion in the form of vowel coalescence and cluster simplification is more likely to occur in mora-based and syllable-based phonologies. Within these systems, phonotactic constraints determine the possible segmental shapes of clitics and other morphemes.
- Syllabic erosion in the form of vowel reduction and finally deletion is more likely to occur in stress-based phonologies. Within these systems, stress reduction goes hand in hand with segmental reduction. Notably, this is not a prosodic process restricted to cliticization.

## **4. Summary & outlook**

- Deaccentuation of clitics can lead to diverse behavior in prosodic systems. Although they don't carry their own stress, they might be stressable or unstressable within the domain into which they are (to varying degrees) integrated. In tonal phonologies, their deficiency is manifested by reduced tonal contrasts in tonal Sandhi. In purely intonational systems, the concept of deficiency might not be applicable at all.

- With respect to segmental effects of cliticization/deaccentuation, both lengthening and reduction depend on the extent to which a language has developed into stress accent prosody. Vowel reduction and vowel harmony in clitics seem to be mutually exclusive, as well as vowel reduction and tone realization.
- In more recent work on grammaticalization, it has been noted again and again that certain grammaticalization processes are not paralleled by phonological erosion (e.g. Wiemer 2004 on the evolution of passive constructions) and that certain languages systematically show little or no erosion in grammaticalization (e.g. Bisang 2004).
- With respect to morphologization, systematic reduction in languages like Maltese leads to subminimal clitics and suffixes, whereas systematic lack of reduction in languages like Tamil leads to bisyllabic clitics and suffixes.

(36) Maltese \**xaj* 'thing' > negative clitic/suffix *x*

ħrig=tu=x mill-forn kmieni wisq?

took.out.2.Sg.=3.Masc.Sg.=NEG from.the=oven early too-much

“Did you take it out of the oven too soon?” (Borg & Azzopardi-Alexander 1997: 4)

(37) Tamil \**iruntu* (past participle of *iru* ‘to be’) > clitic postposition *runtu* ‘ablative’

naan inta pustakatte murugaṅ=kitte=runtu eraval vaṅkaneen

I this book-acc Murugaṅ=loc=abl loan obtain-past-1s

‘I borrowed this book from Murugaṅ.’ (Asher 1982: 111)

Maltese: “How come you guys have such long words?”

Tamil: “Blame it on the Boogie!”

## References

- Aikhenvald, Alexandra Y. 2002. “Typological parameters for the study of clitics, with special reference to Tariana.” In: Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds.). *Word: A cross-linguistic typology*. Cambridge: Cambridge University Press: 42-78.
- Aikhenvald, Alexandra Y. 2003. *A Grammar of Tariana*. (Cambridge Grammatical Descriptions). Cambridge: Cambridge University Press.
- Akinlabi, Akinbiyi & Mark Liberman 2000. “The tonal phonology of Yoruba clitics.” In: Gerlach, Birgit & Janet Grijzenhout (eds.). *Clitics in Phonology, Morphology and Syntax*. Amsterdam: Benjamins: 31-62.

- Anderson, Stephen R. 1992. *A-morphous Morphology*. (Cambridge Studies in Linguistics, 62). Cambridge: Cambridge University Press.
- Asher, R. E. 1982. *Tamil*. (Lingua Descriptive Studies, 7). Amsterdam: North-Holland.
- Auer, Peter 1993. *Is a rhythm-based typology possible? A study on the role of prosody in phonological typology*. (Arbeitspapier, 21). Konstanz: Fachgruppe Sprachwissenschaft, Universität Konstanz.
- Auer, Peter 2001. "Silben- und akzentzählende Sprachen." In: Haspelmath, Martin; Ekkehard König; Wulf Oesterreicher & Wolfgang Raible (eds.). *Language Typology and Language Universals*. Berlin/ New York: Walter de Gruyter: 1391-1399.
- Auer, Peter & Susanne Uhmann 1988. "Silben- und akzentzählende Sprachen. Literaturüberblick und Diskussion." *Zeitschrift für Sprachwissenschaft* 7: 214-259.
- Bauer, Winifred 1993. *Maori*. (Descriptive Grammars). London/New York: Routledge.
- Beckman, Mary E. 1986. *Stress and Non-Stress Accent*. (Netherland Phonetic Archives, 7). Dordrecht: Foris.
- Bisang, Walter 2004. "Grammaticalization without coevolution of form and meaning: The case of tense-aspect-modality in East and mainland Southeast Asia." In: Bisang, Walter; Nikolaus P. Himmelmann & Björn Wiemer (eds.). *What makes Grammaticalization? A Look from its Fringes and its Components*. (Trends in Linguistics; Studies and Monographs, 158). Berlin/New York: Mouton de Gruyter: 109-138.
- Borg, Albert & Marie Azzopardi-Alexander 1997. *Maltese*. (Descriptive Grammars). London: Routledge.
- Bybee, Joan; Revere Perkins & William Pagliuca 1994. *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago/London: University of Chicago Press.
- Bybee, Joan; Paromita Chakraborti; Dagmar Jung & Joanne Scheibman 1998. "Prosody and segmental effect. Some paths of evolution for word stress." *Studies in Language* 22: 267-314.
- Croft, William 2003. *Typology and Universals*. Second Edition. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Dauer, Rebecca M. 1983. "Stress-timing and syllable-timing reanalyzed." *Journal of Phonetics* 11: 51-62.
- Dauer, Rebecca M. 1987. "Phonetic and phonological components of language rhythm." In: *Proceedings of the Eleventh International Congress of Phonetic Sciences*, Volume 5. Tallinn: Academy of Sciences of the Estonian S.S.R.: 447-450.
- Donegan, Patricia J. & David Stampe 1983. "Rhythm and the holistic organization of language structure." In: Richardson, John F.; Mitchell Marks & Amy Chukerman (eds.). *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: Chicago Linguistic Society: 337-353.
- Dufter, Andreas 2003. *Typen sprachrhythmischer Konturbildung*. (Linguistische Arbeiten, 475). Tübingen: Niemeyer.
- Dunn, Michael J. 1999. *A Grammar of Chukchi*. PhD-Dissertation, 1999. Australian National University.
- Evans, Nicholas D. 1995. *A Grammar of Kayardild*. With Historical-Comparative Notes on Tangkic. (Mouton Grammar Library, 15). Berlin/New York: Mouton de Gruyter.
- Faraclas, Nicholas G. 1989. *A Grammar of Nigerian Pidgin*. PhD-Dissertation, University of California, Berkeley.
- Fortescue, Michael 1984. *West Greenlandic*. (Croom Helm descriptive grammars). London: Croom Helm.
- Gil, David 1986. "A prosodic typology of language." *Folia Linguistica* 20: 165-231.

- Harris, Alice C. 2002. "The word in Georgian." In: Dixon, R.M.W. & Alexandra Y. Aikhenvald (eds.). *Word. A cross-linguistic typology*. Cambridge: Cambridge University Press: 227-242.
- Haspelmath, Martin 2002. *Understanding Morphology*. (Understanding Language Series). London: Arnold.
- Heath, Jeffrey 1999. *A Grammar of Koyra Chiini*. The Songhay of Timbuktu. (Mouton Grammar Library, 19). Berlin/New York: Mouton de Gruyter.
- Heine, Bernd; Ulrike Claudi & Friederike Hünemeyer 1991. *Grammaticalization. A Conceptual Framework*. Chicago/London: University of Chicago Press.
- Heine, Bernd & Mechthild Reh 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hopper, Paul J. & Elizabeth Closs Traugott 1993. *Grammaticalization*. (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press.
- Hualde, José Ignacio 1991. *Basque Phonology*. (Theoretical Linguistics). London: Routledge.
- Hyman, Larry M. 1977. "On the nature of linguistic stress." In: Hyman, Larry M. (ed.). *Studies in stress and accent*. (Southern California Occasional Papers In Linguistics, 4). Los Angeles: Department of Linguistics, University of Southern California: 37-82.
- Klavans, Judith L. 1985. "The independence of syntax and phonology in cliticization." *Language* 61: 95-120.
- Kleinschmidt, Samuel 1968 [1851]. *Grammatik der grönländischen Sprache*. Hildesheim: Olms.
- Lehmann, Christian [1982] 1995. *Thoughts on Grammaticalization*. (Lincom Studies in Theoretical Linguistics, 1). München: Lincom Europa.
- Lin, Hua 2001. *A Grammar of Mandarin Chinese*. (Languages of the World/Materials, 344). München: Lincom Europa.
- Mascaró, Joan & Gemma Rigau 2002. "The Grammar of Clitics." *Catalan Journal of Linguistics* 1: 9-15.
- Nevis, Joel A. 1986. "Finnish Particle Clitics and General Clitic Theory." *Ohio State University Working Papers in Linguistics* 33: 1-159.
- Nikolaeva, Irina & Maria Tolskaya 2001. *A Grammar of Udihe*. (Mouton Grammar Library, 22). Berlin/New York: Mouton de Gruyter.
- Pike, Kenneth L. 1945. *The Intonation of American English*. Ann Arbor: The University of Michigan Press.
- Ramus, Franck; Nespors, Marina & Jacques Mehler 1999. "Correlates of linguistic rhythm in the speech signal." *Cognition* 73: 265-292.
- Rice, Keren D. 1989b. "Vowel initial suffixes and clitics in Slave." In: Gerds, Donna B. & Karin Michelson (eds.). *Theoretical Perspectives on Native American Languages*. (SUNY Series in Linguistics). Albany: State University of New York Press: 104-132.
- Rijkhoff, Jan & Dik Bakker 1998. "Language sampling." *Linguistic Typology* 2: 263-314.
- Rijkhoff, Jan; Dik Bakker; Kees Hengeveld & Peter Kahrel 1993. "A method of language sampling." *Studies in Language* 17: 169-203.
- Rischel, Jørgen 1974. *Topics in West Greenlandic Phonology. Regularities Underlying the Phonetic Appearance of Wordforms in a Polysynthetic Language*. Copenhagen: Akademisk Forlag.
- Roach, Peter 1982. "On the distinction between 'stress-timed' and 'syllable-timed' languages." In: Crystal, David (ed.). *Linguistic Controversies: Essays in Linguistic Theory and Practice in Honour of F. R. Palmer*. London: Edward Arnold: 73-79.
- Roberts, John R. 1996. "A Government and Binding Analysis of the Verb in Amele." *Language & Linguistics in Melanesia: Journal of the Linguistic Society of Papua New Guinea* 27: 1-66.

- Ruhlen, Merrit 1991. *A Guide to the World's Languages. Volume 1: Classification*. With a postscript on recent developments. Stanford: Stanford University Press.
- Schiffman, Harold F. 1999. *A Reference Grammar of Spoken Tamil*. Cambridge: Cambridge University Press.
- Trill, Anthony 1985. *Phonetic and Phonological Studies of !Xóõ Bushman*. (Quellen zur Khoisan-Forschung, 1). Hamburg: Buske.
- Wheeler, Max 1979. *Phonology of Catalan*. (Publications of the Philological Society, 28). Oxford: Blackwell.
- Wiemer, Björn 2004. "The evolution of passives as grammatical constructions in Northern Slavic and Baltic languages." In: Bisang, Walter; Nikolaus P. Himmelmann & Björn Wiemer (eds.). *What makes Grammaticalization? A Look from its Fringes and its Components*. (Trends in Linguistics; Studies and Monographs, 158). Berlin/New York: Mouton de Gruyter: 271-331.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington: Indiana University Linguistics Club.
- Zwicky, Arnold M. 1985. "Clitics and Particles." *Language* 61: 283-305.
- Zwicky, Arnold M. 1994. "Clitics." In: Asher, R. E. & J. M. Y. Simpson (eds.). *The Encyclopedia of Language and Linguistics*. Volume 2. Oxford et al.: Pergamon Press: 571-576.
- Zwicky, Arnold M. & Geoffrey K. Pullum 1983. "Cliticization vs. inflection: English *n't*." *Language* 59: 502-513.

## **Appendix A: Language Sample**

Nr.	Language	Macroarea, Area	Classification
1	MALTESE	Europe, Malta	Semitic
2	UDIHE	Eurasia, Eastern Siberia	Tungus-Manchu
3	TARIANA	South America, Northwest Amazonia	Arawak
4	KAYARDILD	Australia, Queensland	Tangkic
5	MAORI	Oceania, New Zealand	Polynesian
6	GEORGIAN	Eurasia, Georgia, Turkey, Iran	Kartvelian
7	CHUKCHI	Eurasia, North Eastern Siberia	Chuckchi-Kamchatkan
8	TAMIL	Asia, India	Dravidian
9	WEST GREENLANDIC	Europe, Greenland	Eskimo-Aleut
10	CATALAN	Europe, Spain, France, Sardinia	Romance
11	AMELE	Oceania, Papua New Guinea	Gum
12	!XÓŃ BUSHMAN	Africa, Botswana, Namibia	Southern Khoisan
13	BASQUE	Europe, Spain, France	Language Isolate
14	SLAVE	North America, Canada	Na-Dene
15	YORUBA	Africa, Nigeria	Niger-Congo
16	KOYRA CHIINI	Africa, Mali, Timbuktu	Songhai
17	NIGERIAN PIDGIN	Africa, Nigeria	Pidgin, English-Based
18	MANDARIN CHINESE	Asia, China	Sinitic
19	FINNISH	Europe, Finland	Uralic

## Appendix B: Sample Entry from the Clitics Database

LANGUAGE	Finnish
AFFILIATION	Uralic
CLITIC	kO
FULL FORM	[NO]
WORD-CLASS	sentence particle
SEMANTICS	interrogative
EXAMPLE	(1) On= <b>ko</b> Pertti naimisissa? is= <b>Q</b> Pertti married 'Is Pertti married?' (2) Pertti= <b>kö</b> on naimisissa? Pertti= <b>Q</b> is married 'Is it Pertti who is married?' (Nevis 1986d: 9f.)
LIAISON	enclitic
PROSODIC PHONOLOGY	unstressed, regularly secondary stressed
SEGMENTAL PHONOLOGY	undergoing vowel harmony
MORPHOPHONOLOGICAL IDIOSYNCRASIES	[NO]
HOST	verb, noun, participle, etc.
ORDER	outside derivational/inflectional morphology
TEMPLATE	=POSS=PARTICLE CLITICS
PARADIGMATICITY	[N.A.]
PARADIGM	[N.A.]
PARADIGM FULL FORM	[N.A.]
DOMAIN	S
PLACEMENT	2P
CLASSIFICATION	clitic, particle clitic, bound word
RATIONALE	phonological behavior, syntactic distribution
SOURCES	Sulkala & Karjalainen (1992), Nevis (1986d)
COMMENT	2P or phrase final?
SAMPLE	A - 19 (Pilot)