词类大小和重新分析
WORD CLASS SIZE AND REANALYSIS

马嘉思 **Matthias Gerner**

第三届语言类型学国际学术研讨会，
上海外国语大学，2017年7月14–16日

The 3rd Biennial Conference on Chinese Linguistic Typology,

Shanghai International Studies University, July, 14-16 2017

Contents:

1. **Introduction**
2. **Reanalysis of Selected Examples**
3. **Correlation between Word Class Size and Reanalysis**

1. Introduction

§1 Definitions

- A *word class* is a set of morphemes or independent words with common semantic-pragmatic, morphosyntactic and (sometimes) phonological properties;

- *Word class reanalysis* (either grammaticalization or lexicalization) is a process whereby the members of the word class undergo **coordinated semantic change**;

- Excluded: Chinese or English prepositions which changed individually not collectively;

- Reminder: *Grammaticalization* (Hopper & Traugott 1993)
  - ✓ Changes lexical items/constructions to serve grammatical functions;
  - ✓ Changes grammatical items/constructions to serve new grammatical functions;

- Reminder: *Lexicalization*
  - ✓ Puts grammatical items outside grammar rules and integrates them into lexicon (Anttila 1989);
  - ✓ Changes lexical items to take on a new form or lexical meaning (Bauer 1983);

2. Reanalysis of Selected Examples

§1 Example: 彝族语系 Burmese-Lolo Group (Gerner 2007)

(1) G Gejiu Nesu (Loloish language in China)

a.

kɚ⁵⁵	a⁵⁵ko³³	tɕʰe²¹	tʰɛ³³	**ba⁵⁵**	sə³³.
3P SG	child	NUM:1	CL	carry	PROG

'She carries a child on the back.'

b.

ŋo³³	a⁵⁵ko³³	gu²¹	kɚ⁵⁵	**pa⁵⁵**.
1P SG	child	PAT	3P SG	make carry

'I make him carry the child.'

(2) A Weining Neasu (Yi language: Weining County in Guizhou)

a.

si³³	ga⁵⁵	dze¹³	kɔ⁵⁵	**lɤ³³**.
tree	DEM:DIST	CL	PROG	move

'That tree is moving (e.g. its branches and leaves by the wind).'

b.

si³³	ga⁵⁵	dze¹³	ɕɪ²¹	le⁵⁵	**ɬɤ³³**.
tree	DEM:DIST	CL	3P SG	PASS	move

'That tree is moved by him (e.g. by shaking the stem).'

(3) I Yongren Lolo (Yi language: Yongren County in Yunnan)

a.

zʮ²¹	bɛ³³tsɨ³³	tsɛ³³lu³³	tʰi²¹	kʰə³³	**ʐi²¹**	do³³.
3PSG	garment	beautiful	NUM:1	CL	wear	PROG

'She wears a beautiful garment.'

b.

ɔ²¹mo³³	zʮ²¹	tʰie²¹	bɛ³³tsɨ³³	gɛ⁵⁵	**ɕi⁵⁵**.
mother	3P SG	BEN	garment	PAT	dress

'Mother dressed her with a garment.'

2. Reanalysis of Selected Examples

§1 Example: 彝族语系 Burmese-Lolo Group (Gerner 2007)

Table 1: Simplex/complex verb pairs in the Loloish group

		@	A	B	C	D	E	F	G	H	I	J
'open' (intr.)	Simplex		bu^{21}	po^{33}	$p^h\mathfrak{o}^{13}$		bu^{33}	po^{33}				
'open' (tr.)	Complex		p^hu^{21}	p^ho^{33}	$p^h\mathfrak{o}^{13}$		p^hu^{33}	po^{33}				
'drink'	Simplex	ndo^{33}	$nd\mathfrak{o}^{33}$	to^{33}	$nt^{fi}o^{21}$	do^{13}		du^{33}	da^{21}	du^{55}	$d\mathfrak{o}^{33}$	$d\mathfrak{o}^{33}$
'give to drink'	Complex	to^{21}	$t\mathfrak{o}^{21}$	to^{55}	$nt^{fi}o^{21}$	to^{33}		tu^{33}	$t\varepsilon^{21}$	tu^{33}	$t\mathfrak{o}^{33}$	$t\mathfrak{o}^{33}$
'be afraid'	Simplex	$g\underset{\sim}{u}^{33}$									$g\underset{\sim}{u}^{33}$	
'frighten'	Complex	$k\underset{\sim}{u}^{33}$									$k\underset{\sim}{u}^{33}$	

The numbers refer to the following Loloish languages:

@ Liangshan Nuosu; A Weining Neasu; B Longlin Ngopho; C Luoping Nase; D Shizong Kopho; E Mile Axi; F Mile Azhee; G Gejiu Nesu; H Weishan Lalo; I Yongren Lolo; J Wuding Aluphu.

2. Reanalysis of Selected Examples

§1 Example: 彝族语系 Burmese-Lolo Group (Gerner 2007)

Summary:

➢ A Proto-Tibeto-Burman prefix became unproductive and disappeared by lexicalizing in the Loloish Group (and beyond);

➢ This process of lexicalization is old;

➢ The process acted upon the open class of verbs;

➢ The process created one lexical meaning, the meaning of causative verbs, though not for all verbs;

➢ Depending on the Loloish language, the lexicalized verbs number between 15 and 50.

Language/Family	Word Class	Size	Output	Age	Type
Tibeto-Burman (TB)	Causative Verbs	open	0/1	old	lexicalization

2. Reanalysis of Selected Examples

§ 2 Example 侗语北部方言 Northern Kam (TK) (Gerner 2006)

Northern Kam							
(4) a.	i^{45}	jiu^{22}	ηa^{45}		b.	ham^{11}	$\underset{\cdot}{t}iu^{22}$	ηa^{45}
	NUM:1	CL-SG	river			NUM:3	CL-PL	river
	Numeral	Noun classifier	Noun			Numeral	Noun classifier	Noun
	'One river'					'Three rivers'		

Southern Kam							
(5) a.	i^{55}	$\underset{\cdot}{t}iu^{11}$	ηa^{55}		b.	sam^{35}	$\underset{\cdot}{t}iu^{11}$	ηa^{55}
	NUM:1	CL-SG	river			NUM:3	CL-PL	river
	Numeral	Noun classifier	Noun			Numeral	Noun classifier	Noun
	'One river'					'Three rivers'		

Northern Kam							
(6) a.	$ma\eta^{55}$	ta^{45}	ai^{33}		b.	to^{22}	$lau^{31}k\partial n^{22}$	i^{44}
	CL-PL	eye	DEM:PROX			CL-PL	friend	DEM:MED
	Classifier	Noun	Demonstrative			Numeral	Noun	Demonstrative
	'This eye'					'that friend'		

2. Reanalysis of Selected Examples

§2 Example 侗语北部方言 Northern Kam (TK) (Gerner 2006)

Derivation	Classifier meaning	Northern		Rongjiang	Sanjiang
		CL-SG	CL-PL	CL	CL
$[p] \rightarrow [w]$	Entities with handle	wa^{33}	pa^{33}	pak^{323}	pak^{323}
	Human	$wəu^{45}$	$pəu^{45}$	---	---
	'bridge'	wu^{44}	pu^{44}	---	pu^{33}
$[m] \rightarrow [w]$	Dual body parts	$waŋ^{24}$	$maŋ^{55}$	$maŋ^{53}$	$maŋ^{53}$
	Clothes	$wəi^{31}$	$məi^{31}$	$məi^{31}$	$məi^{31}$
$[ɬ] \rightarrow [w]$	2-Dim entities	wen^{11}	$ɬen^{11}$	---	---
$[k^w] \rightarrow [w]$	'piece'	wai^{24}	k^wai^{24}	---	$k^{hw}ai^{453}$
$[t] \rightarrow [ẓ]$	'lump, ball'	$ẓa^{22}$	ta^{22}	---	ta^{11}
	Animate entities	$ẓo^{22}$	to^{22}	tu^{11}	tu^{11}
$[n] \rightarrow [ẓ]$	3-Dim entities	$ẓən^{11}$	$nən^{45}$	$nɐn^{55}$	$nɐn^{55}$
	Several versatile entities	$ẓa^{13}$	$naŋ^{13}$	---	---
$[t] \rightarrow [n]$	Several versatile entities	na^{31}	ta^{31}	---	---
$[ɬ] \rightarrow [j]$	1-Dim entities	jiu^{22}	$ɬiu^{22}$	$ɬiu^{11}$	$ɬiu^{11}$
	Drop-shaped entities	jit^{33}	$ɬit^{33}$	$ɬik^{323}$	$ɬik^{323}$
$[ɕ] \rightarrow [j]$	Entities with handle	$jaŋ^{45}$	$ɕaŋ^{45}$	$ɕaŋ^{53}$	$ɬaŋ^{53}$
	Erected/layered entities	$joŋ^{22}$	$ɕoŋ^{22}$	$ɕoŋ^{11}$	$ɕoŋ^{11}$
$[k] \rightarrow [ɣ]$	Several versatile entities	$ɣaŋ^{44}$	$kaŋ^{44}$	---	---
$[ʔ] \rightarrow [ɣ]$	Vehicles & machines	$ɣa^{55}$	$ʔa^{55}$	$ʔa^{53}$	ka^{53}

2. Reanalysis of Selected Examples

§ 2 Example 侗语北部方言 Northern Kam (TK) (Gerner 2006)

Summary:

➢ The singular/plural classifiers were created by contact assimilation with the numeral i^{45} 'one', borrowed from Chinese;

➢ The classifiers with assimilated initial consonant were reanalyzed as singular classifiers;

➢ When the same classifiers occurred with higher numerals, their initial consonant was not assimilated;

➢ These occurrences were reanalyzed as plural classifiers;

➢ This process of lexicalization is young;

Language/Family	Word Class	Size	Output	Age	Type
Northern Kam (TK)	Plural Classifiers	31	2	young	grammaticalization

2. Reanalysis of Selected Examples

§ 3 Sursurunga (Austronesian) (Hutchisson 1975; 1977; 1986)

➢ Sursurunga is an Austronesian language spoken by 3,000 in Papua New Guinea 巴布亚新几内亚;

➢ The numbers 'two', 'three' and 'four' have merged with the plural pronouns to form dual, trial and quadral pronouns;

➢ The process is middle-aged

	PERSON	SG (singular)	DU (dual) ur / ar 'two'	TRL (trial) tul 'three'	QUAD (quadral) hat 'four'	PL(plural) -∅
INCLUSIVE -t-	1		gi-t-ar	gi-t-tul	gi-t-at	gi-t-∅
EXCLUSIVE -m-	1	iau	gi-ur	gi-m-tul	gi-m-at	gi-m-∅
	2	iáu	ga-ur	gam-tul	gam-at	gam-∅
	3	i/on/ái	di-ar	di-tul	di-at	di-∅

Language/Family	Word Class	Size	Output	Age	Type
Austronesian (A)	Plural Classifiers	7	3	middle	grammaticalization

2. Reanalysis of Selected Examples

§ 3 English (Indo-European) (Lightfoot 1979)

➢ In Middle English (中古英语) there was a small class of preterit-present verbs within the class of lexical verbs;

➢ In the 16th century, they were rapidly reanalyzed acquiring and monopolizing two grammatical functions: alternative question and negation. All lexical verb could directly assume these functions

➢ The acquisition of these functions created the category of modal auxiliary verbs:

  ➢ Inversion in direct questions (*can you speak* vs. **speak you*);
  ➢ Negation (*he will not win* vs. **he not win*);
  ➢ Occurrence in sentence-final tags (*he will win, won't he* vs. **he wins, not wins he*);
  ➢ Bare infinitive VP complements (*will meet him* vs. **is meet him*);
  ➢ No *to*-infinitives (**to can speak* vs. *to be spoken*);
  ➢ No participle forms (**hasn't canned* vs. *has been taken*);
  ➢ No *-ing* gerunds (**canning do it* vs. *to be able to do it*);

2. Reanalysis of Selected Examples

§ 3 English (Indo-European) (Lightfoot 1979)

- The category of modal auxiliary verbs ultimately comprised 11 items: *will, would, shall, should, can, could, may, might, do, did, must*;

- This process of lexicalization is young;

Language/Family	Word Class	Size	Output	Age	Type
English (IE)	Modal Auxiliaries	11	2	young	grammaticalization

2. Reanalysis of Selected Examples

§ 4 Tamil (Dravidian) (Asher 1982)

➢ Honorific pronouns in Tamil are reanalyzed plural pronouns

➢ The mechanism is well-attested in languages of the world (e.g. French *vous* 'you plural' and *vous* 'you singular polite')

➢ This process of grammaticalization is young;

Person	Singular		Plural		Honorific	
	PROX	DIST	INCL/PROX	EXCL/DIST	PROX	DIST
1st SG	naan		naampa/naama	naanga	naama (king: 'we')	
2nd SG	nii		niinga		niinga/niir	
3rd MALE	ivan	avan	ivanga	avanga	ivaru	avaru
3rd FEMALE	iva	ava	ivanga	avanga	ivanga	avanga
3rd NEUTER	idu	adu				

Language/Family	Word Class	Size	Output	Age	Type
Tamil (D)	Honorific Pronouns	8	0/1	young	grammaticalization

2. Reanalysis of Selected Examples

§ 5 Summary

Language/Family	Word Class	Size	Output	Age	Direction
Tibeto-Burman (TB)	Causative Verbs	open	0/1	old	lexicalization
Neasu (TB)	Demonstratives	3	1/2	young	grammaticalization
Northern Yi group (TB)	Verb 'say'	1	4	middle	grammaticalization
Sunwar (TB)	Nonpast Tense	open	1	young	grammaticalization
Northern Kam (TK)	Plural Classifiers	31	2	young	grammaticalization
Kam-Tai (TK)	Verb 'touch'	1	3	middle	grammaticalization
Romance (IE)	Demonstratives	3	0/1	middle	grammaticalization
English (IE)	Modal Auxiliary Verbs	11	2	young	grammaticalization
Albanian (IE)	Participles	open	0/1	middle	lexicalization
Sursurunga (A)	Dual, Trial, Quadral	7	3	middle	grammaticalization
Polynesian (A)	Dual Pronouns	6	1	old	grammaticalization
Ahmao (MY)	Inflection classifiers	51	6	young	grammaticalization
Tamil (D)	Honorific Pronouns	8	0/1	young	grammaticalization
Omotic group (AA)	Emphatic Pronouns	6	1	middle	grammaticalization
Tirana (AR)	Evidential Verbs	3	1	young	grammaticalization
Panare (C)	Demonstratives	2	1	young	grammaticalization
Siouan (SI)	Demonstratives	2	2	middle	grammaticalization
Turkic (T)	Auxiliary Postverbs	12	1	old	grammaticalization
Muskogean (M)	Auxiliary Conjugation	50	2	middle	grammaticalization
Abui (TNG)	Demonstratives	2	1	young	grammaticalization
Bantu group (NG)	O agreement markers	6	1	middle	grammaticalization
Bamileke-Dschang (NG)	Tense Auxiliaires	10	1	middle	grammaticalization
Jamul Tiipay (Y)	Possessor Prefixes	3	3	young	grammaticalization
Kawaiisu (UA)	Demonstratives	3	1	young	grammaticalization

3. Correlation of Class Size and Reanalysis

§ 1 Greenberg's Universals (Greenberg 1963)

(7)	Greenberg (1963)'s Universal 2
	"In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes."

(8)	Type of variable	Number of Values
a.	nominal	Two values
b.	ordinal	More than two but finite number of values
c.	continuous	Many values possibly infinite countable or infinite non-countable

(9)	The correlation between two nominal variables
	If +X, then (almost always) +Y.

(10)	The correlation between two ordinal/continuous variables
a.	The more X, the more Y (positive).
b.	The more X, the less Y (negative).

3. Correlation of Class Size and Reanalysis

§ 1 Greenberg's Universals (Greenberg 1963)

(11)	Greenberg (1963)'s Universal 2
	"In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes."

(12)	Type of variable	Number of Values
a.	nominal	Two values
b.	ordinal	More than two but finite number of values
c.	continuous	Many values possibly infinite countable or infinite non-countable

(13)	Variable	Type of variable	Description
a.	adposition	nominal	a language has prepositions or postpositions
b.	possession	nominal	possessor precedes or follows the possessee

(14)	The correlation between two nominal variables
a.	If +X, then (almost always) +Y (positive).
b.	If +X, then (almost always) -Y (negative).

3. Correlation of Class Size and Reanalysis

§ 2 Universal Tendency

(15)		Variable	Type of variable	Description
	a.	size	continuous	number of elements in word class
	b.	output	continuous or ordinal	number of new reanalyzed meanings

(16)	The correlation between two ordinal/continuous variables
a.	The more X, the more Y (positive).
b.	The more X, the less Y (negative).

(17)	Possible correlations between 'size' and 'output'
a.	The greater the size of the word class, the higher the number of new acquired meanings (positive).
b.	The greater the size of the word class, the lower the number of new acquired meanings (negative).

3. Correlation of Class Size and Reanalysis

§ 3 Statistical Significance

Family	Word Class Studies
Tibeto-Burman (TB)	4
Tai-Kadai (TK)	2
Indo-European (IE)	3
Austronesian (A)	2
Miao-Yao (MY)	1
Dravidian (D)	1
Afro-Asiatic (AA)	1
Arawakan (AR)	1
Caribean (C)	1
Siouan (SI)	1
Turkic (T)	1
Muskogean (M)	1
Trans-New Guinea (TNG)	1
Niger-Congo (NG)	2
Yuman (Y)	1
Uto-Aztecan (UA)	1

3. Correlation of Class Size and Reanalysis

§ 3 Statistical Significance

Language/Family	Word Class	Size	Output
Tibeto-Burman (TB)	Causative Verbs	300	0.06
Neasu (TB)	Demonstratives	3	1.33
Northern Yi group (TB)	Verb 'say'	1	4
Sunwar (TB)	Nonpast Tense	300	1
Northern Kam (TK)	Plural Classifiers	31	2
Kam-Tai (TK)	Verb 'touch'	1	3
Romance (IE)	Demonstratives	3	0.5
English (IE)	Modal Auxiliary Verbs	11	2
Albanian (IE)	Participles	300	0.8
Sursurunga (A)	Dual, Trial, Quadral	7	3
Polynesian (A)	Dual Pronouns	6	1
Ahmao (MY)	Inflection classifiers	51	6
Tamil (D)	Honorific Pronouns	8	0.5
Omotic group (AA)	Emphatic Pronouns	6	1
Tirana (AR)	Evidential Verbs	3	1
Panare (C)	Demonstratives	2	1
Siouan (SI)	Demonstratives	2	2
Turkic (T)	Auxiliary Postverbs	12	1
Muskogean (M)	Auxiliary Conjugation	50	2
Abui (TNG)	Demonstratives	2	1
Bantu group (NG)	O agreement markers	6	1
Bamileke-Dschang (NG)	Tense Auxiliaires	10	1
Jamul Tiipay (YU)	Possessor Prefixes	3	3
Kawaiisu (UA)	Demonstratives	3	1

3. Correlation of Class Size and Reanalysis

§ 3 Statistical Significance

Scattergram

$y = -0.0032x + 1.83$
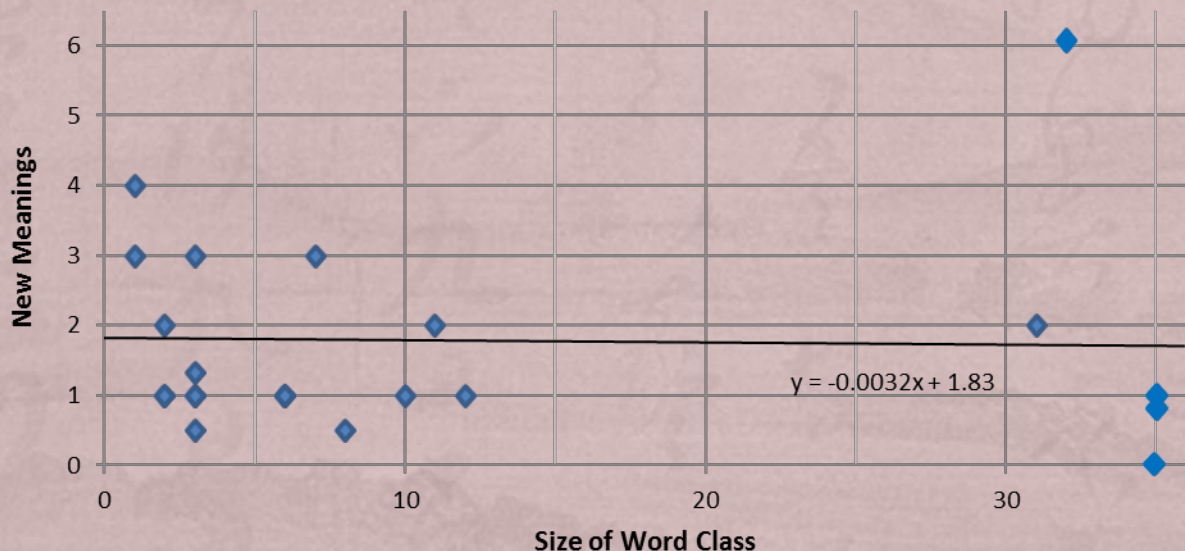
Pearson correlation coefficient

$$r = \frac{n\Sigma(xy) - \Sigma x \Sigma y}{\sqrt{\left(n\Sigma x^2 - (\Sigma x)^2\right)\left(n\Sigma y^2 - (\Sigma y)^2\right)}} = \frac{24 \cdot 1148.49 - 45052.99}{\sqrt{5384327 \cdot 974.6639}} = -0.241$$

- If 0 < r < 1, then there is (potential) positive correlation;
- If r = 0, then there is no correlation;
- If -1< r < 0, then there is (potential) negative correlation.

3. Correlation of Class Size and Reanalysis

§ 3 Statistical Significance

Result:

- The critical value of -0.241 suggests a fair correlation approximately at the significance level of 10%;

- In Social Sciences, in order to establish a new correlation between two variables, the significance level should be 5% or lower;

- The number of case studies is only 24; when we enlarge the corpus to 100 or 150 word class case studies, the correlation should be more pronounced.

(18)	Correlations between 'size' and 'output'
	The greater the size of the word class, the lower the number of new acquired meanings (negative).

谢 谢！