

# Word-Formation *in the* World's Languages

A Typological Survey

Pavol Štekauer  
Salvador Valera  
Livia Körtvélyessy

# Word-Formation in the World's Languages

## A Typological Survey

*By Pavol Štekauer, Salvador Valera, and Lívia Körtvélyessy*

PAVOL ŠTEKAUER, SALVADOR  
VALERA, LÍVIA KÖRTVÉLYESSY



CAMBRIDGE  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,  
São Paulo, Delhi, Dubai, Mexico City

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521765343](http://www.cambridge.org/9780521765343)

© Pavol Štekauer, Salvador Valera and Lívia Körtvélyessy 2012

This publication is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without the written  
permission of Cambridge University Press.

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

*A catalogue record for this publication is available from the British Library*

ISBN 978-0-521-76534-3 Hardback

Cambridge University Press has no responsibility for the persistence or  
accuracy of URLs for external or third-party internet websites referred to  
in this publication, and does not guarantee that any content on such  
websites is, or will remain, accurate or appropriate.

# Acknowledgements

The authors wish to express their gratitude to all the informants who were so kind as to supply the necessary data for this research. Without their invaluable help, this book would have been unthinkable. Unless otherwise specified, the contents of this book on each of the following languages were provided as personal communications and/or in the form of questionnaires by the informants. The source of examples is cited in the text only in languages for which two informants supplied data. Where the same data were provided by the two informants, the source is not explicit in the text. The informants are, in alphabetical order by language:

Afrikaans	Suléne Pilon and Gerhard B. van Huyssteen
Amele	John R. Roberts
Amharic	Grover Hudson
Anejom	John Lynch
Bardi	Claire Bower
Belorussian	Svetlana Rudaja and Alena Rudenka
Breton	Greg T. Stump
Catalan	Max Wheeler
Cirecire	Andy Chebanne
Clallam	Timothy Montler
Dangaléat	Erin Shay
Datooga	Roland Kießling
Diola-Fogny	Kirsten Fudeman
Dutch	Jan Don
English	Andrew Carstairs-McCarthy
Estonian	Annika Kilgi
Finnish	Vesa Koivisto and Johana Laakso
French	Dany Amiot
Gã	M.E. Kropp Dakubu
Georgian	Nino Amiridze
German	Christiane Dalton-Puffer
Greek	Angela Ralli

---

Hausa	Paul Newman
Hebrew	Ora Schwarzwald
Hungarian	Ferenc Kiefer
Ilocano	Carl Rubino
Indonesian	Franz Müller
Italian	Livio Gaeta
Japanese	Mark Volpe
Kalkatungu	Barry Blake
Karao	Sherri Brainard
Ket	Edward Vajda
Konni	Michael Cahill
Lakhota	Regina Pustet
Luganda	Xavier Luffin
Luo	Joost Zwarts
Maipure	Raoul Zamponi
Malayalam	K.-P. Mohanan
Mandarin Chinese	Karen Steffen Chung
Māori	Ray Harlow
Marathi	Veena Dixit
Nelemwa	Isabelle Bril
Portuguese	Rui Marques
Romanian	Niculina Iacob and Gina Măciucă
Russian	Peter Arkadiev
Serbian-Croatian	Gordana Štasni
Slavey	Keren Rice
Slovak	Ján Horecký and Martin Ološtiak
Spanish	Laura Malena Kornfeld
Swahili	Ellen Contini-Morava
Swedish	Arne Olofsson
Tamil	Harold F. Schiffman
Tatar	Uli Schamiloglu and Susan Wertheim
Telugu	Sailaja Pingali
Tibetan	Nathan W. Hill
Totonac	David Beck
Trinidadian Creole	Donald Winford
Ukrainian	Peter Lizanec
Vietnamese	Mark Alves and Nguyễn Thái Ân

West Greenlandic	Michael Fortescue
Wichí	Verónica Nercesian
Yoruba	Oye Taiwo
Zulu	Andrew van der Spuy

This publication was supported by the the Slovak Academic Grant Agency (VEGA), grant No. 1/2236/05, and Spanish Junta de Andalucía (Incentivos de Carácter Científico y Técnico de la Junta de Andalucía, 3/2008). We also owe a great deal to the linguists who generously shared their expertise with us and gave such insightful advice, recommendations and comments. They are, in alphabetical order:

Laurie Bauer  
 Winifred Bauer  
 Ruth A. Berman  
 Shmuel Bolozky  
 Geert Booij  
 Klára Buzássyová  
 Adelia Carstens  
 Taro Kageyama  
 Viktor Krupa  
 Rochelle Lieber  
 Lubor Mojdl  
 Jonathan Owens  
 Sergio Scalise  
 Dana Slančová  
 Chad Thompson  
 Ghil'ad Zuckermann

Special thanks are owed to four anonymous referees and to Leanne V. Bartley, Fraser T. Bayles, Miguel Ángel Benítez Castro, Karen Steffen Chung, Ana Díaz-Negrillo, Jesús Fernández-Domínguez, Encarnación Hidalgo Tenorio, Arne Olofsson, Sailaja Pingali, Franz Rainer, Francisco Valera Hernández and Mark Volpe for reading and commenting on the manuscript and to all those who were so kind as to provide us with the examples in their original script: Mark Alves (Vietnamese), Nino Amiridze (Georgian), Karen Steffen Chung (Mandarin Chinese), Niladri Dash (Hindi), Veena Dixit (Marathi), Taro Kageyama (Japanese), Harold Schiffman (Tamil), Ghil'ad Zuckermann and Ruth Berman (Hebrew) and Joost Zwarts (Luo).

# Notes on language-specific symbols

## Cirecire

This language has two clicks:

- |     |         |     |
|-----|---------|-----|
| (1) | dental  | [ǀ] |
| (2) | lateral | [ǁ] |

Clicks can appear with:

- |     |                       |            |
|-----|-----------------------|------------|
| (3) | aspiration            | [ǀh], [ǁh] |
| (4) | voicing               | [ǀg], [ǁg] |
| (5) | rarely, uvularization | [ǀx], [ǁx] |

Other symbols used are:

- |      |  |
|------|--|
| [c]  | for non-voiced affricated palatal stop |
| [dj] | for voiced affricated palatal stop     |
| [q]  | for affricated uvular stop             |

## Jaqaru

- |   |            |
|---|------------|
| “ | aspiration |
|---|------------|

## Kwakw'ala

- |                 |                  |
|-----------------|------------------|
| !               | glottalization   |
| = before suffix | sonantized sound |
| = before suffix | sonantized sound |

## Telugu

- |                     |   |                          |
|---------------------|---|--------------------------|
| <i>T, D, N, L</i>   | = | retroflex sounds         |
| <i>t, d</i>         | = | dentals                  |
| <i>ph, th, etc.</i> | = | aspirated sounds         |
| <i>c, j</i>         | = | post-alveolar affricates |

**Totonac**

## Orthography:

x	=	/j/
h	=	/ʔ/
j	=	/x/
lh	=	/t̺/
tz	=	/ts/
ch	=	/tʃ/
'	=	laryngealization of preceding vowel
,	=	glottalization of preceding consonant

**Zulu**

C1, C2, C3, ... Zulu nouns are divided into classes, each with its own agreement pattern. Many of the classes are marked with an identifying prefix. These prefixes are glossed here as C1, C2, C3, etc. Odd-numbered classes from 1 to 11 are singular classes. Classes 2, 4, 6 and 10 are plural classes. Other classes (14 and 15) are used exclusively for non-count nouns. There are no classes 8, 12 or 13 in Zulu.



# Abbreviations

A	Active personal pronominal affix (followed by 1, 2, 3, 0 – referring to the 1st, 2nd, 3rd person and zero)
ABL	Ablative
ABN	Abstract noun
ABS	Absolutive
ACC	Accusative
ACN	Action noun
ACT	Active
ADE	Adessive
ADJ	Adjective
ADR	Adverb
ADV	Adverse
AFF	Affective
AG	Agent
ALL	Allative
AMB	Ambulative
ANP	Anaphoric
ANX	Action nominal affix
APP	Applicative
APS	Antipassive
ATTR	Attribute
AUG	Augmentative
AV	Agent/Actor voice
+ ATR	Advanced (retracted) tongue root
– ATR	Non-advanced (retracted) tongue root
BEN	Benefactive
BFR	Back-formation
BLN	Blending
CAL	Consonant alternation
CAU	Causative
CFX	Confix
CHN	Change
CLA	Classifier

CLS	Classifying particle
CMP	Comparative
CMT	Comitative
CNV	Conversion
COL	Collective
COM	Companionate
CON	Continuous
COO	Coordination
COU	Countable
CPL	Completion
CPN	Compound
CRX	Circumfix
CV	Consonant-vowel (indicates reduplication)
DEF	Definite marker
DET	Determinative
DIM	Diminutive
DIR	Directional
DIS	Discontinuity
DIV	Diversative
DNL	Denominal
DST	Distal
DTB	Distributive
DUA	Dual
DUR	Durative
EMP	Empty particle
ERG	Ergative
ESS	Essive
EVN	Event
EXC	Excessive
EXP	Experiential
F	Feminine
FCT	Factitive
FEA	Feature
FRE	Frequentative
FUT	Future
GEN	Genitive
GND	Gender
GRC	Generic
GRN	Gerund
ICL	Inclusive
ICP	Inceptive

---

IDF	Indefinite
IFL	Inflection
IFX	Infix
IMP	Imperative
INC	Inchoative
IND	Indicative
INF	Infinitive
INS	Instrumental
INT	Intensive
INTR	Intransitive
IO	Indirect object
IPF	Imperfective
IPS	Impersonal
IRR	Irrealis
ITE	Iterative
ITX	Interfix
LNK	Linking element
LOC	Locative
M	Masculine
MOM	Punctual, momentaneous or single-event verb
MRK	Marker
MS	Morphotactic separator (connector in finite verb forms)
MUT	Mutual
N	Noun
NEG	Negative, negation
NIN	Noun incorporation
NMR	Nominalizer
NOM	Nominative
NTR	Neuter gender
O	Direct object pronoun
OBJ	Object
OBL	Oblique
PAS	Passive
PAT	Patient
PEF	Personifier
PER	Personal article
PFV	Perfective
PK	Personal knowledge
PL	Plural
PLC	Place

---

POS	Possessive
PRG	Progressive
PRI	Privative
PRS	Present
PRX	Prefix
PST	Past
PTC	Participle
QNT	Quantity
QUO	Quotation
QSN	Question
RAP	Root-and-pattern
RCP	Reciprocal
RDP	Reduplication
REA	Realis
REL	Relational
REV	Reversative
RFL	Reflexive
S	Stative personal pronominal affix (followed by 1, 2, 3, 0 – referring to the 1st, 2nd, 3rd person and zero)
SBJ	Subject
SFX	Suffix
SG	Singular
SGT	Singulative
SIM	Simultaneity
SOC	Sociative
SPE	Specific
SS	Sentence suffix
ST	First part of stem, probably part of a compound of unclear meaning, preceding pronominal prefixes
STA	Stative
STM	Stem
STR	Stress
STT	State
SUP	Supine morpheme
TAM	Tense/Aspect/Mood marker
TH	Thematic prefix/suffix
THM	Theme
TI	Time
TOP	Tone/pitch
TOT	Totality

---

TR	Transitive
TRN	Translative
V	Verb
VAL	Vowel alternation
VL	Verbal
VR	Verbalizer
1	1st person
1>2	Verbal grammatical first to second person 'I>you'
2	2nd person
3	3rd person
3o	3rd person direct object
3>3	Verbal grammatical 3rd to 3rd person 'she>him'
<sub>c</sub> + Suffix	the preceding morpheme drops its vowel
<sub>v</sub> + Suffix	the preceding morpheme retains its vowel

# Contents

<i>List of figures</i>	viii
<i>List of tables</i>	ix
<i>Acknowledgements</i>	xiv
<i>Notes on language-specific symbols</i>	xvii
<i>List of abbreviations</i>	xix
<b>Introduction</b>	1
Antecedents	2
<i>Word-Formation in the World's Languages</i> : purpose, method and scope	4
Purpose	4
Method	6
Scope	9
The structure of the book	14
<b>PART I. The field of word-formation</b>	17
<b>1 The scope of word-formation</b>	19
1.1 Inflection vs. derivation	19
1.1.1 General	19
1.1.2 The scope of word-formation	24
1.1.3 Ambiguities in the supercategory of quantity	26
1.1.4 Summary	35
<b>2 Word-formation vs syntax</b>	36
2.1 Different notions of word	36
2.2 The status of compounds	41
2.3 Compounding and noun incorporation	43
2.4 Summary	47

---

<b>PART II. Cross-linguistic analysis</b>	<b>49</b>
<b>3 Word-formation processes combining free morphemes</b>	<b>51</b>
3.1 Compounding	51
3.1.1 Types of compounds	53
3.1.2 Recursiveness in compounding	93
3.1.3 Word-formation base modification in compounding	99
3.2 Reduplication	101
3.2.1 Types of reduplication	107
3.3 Blending	131
3.4 Summary	134
<b>4 Word-formation processes with bound morphemes</b>	<b>135</b>
4.1 Affixation	135
4.1.1 Suffixation and prefixation	138
4.1.2 One-to-many relation in affixation	168
4.1.3 Many-to-one relation in affixation	183
4.1.4 Suffixation, prefixation and word order	195
4.2 Minor types of affixation	197
4.2.1 Infixation	198
4.2.2 Prefixal-suffixal derivation	203
4.2.3 Circumfixation	208
4.2.4 Prefixal-infixal and infixal-suffixal derivation	210
4.3 Summary	212
<b>5 Word-formation without addition of derivational material and subtractive word-formation</b>	<b>213</b>
5.1 Word-formation without addition of derivational material	213
5.1.1 Conversion	213
5.1.2 Stress	225
5.1.3 Tone/pitch	226
5.1.4 Word-formation by internal modification	229
5.2 Subtractive word-formation processes	234
5.2.1 Back-formation	234
5.3 Summary	236

<b>6. An onomasiological description</b>	237
6.1 Introduction	237
6.2 Nominal categories	240
6.2.1 Agents	240
6.2.2 Patients	248
6.2.3 Instrumentals	253
6.2.4 Locatives	256
6.2.5 Gender in animate beings	260
6.3 Evaluative categories	264
6.3.1 Augmentatives and diminutives	264
6.3.2 Phonetic iconicity	270
6.3.3 Word-classes	273
6.4 Verbal categories	275
6.4.1 Causatives	275
6.4.2 Transitivity	280
6.4.3 Intransitivity	286
6.4.4 Iterativity and/or intensification	289
6.5 Word-class changing categories	293
6.5.1 Action nouns	293
6.5.2 Abstract nouns	297
6.6 Summary	303
<b>7 Results and discussion</b>	304
7.1 Introduction	304
7.1.1. Method	305
7.2 Results and discussion	308
7.3 Summary	326
<b>Epilogue</b>	327
<i>Appendix I. The distribution of word-formation processes in the study sample</i>	330
<i>Appendix II. The questionnaire</i>	332
<i>References</i>	340
<i>Author index</i>	352
<i>Language index</i>	355
<i>Subject index</i>	360



# Figures

0.1	Geographic distribution of the basic sample languages (seventy languages) as in Haspelmath, Dryer, Gil and Comrie (2005)	10
0.2	Geographic distribution of the study sample languages (fifty-five languages) as in Haspelmath, Dryer, Gil and Comrie (2005)	12
7.1	Cross-linguistic use of word-formation processes in the study sample (absolute values with respect to fifty-five languages)	309
7.2	Cross-linguistic use of word-formation processes (percentages)	309
7.3	Language families by word-formation process occurrence or not	314
7.4	Presence (represented as .1) and absence (represented as .0) of word-formation processes in the study sample.	315
7.5	Associations between language families and word-formation processes	316
7.6	Associations between language families and types of prefixation	317
7.7	Associations between language families and types of suffixation	318
7.8	Associations between language families and types of compounding	319
7.9	Associations between language families and types of reduplication	320
7.10	Word-formation relevance of semantic categories in the study sample (absolute values with respect to fifty-five languages)	323
7.11	Word-formation relevance of semantic categories (percentages)	324
7.12	Percentage of occurrence of semantic categories with respect to the most frequently used word-formation processes in the languages sampled	326