

Nominal Compound Acquisition

Language Acquisition & Language Disorders

61

Edited by

Wolfgang U. Dressler

F. Nihan Ketrez

Marianne Kilani-Schoch

John Benjamins Publishing Company

This book offers a systematic study of the emergence and early development of compound nouns in first language acquisition from a cross-linguistic and typological perspective. The language sample is both genealogically and typologically diversified, ranging from languages rich in compounds, such as German, Saami, Estonian and Finnish, to languages poor in compounds, such as French. Some of them differ in compound richness according to genres of adult-directed speech in contrast to child-directed speech and thus also child speech, like Russian, Lithuanian and especially Greek. Differences in the delimitation and transition between compounds and phrases and in the distribution of subtypes of compounds in these languages involve great typological variety and thus different tasks for children acquiring them. The eleven languages investigated in the volume and the common methodology of longitudinal collection of spontaneous speech data concerning the interaction between children and their caretakers or peers, supplemented by lexical typology as a new means of cross-linguistic comparison of language acquisition, allow new generalizations and make the volume a unique contribution.

ISBN 978 90 272 5324 8



9 789027 253248

JOHN BENJAMINS PUBLISHING COMPANY

LALD
61

DRESSLER, KETREZ & KILANI-SCH

Normal Compound Acquisition



Nominal Compound Acquisition

Edited by

Wolfgang U. Dressler

University of Vienna

F. Nihan Ketrez

Istanbul Bilgi University

Marianne Kilani-Schoch

University of Lausanne

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/lald.61

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2017031636 (PRINT) / 2017046277 (E-BOOK)

ISBN 978 90 272 5324 8 (HB)

ISBN 978 90 272 6497 8 (E-BOOK)

© 2017 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · <https://benjamins.com>

Nominal Compound Acquisition

Language Acquisition and Language Disorders (LALD)

ISSN 2213-428X

Volumes in this series provide a forum for research contributing to theories of language acquisition (first and second, child and adult), language learnability, language attrition and language disorders.

For an overview of all books published in this series, please see
<http://benjamins.com/catalog/lald>

Series Editors

Roumyana Slabakova
University of Southampton
and University of Iowa

Lydia White
McGill University

Editorial Board

Kamil Ud Deen
University of Hawaii at Manoa

Katherine Demuth
Macquarie University

Naama Friedmann
Tel Aviv University

Heather Goad
McGill University

Barbara Höhle
University of Potsdam

Nina Hyams
University of California at Los Angeles

Jürgen M. Meisel
University of Calgary

Mabel Rice
University of Kansas

Luigi Rizzi
University of Siena

Petra Schulz
Johann Wolfgang Goethe-Universität

Bonnie D. Schwartz
University of Hawaii at Manoa

Antonella Sorace
University of Edinburgh

Ianthi Maria Tsimpli
University of Cambridge

Volume 61

Nominal Compound Acquisition

Edited by Wolfgang U. Dressler, F. Nihan Ketrez and Marianne Kilani-Schoch

Abbreviations

A, ADJ	adjective	GSC	Grammatical Smixut
ABL	ablative		Construction
ACC	accusative	HSES	high socio-economic
ADESS	adessive		status
ADS	adult directed-speech	ILL, Ill	illative
ADV, ADV	adverb	IMP	imperative
ALL	allative	INCH	inchoative
AN, AN	adjective noun	INF	infinitive
Anc.	Ancient (Greek)	INTERF	interfix
ASC	Adjacency Smixut	Isuf	inflectional suffix
	Compound	Lem, LEM	lemma
ATTR	attributive	LI	language-impaired
CAR	caritive	LOC, Loc	locative
CAUS	causative	LSES	low socio-economic
CDS	child directed-speech		status
CHI	child	m	mean value
CLIT	clitic, enclitic particle	M, MASC	masculine
CM	compound marker	MCT	main caretaker
COM	comitative	MLU	mean length of utterance
COMP, COMP	compound	Mod.	Modern (Greek)
CS	child speech	Mot	mother
DEM	demonstrative	N	noun
DERIV	derivative element	N	number of raw data
DIM, DIM	diminutive	N.A.	non available
Dsuf	derivational suffix	NEUT	neuter
ELA	elative	NN, NN	noun noun
Eng.	English	N^N	bound noun compound
F, FEM	feminine	NOM, Nom	nominative
FCC	Free Compound	NR	nominalizer
	Construction	NUM, Num	numeral
FILL	filler	O, ONOM	onomatopoeia
FREQ	frequentative	P, PREP	preposition
GA	genitive-accusative	PART	participle
GEN, Gen	genitive	PARTIT	partitive

PAST	past tense	REFL	reflexive
PFV	perfective	SES	socio-economic status
PL, Pl	plural	SG, Sg	singular
PN	preposition noun	SLI	specific language impairment
POSS	possessive	SUBIT	subitive
PP, pp	past participle	Tok	token
PREP, prep	preposition	TTR	type-token ratio
PRO	pronoun	V	verb
PROG	progressive	VN, vn	verb noun
PROP	proper noun	VR	verbalizer
PTL	particle		

Table of contents

Abbreviations	vii
Introduction	1
<i>Wolfgang U. Dressler, F. Nihan Ketrez, Marianne Kilani-Schoch & Ursula Stephany</i>	
CHAPTER 1	
Emergence and early development of German compounds	19
<i>Katharina Korecky-Kröll, Sabine Sommer-Lolei & Wolfgang U. Dressler</i>	
CHAPTER 2	
Compound nouns in Danish child language	39
<i>Laila Kjærbaek & Hans Basbøll</i>	
CHAPTER 3	
Acquisition of nominal compounds in Russian	63
<i>Victoria V. Kazakovskaya</i>	
CHAPTER 4	
Early development of compounds in two French children's corpora	91
<i>Marianne Kilani-Schoch</i>	
CHAPTER 5	
Compounding in early Greek language acquisition	119
<i>Ursula Stephany & Evangelia Thomadaki</i>	
CHAPTER 6	
The early production of compounds in Lithuanian	145
<i>Ineta Dabašinskienė & Laura Kamandulytė-Merfeldienė</i>	
CHAPTER 7	
Acquisition of noun compounds in Estonian	165
<i>Reili Argus</i>	
CHAPTER 8	
Acquisition of compound nouns in Finnish	191
<i>Klaus Laalo</i>	

CHAPTER 9	
The acquisition of compound nouns in North Saami	209
<i>Johanna Johansen Ijäs</i>	
CHAPTER 10	
The emergence of nominal compounds in Turkish: A case study on structural simplicity vs. input frequency	233
<i>F. Nihan Ketrez</i>	
CHAPTER 11	
Compounding in early child speech: Hebrew peer talk 2–8	251
<i>Dorit Ravid & Noa Assulin Tzabar</i>	
CHAPTER 12	
Contrastive lexical typology of German and Greek child speech and child-directed speech	275
<i>Ursula Stephany</i>	
CHAPTER 13	
Discussion and outlook	287
<i>Wolfgang U. Dressler, F. Nihan Ketrez & Marianne Kilani-Schoch</i>	
Index	307

Introduction

Wolfgang U. Dressler, F. Nihan Ketrez, Marianne Kilani-Schoch & Ursula Stephany

University of Vienna / Istanbul Bilgi University / University of Lausanne /
University of Cologne

1. Compounding

1.1 Aim of the volume

Compounding appears to be the most wide-spread technique for word formation in the languages of the world (Booij 2013:258; Dressler 2006:23). In a first approximation, compounds may be defined as “the juxtaposition of two words to form a new one (Bauer 2009)” (Booij 2013:258) or as “grammatical combinations of words, that is of lexical items or lexemes, to form new words” (Dressler 2006:24), whereas the combination of grammatical words, as in English *therefore*, *herein*, is marginal. This basic property assigns compounding a transitional position between morphology and syntax in the continuum of inflection – derivation – compounding – phrasal syntax (Dressler, Mayerthaler, Panagl & Wurzel 1987:4–7; Ralli 2013:244 et passim) and has consequences for morphology-theoretic modeling (Dressler 1989), language typology (Dressler 2010, among others) and acquisition. This volume aims at studying the emergence and early development of compounding in first language acquisition from a cross-linguistic and typological perspective with a focus on compound nouns.

It complements the volumes on morphology in first language acquisition edited by Bittner, Dressler and Kilani-Schoch (2003) on the development of verb inflection, by Stephany and Voeikova (2009) on the development of nominal inflection, by Savickienė and Dressler (2007) on the acquisition of diminutives, and by Tribushinina, Voeikova and Noccetti (2015) on semantics and morphology of early adjectives in first language acquisition, all resulting from the international Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition (cf. also Xanthos et al. 2011), which focuses on longitudinal studies of early first language acquisition in connection with parental input (child-directed speech).

1.2 Compounding preferences

The focus of this volume is on the acquisition of nominal compounding, because in most languages there seems to exist a preference for forming compounds which prototypically consist of two lexical elements (Libben & Jarema 2006) by employing nouns as one or both parts of a nominal compound. According to Dressler (2005) other general compounding preferences are

1. morphotactic transparency, i.e. easy formal decomposition of compounds;
2. morphosemantic transparency, i.e. easy recoverability of the meaning of a compound's parts from the holistic meaning of the compound and their relation;
3. iconicity of head-non-head relations so that the formal head of a compound is also its semantic head: for instance, the plural of English *corner stone* is formed from the head *stone*, in other words, preferentially only a head is inflected and semantically a *corner stone* is a stone; the same holds with reversed roles of the compound elements for the head of *stone corner*;
4. a preference for a compound to have a single head, as in the previous examples of subordinative compounds, whereas double headedness of coordinative compounds, as in *prince-consort*, is dispreferred;
5. a further preference for the non-final parts of compounds, which tend to represent lexical base forms or lexemes rather than smaller forms (roots) or longer forms (inflected forms, phrases, clauses or even sentences);
6. preferentially binary status of compounds, i.e. consisting of two parts. This may either mean that they are not recursive (Štekauer, Valera & Körtvélyessy 2012: 93–98) in that they consist of only two elements or that they have a hierarchical binary structure. Thus the subordinative left-branching compound *three star general* consists of the two constituents *general* (head) and *three+star* (non-head), the latter of which in turn contains a non-head and a head so that the overall structure of the construction is [[[three][star]][general]].

All of these preferences are reflected in the relative type and token frequencies of compounds and their respective degree of productivity in a given language. But the question which subtypes of compounds are more or less productive and/or frequent can only be answered in a language-specific way. For example, nominal compounds consisting of a verb and a noun such as French *porte+avion* 'aircraft-carrier' are much more essential in Romance languages (cf. Bauer 2011; Villoing 2009) than in the Germanic and other languages studied in this volume.

1.3 Compounding vs. phrasal syntax

Whereas the domain of transition between compounding and derivation (e.g., suffixoids) is relatively unimportant for the early phases of acquisition investigated in

this volume, patterns belonging to the transitional domain between phrasal syntax and compounding may emerge very early in child speech.

Noun phrases carrying a fixed lexicalized and partially opaque meaning, e.g., *Third World*, are to be considered as multi-lexical words. They may also contain a syntactic indicator as in the French multilexical words *clair de lune* ‘moonlight’ or *pomme de terre* ‘potato’ (called “synapsies” by Benveniste 1966/1974: 172). They are sometimes considered to be loose compounds (Dressler 2006: 28; Scalise 1984) or phrasal compounds (Booij 2010; Ralli 2013). Ralli (2013: 244) proposes “a continuum of morphologically complex nominal formations” with “typical one-word compounds placed on one of its ends, while the other end contains syntactically built noun phrases sharing with compounds the same grammatical categories.” The features of lexicalization and an idiomatic, partially opaque meaning are fundamental for them to qualify as phrasal compounds rather than syntactic phrases (Dressler 2006: 28; Ralli 2013: 246). In spite of the fact that such formations carry a phrasal rather than a compound accent, the following characteristics distinguish them from syntactic phrases and make them resemble ‘morphological compounds’ (Booij 2010): The order of constituents cannot be reversed even in languages where order reversal is possible in syntactic phrases (e.g., Greek *atomikí vómva* ‘atomic bomb’ vs. *?vómva atomikí* ‘?personal bomb’, but *meyáli vómva* and *vómva meyáli* ‘big bomb’). No other element can be inserted between their constituents and neither the dependent element nor the head can be independently modified (e.g., **Third nice World*). Finally in many languages the order of elements in such multilexical words is different from that of semantically comparable compounds (cf. Ralli 2013: 246–247 for Greek).

Phrasal compounds may be ordered on a scale depending on whether they are to a higher degree governed by syntactic or morphological characteristics (see Stephany & Thomadaki 2016: 23). Thus, Greek *çimós portokáli* (lit. juice orange) ‘orange juice’ or *sálsa domáta* (lit. sauce tomato) ‘tomato sauce’ may be considered to be special NPs and not to belong to compounding (Ralli 2013: 267). The reason is that, in spite of the fact that “their meaning does not completely originate from the meaning of the constituent parts”, they “can accept insertion of a parenthetical element” and “the order of their constituents may be reversed” (Ralli 2013: 258). But in such continua there may be no parallel asymmetric and transitive order of features distinguishing syntax and morphology so that the same properties may not have the same rank in each continuum.

1.4 Classification of compounds

All the chapters contained in this volume more or less directly refer to one of the following five classifications of compounds:

A first classification of compounds based on Scalise (1992), Scalise and Vogel (2010), and Štekauer, Valera and Körtvélyessy (2012:73–101) distinguishes

- Subordinative endocentric compounds, e.g., English *love story*, where, within the compound, the non-head *love* is subordinated to the head *story*. This is the default construction in many languages. Although many languages allow both orders of head and non-head, right-headed compounds are typical of a larger number of languages than left-headed ones.
- Subordinative exocentric (or bahuvrihi) compounds, e.g., *pickpocket*, where the semantic head (a person) is outside the compound, but the verb (as a sort of subhead) dominates the noun. This is a rare type in many languages, and so are the following two.
- Coordinative (or appositional) endocentric compounds, e.g., *speaker-hearer*, i.e. a speaker and a hearer at the same time, with two semantic heads.
- Coordinative exocentric compounds, e.g., *nature-nurture (debate)* with an external head (*debate*), whereas the compound itself consists of two coordinate nonheads.

A second classification identifies the word classes (and their subcategories) to which the members of the compound belong in their autonomous use: i.e. although a *sit-in* is a noun, *sit* is a verb and *in* a preposition.

The third one distinguishes transparent compounds consisting of two simply concatenated elements from morphotactically less transparent ones where the final part of the first element is truncated or an interfix (or a meaningless linking element) is inserted. Both opacifying operations occur in German *Gebirg+s+bach* ‘mountain creek’, with the first element truncated from *Gebirge* ‘mountain’ and an interfix -s- having been inserted between the two constituents.

Fourth, the prefinal lexical member of a compound may be identical to one of the following: the inflectional base, some other of its inflected forms, the stem, or the root of a word. In this way, word-based, stem-based and root-based compounds are distinguished from each other. Thus, the first element *Gebirg-* of German *Gebirgsbach* may also be identified as the root of the noun *Gebirge*; the first element of Lithuanian *nakti+piečiai* ‘night snack’ is the stem of NOM SG *nakt-i+s* ‘night’.

Fifth, in synthetic compounds a noun phrase or verb phrase constitutes the non-head, e.g., English *three+star+general* and *street+sweep+er*.

The contents of this volume are limited to the acquisition of grammatical compound patterns, in contrast to superficially similar extragrammatic blends, such as *smog*, which— similar to coordinative compounds— combines the meaning of *smoke* and *fog*. However, the morphotactics of such forms differ strikingly from

grammatical compound formation, even in case of truncation of the first element, as in German *Fried+hof* ‘cemetery’ from *Friede(n)* ‘peace’ and *Hof* ‘court’. But since compound amalgams occur in early phases of compound acquisition (see Section 2.2.1), their superficial formal similarity to blending will require investigation in early child speech. This topic has not yet received much attention in the literature (cf. Jaeger 2005) and will be investigated in this volume (see Chapter 1).

2. Methodology and theoretical approach

2.1 Longitudinal study of spontaneous speech

Research on the acquisition of compounds started with diary observations (a recent example with rich data is Rainer 2010). However, experimental studies are much more common (Berman 2009; Clark 1993/2009; Nicoladis 2006). Since longitudinal case studies (Argus & Kazakovskaya 2013; Argus, Johansen Ijäs & Laalo 2014; Becker 1994; Bilev 1985; Clark 1993/2009; Dressler, Lettner & Korecky-Kröll 2010; Mellenius 1997; Nicoladis 1999) are much rarer than studies based on transversal tests and are mainly about single languages, the present volume intends to fill this lacuna. In order to reach a high ecological validity by tracing the developmental paths of the early acquisition of languages of different types, the investigation of parental input and child output must be as systematic as possible. Studies on the acquisition of compounds have so far rarely included a discussion of child-directed speech (CDS). Since CDS differs greatly from adult-directed speech (ADS), particularly so in the early phases of acquisition, it does not seem reasonable to compare acquisition to the respective target language or to electronic corpora of ADS (and even often only written speech), as has become increasingly evident in studies on language acquisition (Clark 2010; Kilani-Schoch, Balčiuniene, Korecky-Kröll, Laaha & Dressler 2009; MacWhinney 2004; Ravid et al. 2008; Stephany 1985, to appear). Moreover the intermediate levels of intake and uptake (cf. Harris 1992; Mitchell, Myles & Marsden 2013) have to be taken into consideration. Intake refers to what the young child actually can be considered to perceive or notice of the input, for example in early phases often only the final part of a compound according to positional and prosodic salience (recency effect) or its stressed syllable or a combination of the initial and the final part (bathtub effect, i.e. combination of primacy and recency effect). Uptake is what children abstract from the intake and store in their implicit memory. At this point they may even construct patterns (in terms of constructivist self-organisation, cf. Bittner, Dressler & Kilani-Schoch 2003: xviii) which may differ from patterns reflected in the input, as evidenced by overregularizations and blind-alley developments (see

Section 2.2.3 below). Intake and uptake are, of course, not directly observable but rather constitute black boxes whose contents can be reconstructed only partially by the analyst.

Our Crosslinguistic Project and the present volume accordingly are based on longitudinal studies of spontaneous interaction between children and their caretakers, thus focusing on the development of children's output in relation to their input. This contrasts with most studies of compounding which rely on tests and thus on children's output in usually rather formal situations that demand more language awareness than is needed in spontaneous speech. It is not only impossible for such studies to compare output with input, but they also usually only adopt a secondary developmental perspective by comparing distant age brackets which most often concern different groups of children. Therefore, the results concerning the developmental phases of the children tested have only limited ecological validity.

2.2 Distinctions between phases

In the study of longitudinal data on interactive spontaneous speech we differentiate the following developmental phases:

2.2.1 Premorphology

In the premorphological phase, i.e. the phase before the detection of morphological decomposition and composition, children produce only isolated rote-learned compounds without recurrence of their lexical members either as autonomous words or parts of other compounds (thus not in word families), i.e. as input compounds not yet analysed in the intake and uptake. As a consequence of the unawareness of morphological boundaries and probably also of prosodic constraints on word length, these productions can include amalgams which may neither contain the end of the first member nor the beginning of the second (plus possible further deformations), e.g., German *Bausen* < *Bau+stein-e* 'lit. construction stones' (Jan 1;5). If it were just a question of prosodic shortening, then a child aware of morphological structure should reduce compounds to one of its lexical members. We have not, however, found clear instances of this during the premorphological phase. Amalgams which start with the beginning of the first member and finish with the end of the last member resemble similar early reductions of diminutives, such as Italian *pinna* < *panc+ina* 'belly-DIM' (Noccetti, de Marco, Tonelli & Dressler 2007: 134).

2.2.2 Protomorphology

The protomorphological phase starts with the child's detection of morphology as a means of decomposing and composing meaning and form, even if only partially