

The Acquisition of Derivational Morphology

A cross-linguistic perspective

Language Acquisition & Language Disorders
66

Edited by
Veronika Mattes
Sabine Sommer-Lolei
Katharina Korecky-Kröll
Wolfgang U. Dressler

John Benjamins Publishing Company

The Acquisition of Derivational Morphology

Language Acquisition and Language Disorders (LALD)

ISSN 0925-0123

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 benjamins.com/catalog/lald

Series Editors

Roumyana Slabakova
University of Southampton and NTNU,
Norwegian University of Science and Technology

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 66

The Acquisition of Derivational Morphology. A cross-linguistic perspective
Edited by Veronika Mattes, Sabine Sommer-Lolei, Katharina Korecky-Kröll
and Wolfgang U. Dressler

The Acquisition of Derivational Morphology

A cross-linguistic perspective

Edited by

Veronika Mattes

University of Graz

Sabine Sommer-Lolei

Katharina Korecky-Kröll

Wolfgang U. Dressler

University of Vienna

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.66

Cataloging-in-Publication Data available from Library of Congress:

LCCN 2021034420 (PRINT) / 2021034421 (E-BOOK)

ISBN 978 90 272 0982 5 (HB)

ISBN 978 90 272 5888 5 (E-BOOK)

© 2021 – 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>

Table of contents

List of abbreviations	VII
Acknowledgements	IX
CHAPTER 1	
Introduction	1
<i>Wolfgang U. Dressler, Veronika Mattes and Laila Kjærbaek</i>	
CHAPTER 2	
The development of derivation in early Greek first language acquisition	21
<i>Ursula Stephany</i>	
CHAPTER 3	
Derivational patterns in spontaneous data of French-speaking parent-child interactions before age three	53
<i>Marianne Kilani-Schoch and Aris Xanthos</i>	
CHAPTER 4	
Emergence and early development of derivatives in Danish child language	85
<i>Laila Kjærbaek and Hans Basbøll</i>	
CHAPTER 5	
Early phases of development of German derivational morphology	109
<i>Sabine Sommer-Lolei, Veronika Mattes, Katharina Korecky-Kröll and Wolfgang U. Dressler</i>	
CHAPTER 6	
Derivational morphology in Croatian child language	141
<i>Gordana Hržica</i>	
CHAPTER 7	
Acquisition of derivational morphology in Russian	169
<i>Victoria V. Kazakovskaya and Maria D. Voeikova</i>	

CHAPTER 8	
The acquisition of the Lithuanian derivational system	197
<i>Laura Kamandulytė-Merfeldienė, Ingrida Balčiūnienė and Ineta Dabašinskienė</i>	
CHAPTER 9	
Acquisition of noun and verb derivation in Estonian	217
<i>Reili Argus</i>	
CHAPTER 10	
Derivation in Finnish child speech and child-directed speech	237
<i>Klaus Laalo</i>	
CHAPTER 11	
Noun and verb derivations in early Turkish child and child-directed speech	263
<i>F. Nihan Ketrez and Ayhan Aksu-Koç</i>	
CHAPTER 12	
Conclusions	289
<i>Veronika Mattes and Wolfgang U. Dressler</i>	
Subject index	305

List of abbreviations

1	first person	ind.	indicative
2	second person	INF, inf.	infinitive
3	third person	INFL	inflection / inflectional
ABL	ablative	INSTR, instr.	instrument(al)/ instrumentative
ACC	accusative		
add	addition	INTERF	interfix
ADJ, ADJ	adjective	INTERJ	interjection
ADJN	adjective-noun compound	INTR	intransitive
ADS	adult-directed speech	IPFV	imperfective
ADV, ADV	adverb(ial)	ITER	iterative
AG	Ancient Greek	lem.	lemmas
ALL	allative	lit.	literally
AUG	augment(ation)/ augmentative	loc.	local
		LOC	locative/location
AUX	auxiliary	m	male
CAUS	causative	M	masculine / male
CDS	child-directed speech	MEDIOPASS	mediopassive
CHI	(target) child	MG	Modern Greek
CI	confidence interval	MLU	mean length of utterance
CM	compound marker	mod.	modification
COLL	collective	MOT	mother
COND	conditional	MPC	mini-paradigm criterion
CPD	compound	N, N	noun
CS	child speech	NEG	negation
DAT	dative	NEUT	neuter
DER, DERIV,	derivational/	NN	noun-noun compound
DERIV	derivation/derived	N-N	noun to noun derivation
DIM, DIM	diminutive	NOM	nominative
DIR	directive	NP	noun phrase
em.	emergence	NPST	nonpast
f	female	NSUFF	nominal suffix
fem. / F	feminine	NUM, NUM	numeral/number
FACT	factitive	N-V	noun to verb conversion
FUT	future	OBJ	object
GEN	genitive	omiss	omission
IMP	imperative	ONOM	onomatopoeic / onomatopoeia

PARTIT	partitive	Q	question particle/marker
PASS	passive	rec.	recording(s), recorded
PFV	perfective	res.	result / resultative
phon	phonotactic	REVERS	reversative
PL	plural	SES	socioeconomic status
pm	premorpheme	SG, sg.	singular
POSS	possessive	SLI	Specific Language Impairment
PP, PP	perfect participle	SMG	Standard Modern Greek
PPres	present participle	SBJ	subject
PREF, PREF	prefix	SUFF, SUFF	suffix
PRIV	privative	TAM	tense-aspect-mood
PREP	preposition	TEMP	temporal
prod.	productive	tok.	tokens
PRON	pronoun	typ.	types
PRS, pres.	present	TR	transitive / transitivity
PST	past	V, v	verb
PTCP	participle	VV	verb-verb compound
PTL	particle	V-N	verb to noun derivation

Acknowledgements

This book evolved in the framework of the international Cross-linguistic Project on Pre- and Protomorphology, founded and guided by W.U. Dressler. We are very grateful to the members of this group for their fruitful discussions and intensive collaboration on the topic of acquisition of derivation. They accompanied the process of editing this volume and helped us to understand the issues more clearly by their valuable feedback on the various stages which were regularly presented to the group.

We would like to thank the contributors to this volume for their thorough work, for cooperating closely and reliably, for their suggestions and comments during all phases of the development of this volume. Furthermore, we are grateful for their critical reading and commenting of the Introduction and Conclusions chapters.

We would also like to express our gratitude to the reviewers of the chapters who agreed to share their time and their expertise with us, for their useful suggestions.

We appreciate the pleasant, productive and helpful cooperation with the LALD series editors.

Veronika Mattes
Sabine Sommer-Lolei
Katharina Korecky-Kröll
Wolfgang U. Dressler

Introduction

Wolfgang U. Dressler, Veronika Mattes and Laila Kjærbæk

University of Vienna / University of Graz / University of Southern Denmark

1. Aim of the volume

This is the first cross-linguistic volume on the early acquisition of derivational morphology. It aims at studying the emergence and early development of patterns in derivational morphology for first language acquisition from a cross-linguistic and typological perspective. This will allow us to propose implications for theoretical morphology and for acquisition studies in general.

This volume complements a series of volumes in first language morphology acquisition: on the development of verb inflection edited by Bittner, Dressler & Kilani-Schoch (2003), by Stephany & Voeikova (2009) on the development of nominal inflection, by Savickienė & Dressler (2007) on the acquisition of diminutives, by Tribushinina, Voeikova & Noccetti (2015) on semantics and morphology of early adjectives and by Dressler, Ketrez & Kilani-Schoch (2017) on nominal compound acquisition. They all result from the international Cross-linguistic Project on Pre- and Protomorphology in Language Acquisition (see also Xanthos et al. 2011), which consists of longitudinal studies of early first language acquisition in interaction with parental caregivers (input or child-directed speech, henceforth CDS). Our volume focuses on that part of morphology which has not been described in the previous volumes, namely derivational morphology, and that had been (with the exception of diminutives) rather understudied in research on language acquisition so far, as opposed to inflection and compounding.

2. Derivational morphology

2.1 Delimitations of the domain of derivational morphology

Derivational morphology is part of the lexicon and a component of morphology. It has five transition areas in the form of gradual continua which share properties both of derivational morphology and one of the following domains: compounding,

clitics, inflection, the lexicon of simplex words and extragrammatical morphology. The latter has different properties than grammatical morphology, as in different types of shortenings, such as acronyms or clippings, and, partially, in onomastic morphology (see Bybee 1985, 1995; Doleschal & Thornton 2001; Dressler, Dziubalska-Kořaczyk, Gagarina & Kilani-Schoch 2014; Dressler, Mayerthaler, Panagl & Wurzel 1987: 4–7; Haspelmath 1996; Spencer & Luís 2012).

The core of derivational morphology is constituted by productive prototypical rules (Dressler 2007) and by the word families they motivate morphosemantically and morphotactically, but the domain of derivation also includes many patterns that are weakly productive or unproductive and weakly transparent or non-transparent (fully opaque), and thus in transition to the lexicon of simplex words.

The transition to inflection is represented by diminutives, which are already of great importance in the early phases of language acquisition. These have already been dealt with abundantly in Savickienė & Dressler (2007). This is the first reason why we deal only marginally with diminutives in this volume. Other reasons are that their dominant function in child language is pragmatic, in contrast to all other derivational patterns. This, and their transitional properties to inflection, leads to several contrasts with other derivational patterns, notably in that they are much more word-class maintaining and semantically empty. Moreover, word-class maintaining denominal and even deadjectival diminutives are frequent or very frequent in our target languages, whereas deverbal diminutive verbs are very rare (e.g. German *lieb+eln* ‘to love a bit, not seriously’). The inclusion of diminutives would thus greatly distort the picture when comparing nominal with verbal and adjectival derivation in language acquisition. We will document the exceptional character of diminutive acquisition in each case of its occurrence.

Another case of transition to inflection is represented by so-called postfixes in Russian. They have the non-prototypical position after all inflectional verb suffixes, like sometimes diminutive suffixes after plural suffixes (e.g. in German *Kind-er-chen* ‘child-PL-DIM’), but like derivational suffixes, often change the meaning of the non-postfixed verb in a morphosemantically opaque manner (Say 2005), as in *pytat* ‘to torture’ vs. *pytat’sja* ‘to try’ (Sergey Say, p.c.). In contrast to diminutive suffixes, however, they occur after any verb form and thus also resemble leaners (they share with clitics that they are unstressed and therefore lean prosodically on a preceding or following stressed word, but in contrast to clitics which have different positions in a sentence according to the syntactic construction, postfixes have a fixed position after the verb). In several languages, infinitives may also have some derivational properties, and participles are often more derivational than inflectional.

The domain of transition between derivation and clitics is relatively unimportant for the early phases of acquisition (see above also on postfixation). The transition between compounding and derivation is also unimportant as far as affixoids

are concerned (e.g. in German *hoffnung-s-voll*, English *hope-ful*), whereas interfixes (e.g. *-s-* in the German example) have been dealt with already in our compound volume (Dressler, Ketrez & Kilani-Schoch 2017). But prefixed and particle verbs with prefixes and particles that are identical (or, more cautiously, semantically related and homophonous) to prepositions or adverbs are relevant in this respect. Complex verbs of this kind thus share several features with compounds. In compounding languages, where compounds emerge earlier than prototypical derivatives, such compound-like verbs may also emerge earlier than prototypical derivatives.

The transition between grammatical and extragrammatical morphology can be exemplified by the area of submorphemic units in expressive lexemes (Klamer 2001) or reduplication (Mattiello 2013). This topic is beyond the scope of our volume, however, since the earliest reduplication types in child language are clearly extragrammatical (Dressler, Dziubalska-Kořaczyk, Gagarina & Kilani-Schoch 2005, 2014), or are so at least in the languages of this volume.

2.2 Preferences within derivational morphology

General word-formation preferences, which by their nature should hold more for child speech (henceforth CS) than for adult speech, are (Dressler & Kilani-Schoch 2017): morphotactic transparency, i.e. easy formal decomposition of derivations and compounds; morphosemantic transparency, i.e. easy recoverability of the meaning of constituents of derivations and compounds from the holistic meaning of the whole complex word (see Section 3.2.1. a); iconicity of head-non-head relations, i.e. the formal head of a derivation and compound is also its semantic head, as in *teach-er*, where the suffix is responsible for the whole word being an agent and to which inflectional suffixes attach. All languages in our volume have inflectional prefixation either never or only sporadically. In the case of inflectional suffixation, derivational prefixes cannot be morphotactic heads and therefore there is no iconicity with their morphosemantic headhood. This may explain why prefixes are often also morphosemantic non-heads.

Some additional and very strong preferences demand that the bases should be lexical and not grammatical (as in German *Ich-heit* 'egotism, lit. I-ness', Latin *nos-ter* 'our' from its base *nos* 'we'). Moreover, lexical bases should be lexical base forms representing lexemes rather than smaller forms (roots), and simple rather than complex words (phrases, clauses or even sentences). Another frequently debated preference, which holds in the majority of languages, including all of those discussed in our volume, is the preference of suffixation over prefixation (Clark 1993: 122; Hall 2000). The preference for the continuity of morphemes predicts a dispreference for infixation (discontinuity of the lexical base) and circumfixation (discontinuity of the circumfix, Hall 2000).

2.3 Classification of patterns of derivational morphology

A first classification opposes concatenative derivational morphology, i.e. by affixation, to non-concatenative internal modification (e.g. by apophony or metaphony) and conversion, whereby affixation and modification can combine. Concatenative morphology is preferred because it is morphotactically transparent. In early language acquisition, however, other general principles such as simplicity (see Section 3.2.1 b, e) might interact with this preference, which may explain Clark's finding (1993: 120) that converted verbs emerge before affixed verbs in English and German. Conversion (zero-derivation, root-derivation) is the simplest option (Clark 1993: 122), but it is a non-iconic operation since there is no formal change corresponding to a change in meaning. Thus, its early importance in English CS depends also on the great importance of conversion in the target language, which is much greater in English than in any language discussed in this volume except French.

A second classification identifies the word classes (and their subcategories) to which the base and the derived word belong. All the languages in our volume have a noun bias in this respect, which should be even stronger in young children (Dhillon 2010). This preference should hold rather for the derived words than for their lexical bases. Since adjectives typically emerge later than nouns and verbs (Caselli et al. 1995; Salerni, Assanelli, D'Odorico & Rossi 2007; Tribushinina, Voeikova & Noccetti 2015), this must also have an effect on the order of acquisition within derivational morphology (Berman 1999). At any rate, derivational morphology helps to establish and reinforce word-class identification by children (Vapnarsky & Veneziano 2017).

A dependent third classification opposes word-class changing to word-class maintaining derivations (e.g. Bauer 2004). Conversion is by definition word-class changing (Vapnarsky & Veneziano 2017), this is much less prototypical for suffixation due to suffixes being prototypically heads, whereas neither of these are true for prefixation.

A fourth classification opposes simple derivatives, i.e. those which have undergone a single morphological operation (mainly affixation) to complex ones, i.e. those which have undergone more than one derivational operation (mainly two or more affixes). For acquisition, a linear rise of complexity can be predicted (see Clark 1993: 119–122).

The fifth classification opposes concrete meanings to abstract meanings, the latter being cognitively more complex. This leads to the prediction that concrete meanings of derivatives will emerge before abstract ones, as generally occurs in lexical acquisition (Gleitman et al. 2005; Langer 2001: 29).

3. Methodology

3.1 Longitudinal study of spontaneous speech

Research on the acquisition of word formation in general began with diaries (Clark 1993 *passim*, reviewing earlier work); a recent and very data-rich example is Rainer (2010). Quantitative transversal studies on derivation started only later, much less than on compounds, focussed mainly on single languages and are only rarely cross-linguistic (e.g. Anglin 1993; Becker 1994; Behrens 1998; Mattes 2019; Ravid & Avidor 1998). Nearly all of these studies involved a limited number of children, which is also the case for our volume.

Our volume aims at filling the lacuna of cross-linguistic longitudinal studies, because only systematic developmental investigations of the interaction between parental input and child output have a high ecological validity for showing the developmental paths of early first language acquisition in languages of different types. Studies on the acquisition of derivational morphology rarely include a discussion of CDS. However, CDS can differ greatly from adult-directed speech (ADS), at least in early phases of acquisition (Soderstrom 2007). Therefore, we must always compare CS with CDS and if possible CDS with ADS, in order to connect our findings with the target language.

As a consequence, it is unreasonable to measure acquisition in relation to the respective target language descriptions or to electronic corpora of ADS (often only written speech), as has become more and more evident in acquisitionist literature (Clark 2010; Kilani-Schoch et al. 2009; MacWhinney 2004; Ravid et al. 2008). ADS descriptions and corpora include a sizeable amount of Latinate word formation (despite of cross-linguistic differences), which is relatively unimportant in early CDS and even more in CS.

Moreover, many acquisition studies make direct input and output comparisons, omitting an explicit consideration of the intermediate levels of intake and uptake (Clark 2010; Harris 1992; Mitchell, Myles & Marsden 2013: 164 ff.). What the young child actually takes in from derivational morphology, i.e. perceives or notices of the input, may be limited to the suffix because of its positional salience (recency effect). In addition, prosodic salience may reduce the intake to the stressed syllable of the derivative. What children take up, i.e. abstract from the intake and store in their implicit memory, represents a black box to be reconstructed by the analyst.

The approach of our Cross-linguistic Project on Pre- and Protomorphology in Language Acquisition and thus also of this volume is based on longitudinal studies of spontaneous interaction between children and their caretakers, thus focusing on the development of children's input and output. This contrasts with most studies of word formation, which rely on tests and therefore on the output of children

in relatively formal situations usually demanding more language awareness than is needed in spontaneous speech. Such studies usually cannot compare output with input and have only a secondary developmental perspective in comparing distant age brackets, most often of different groups of children only. The results concerning the developmental phase of the children tested are thus only of limited ecological validity.

A disadvantage of longitudinal studies of individual interactions between children and caregivers is that there is necessarily much individual variation in the properties of speech situations, especially in the topics discussed and in speech acts, which renders generalisations more difficult, not to mention that only few such case studies can be financed.

3.2 Phases of development

3.2.1 Generalities

Children treat the first derivatives they meet as whole lexical items (Clark 2014). Only later do they start to recognize the word constituents and the combinatorial processes in derivational word formation and link the roots (or other bases) and affixes to a specific meaning – presumably this results from repeated exposure to bases such as *forgive*, *happy* and *kind* and derivatives such as *forgive-ness*, *happi-ness* and *kind-ness* (Clark 2014). This means that the spontaneous use children make of conventional derivatives (e.g. *kindness*) does not inform us whether or not the child concerned has analysed these forms, or whether or not the child is able to form them productively. According to Clark (1993) a number of input factors, in conjunction with conventionality and contrast, account for when and how children build up a repertoire of word-formation devices for extending their vocabulary. These factors are:

- a. *Morphosemantic transparency* (also referred to as transparency of meaning): Speakers, and especially children, try to interpret and coin new words that are transparent in meaning – that is, words based on known roots and affixes (Clark 1993: 116; see also Clark 2014).
- b. *Simplicity of form*: Speakers find it easier to interpret and coin a new word the simpler it is in form – that is, the less its basis changes in its construction (Clark 1993: 120; see also Clark 2014). This has two aspects: morphotactic transparency and the number of constituents which is the basic factor of complexity.
- c. *Productivity*: The degree to which speakers can use a morphological pattern to coin new forms (Bauer 1983: 18; see also Clark 1993, 2014; Dressler 2007). Baayen (2009) refers to affix category size as ‘realized productivity’, i.e. (with regard to derivation) lemma frequency reflects the relative success with which

the affix has produced new derivatives over time (which may have been long, so that also now unproductive rules may have a high lemma and token frequency). The larger the affix category size, the more productive the morphological pattern is considered to have been. The affix category size, however, is also dependent on the context in which it is measured (Laws & Ryder 2018).

- d. *Frequency*: The effects of frequency are pervasive in first language acquisition, i.e. words and constructions with a high frequency in the child's input will (all else being equal) be acquired earlier than words and constructions with a low input frequency (e.g. Ambridge, Kidd, Rowland & Theakston 2015; Behrens 2006; Gülzow & Gagarina 2007). A distinction is made between token frequency effects on the one hand, i.e. the more often you hear a word, the more likely you are to learn it and type frequency effects on the other hand, i.e. the more you hear a particular morphosyntactic pattern, the more likely you are to learn it (Ambridge et al. 2015). For the acquisition of grammar (including the derivational system), type and lemma frequency is more important than token frequency (Bybee 1995: 430). Frequency of the pattern in ADS can be compared with frequency in CS and CDS only, if the respective pattern has at least a minimal frequency in CDS (see Section 3.1).

In our cross-linguistic project we differentiate the following developmental phases:

3.2.2 *Premorphology*

In the premorphological phase, i.e. the phase before the detection of morphological segmentability and composition, children produce only isolated rote-learned derivatives, if at all, without recurrence of either their lexical base in other words or of their affixes, i.e. they remain unanalysable and are therefore of very little interest for our approach.

3.2.3 *Protomorphology*

The protomorphological phase begins when a child detects, within the expanding lexicon, morphology as a means of decomposing and composing the meaning and form of inflection, compounds, diminutives and other derivations, if such patterns have sufficient frequency in CDS. So far there has been no comparison between the dates of the detection of inflection vs. derivation (Swan 2000: 202–203). This is a lacuna we want to fill.

Segmentation usually proceeds from the periphery to the centre of a word; the process occurs first word-finally and only later word-initially (due to the recency effect, see Griffin 2002; Peters 1997). This corresponds to the claim within the Usage-Based Approach, that the internal structure of complex words is discovered by relating them to similar morphological forms with similar grammatical function in the language (Bybee 2010; Stephany 2012). This is a process of language structure

induction based on growing awareness of repeated internal chunks of words, the alignment of such chunks and the comparison of these with similarly distributed chunks (for syntax see Edelman & Waterfall 2007).

Emergence of a derivational pattern, even if it is not totally isolated, does not mean that it is already acquired. According to which criteria can children be assumed to have detected the morphological make-up of patterns, in our case of derivatives? For this purpose, we have developed first for inflection, then for compounding, the concept of the mini-paradigm criterion as criterion of acquisition of a pattern (Bittner, Dressler & Kilani-Schoch 2003: xvi; Dressler, Ketrez & Kilani-Schoch 2017: 7, 26–27; Dressler, Kilani-Schoch & Klampfer 2003), of course before attaining what Berman (2004) calls full mastery.

When adapted to derivational morphology, this criterion means that a child has to produce spontaneously at least three derivatives with each lexical base, which itself occurs twice as frequently as the autonomous word or, more importantly, as part of other derivatives or compounds in the same or adjacent recordings, neither in formulaic nor in immediately imitating speech and that the same affix or non-concatenative process recurs on two more occasions.

Recurrence does not mean strict identity, insofar as, first, the meaning of the base of morphosemantically fully transparent derivatives often does not equal the meaning of the corresponding autonomous word: for example, only different parts of the meanings of *earth* recur in the meanings of *earth-en* and *earth-ly*. Second, due to phonological processes or morphonological alternations, the forms may not be identical, as the root-final consonant of *electric* and *electric-ity* exemplifies. Children establish relations not only between identical, but also between similar items.

In contrast to our project's inflectional data (Bittner, Dressler & Kilani-Schoch 2003; Stephany & Voeikova 2009), the derivational data are not dense enough to always allow the complete fulfilment of the mini-paradigm criterion. This criterion will therefore often be weakened to the emergence of at least one recurrence of the same base or of the same affix or morphological operation.

Soon after their detection of word-formation morphology, children typically produce novel words (neologisms) in analogy to acquired word-formation patterns. This supports the assumption that they have acquired at least the basics of these patterns (Clark 1993: 291, 301; Berman 1999). However, this criterion has been proven not to be always reliable (Dressler, Ketrez & Kilani-Schoch 2017: 288).

Identifying the absolute chronology of when a representative pattern of one morphological component emerges and becomes productive is not enough. We also investigate the relative chronology of whether patterns of two or all morphological components emerge or become productive (approximately) at the same time, for example diminutive formation, inflection and compounding in German (Dressler, Lettner & Korecky-Kröll 2010, 2012).

Compounds are more descriptive, and therefore, morphosemantically more transparent than derivations. Thus, children may prefer compounding even after the emergence of derivational patterns and replace existing derivatives with newly invented compounds, such as agentive and instrumental English *open-er* with *open+man* and *open+thing* respectively (Clark & Hecht 1982; Clark 2016: 294, 300). This is documented even for Italian, a much more derivational than compounding language (Lo Duca 1990: 125 ff.).

3.2.4 Core morphology

The third developmental phase is core morphology (Ravid et al. 2008), the beginning of what we call “morphology proper”, i.e. adult-like morphology. For derivation, core morphology means that the major derivational patterns of the target language have emerged, and productive patterns are used productively by the child. Further development of morphology proper beyond core morphology consists in what Berman (2004: 14) calls “proficient integration of knowledge and use” towards full mastery of morphology. However, this phase is beyond the longitudinal data collection and analysis dealt with in this volume.

4. Relations between input and output

The authors of this volume agree with the basic tenets of the Usage-Based Approach (Tomasello 2003) for early language development. The area where approaches differ is in assumptions about when and according to which criteria children extract schemas and later on possibly rule-governed patterns from their input. Of particular interest are empirical questions concerning the relation of input frequency to the development of children’s output (see Section 3.2.1 c, d). Do output frequencies approach input frequencies in the course of development? Do caretakers fine-tune the frequency of using a derivational pattern in the input according to the development in the output? Thus, do caretakers increase their usage of derivatives when their children have started to produce them themselves? To what extent can input frequency distributions explain the order of emergence of derivational patterns? Furthermore, the impact of the degree of productivity of derivational patterns should not be reduced to frequency in language acquisition.

There are several aspects of the impact of frequency that are controversial (see Gülzow & Gagarina 2007). As mentioned above, frequencies in electronic corpora are often not very reliable, since they generally contain only samples of adult-directed written language (see Section 3.1). In addition, the proposal to replace markedness with frequency (Haspelmath 2006) is flawed by the erroneous assumption that text frequency is the relevant independent variable rather than the

frequency with which speakers think verbally of an item, as can be illustrated with the pragmatically important compound ‘tooth+brush’, which we pronounce and even write in ADS infrequently, but of which we think at least twice a day in most of our societies (Dressler, Libben & Korecky-Kröll 2014: 187 f.). Thus, continuing the tradition initiated by Jakobson (1938/1968), we can assume that unmarked morphological structures are acquired before the corresponding marked ones, and they are not necessarily the most frequent ones. We posit that this also holds empirically for the preferences in language acquisition listed in Section 2.2.

In the course of our investigation of derivation in CS and CDS, we measured lemma and token frequencies in CS and CDS, but not inflectional type frequencies, which is of minor relevance for word formation (see Section 3.2.1 d).

5. Linguistic typology

A central aim of the Cross-linguistic Project on Pre- and Protomorphology in Language Acquisition is the comparison of acquisition in many different languages with different structural features. This comparative view is an important prerequisite to understanding which observations in child language development can be explained by universal processing mechanisms, and, conversely, to what extent developmental courses are guided by the typological and/or individual features of the acquired language, or, more precisely, how all these factors interact (see also Norcliffe, Harris & Jaeger 2015: 1009). Such findings “can critically advance psycholinguistic theory building and empirical coverage” (Norcliffe, Harris & Jaeger 2015: 1023).

The ten languages covered by this volume belong to four language families and eight branches of these families, as shown in Table 1:

Table 1. Language families, branches and languages

Language families	Branches	Languages
Indo-European	Germanic	Danish, German
	Baltic	Lithuanian
	Slavic	Russian, Croatian
	Romance	French
	Greek	Greek
Finno-Ugric	Fennic	Estonian, Finnish
Altaic	Turkic	Turkish

The diversity of languages in our cross-linguistic project allows a passage from mere cross-linguistic comparison to linguistic typology, which permits studying the possible impact of linguistic types on first language acquisition. Avoiding the outdated first stage of linguistic typology, which assigns existing languages to different classes of languages, we adhere to the superior second stage of typology, namely ordering typology (Hempel & Oppenheim 1936), which identifies whether language A is more of a certain type (e.g. agglutinating) than language B. Moreover, we do not compare whole languages, but corresponding components of a language, in our case derivational morphology within word formation.

For compounding, we have observed that the strongest typological impact on its acquisition is the specific wealth of compounding in CDS (Dressler, Ketrez & Kilani-Schoch 2017; Xanthos et al. 2011), analogously to what we have shown for the acquisition of diminutives (Dressler, Lettner & Korecky-Kröll 2012; Savickienė & Dressler 2007). This allows us to compare the respective wealth of derivational morphology with compounding in each specific language and in groups of languages.

In regard to the languages of our volume, language typology coincides largely with genealogical relationship. Within the framework of ordering typology (see above), the Germanic languages are more compounding than Romance and Slavic languages, while Greek and Turkish are more derivational (Dressler, Ketrez & Kilani-Schoch 2017; Posner 1996; Pounder 2000). Finnish and Estonian are rich in both components, although Estonian appears to be the richer of the two in compounding. The compounding vs. derivational character of Modern Greek, however, depends to a great extent on register and style (Stephany 2017; Stephany & Thomadaki 2017).

By contrast, with respect to the inflectional vs. the agglutinating type, the ten languages studied here can be ordered in the following way (Dressler 2010; Dressler, Ketrez & Kilani-Schoch 2017; Stephany & Voeikova 2009: 2 f.): Turkish is the strongest agglutinating language, less strongly agglutinating is Finnish, even less strongly agglutinating but still morphologically very rich is Estonian (the most inflecting-fusional one among major Fennic languages), the strongest inflecting-fusional Lithuanian, then Russian and Croatian, weakly inflecting-fusional are German, Danish, Greek, French (the weakest inflecting-fusional). However, this typological ordering plays a minor role in word formation, since its role is restricted to the level of morphosemantic transparency, which is higher the more agglutinating a language is, and lower the more inflecting an inflecting-fusional language is. The weaker inflecting languages are, the more conversions occur, as in French, whereas agglutinating languages are nearly devoid of conversions.

6. Summaries of chapters

The chapter by Ursula Stephany presents an analysis of the speech of one Greek girl from 1;8 to 3;0. The first derivational prefixes for nouns and verbs and diminutive suffixes emerge from 1;8 onwards, simultaneously with compounding. The most numerous derivational formations occurring in CS (besides diminutives) are suffixed verbs. Whereas there is evidence that verbal prefixes become productive in the third year, the frequently occurring but mostly opaque suffixes remain unproductive. Adjectives derived from nouns by suffixes appear in CS from 1;8 and increase in frequency in the second year, but remain unproductive. Suffixed nominal derivatives (with the exception of diminutives) are even more scarce than derived adjectives. The comparison of CS and CDS reveals that derivational affixes are highly frequent and much more various in CDS than in CS from the beginning. There is evidence that lemma and type frequency in CDS promotes productivity in CS.

Marianne Kilani-Schoch and Aris Xanthos examine the morphological structure of the early French lexicon before neologisms are coined. The data come from the same corpus of two Swiss girls (up to 2;11) that has been used in the other volumes of our cross-linguistic project. The results show that derivational morphology compensates for the later emergence of compounding but still develops slowly, starting earlier with nouns than verbs and earlier in noun suffixation and conversion than in prefixation. This chapter compares CS not only with CDS but also with ADS and discusses on methodological issues such as how to identify early signs of derivational development, including statistical measures.

The Danish chapter by Laila Kjærbaek and Hans Basbøll, presenting the results of four Danish-speaking children up to age 3;11, shows that derivation appears almost simultaneously with compounding and the first inflectional categories. However, even though a large number of derivational patterns appear in CDS, only a few derivational patterns emerge in CS, and no innovative derivatives appear in the corpus. Conversion and implicit derivation are by far the most frequent derivational patterns in CS. The children only produce a few suffix and prefix derivations, out of which the adjective suffix *-ig/-lig* is clearly the most frequent.

The German chapter by Sabine Sommer-Lolei, Veronika Mattes, Katharina Korecky-Kröll and Wolfgang U. Dressler, presenting the results for three Viennese children (up to 3;0), shows that diminutives emerge and develop at the same time as compounding and first inflectional categories. The next category that emerges is that of particle verbs, the status of which is transitional to compounding, whereas the similar but more clearly derivational prefixed verbs develop even later. Within suffixation deverbal instrument nouns develop earlier than agent nouns, which in turn are earlier acquired than result nouns and action nouns.

Gordana Hržica studies the early development in the closely cognate South Slavic language Croatian on the basis of three children (up to 3;2 for the richest corpus). They show a similar rise in early verbal vs. early nominal derivation. Verbs are mainly prefixed and nouns suffixed. The most prevalent semantic categories in nouns are objects and instruments, and derivation in verbs marks differences in aspect and/or aktionsart. There is also a preponderance of the verbal prefix *po-*, although it plays a much smaller aspectual role than in Russian. In early acquisition of Croatian, the number of acquired affixes and proportion of derived words increase with age. Later child speech shows productive usage of derivation, including derivational pairs and mini-paradigms, neologisms and neosemantisms. The frequency of derived words in CDS had a strong impact on the acquisition of derivational morphology.

Victoria V. Kazakovskaya and Maria D. Voeikova present their study of two Russian boys (up to 2;8 and 3;0) from Saint-Petersburg. Similar to other languages, they found an earlier development of prefixation for verbs and suffixation for nouns and adjectives. This is related to the importance of verb prefixation both for lexical semantic change and for the expression of aktionsarten. Interindividual variation can be strong, for example one of the boys produces twice the percentage of derived nouns than the other, the reason being that he starts to produce them much earlier. This preference for noun derivatives is characteristic of a child who is an early speaker whereas the late talker prefers verb derivatives. A later spurt of derived adjectives coincides in both boys with a general adjective spurt.

The chapter on the Baltic language Lithuanian by Laura Kamandulytė-Merfeldienė, Ingrida Balčiūnienė and Ineta Dabašinskienė examines the early phases of acquisition by two children up to 2;7. Derivational morphology, especially and most massively diminutives, emerges earlier than compounding, but about the same time as inflection, always latest adjectives. Prefixation emerges also very early, but only with verbs, first often with the pre-morpheme *a*. Verb derivation also showed greatest variation among word classes and was the first to develop productivity and expand word families soon after emergence, first with many erroneous commissions.

The Estonian chapter by Reili Argus presents the development of derivation in three children (from 1;3 to 3;1). They start to produce derived words at the same time as compounding and inflection. In Estonian, only suffixation plays a role for derivation in CS. In nouns, diminutives clearly dominate, but also some very general and not very transparent nominalization suffixes (for persons, entities, results) are used early, but do not become productive before age 3. With verbs, which have a much more transparent derivation system than nouns, the valency-changing suffixes (causatives and later resultatives) emerge and become productive before age 3. The distribution of derivations in CS and CDS is very similar.

Klaus Laalo analyses the development of two Finnish children. Of the very rich derivational system of Finnish derived nouns, verbs and adjectives appear in CS more or less at the same time and with several suffixes in all word classes, between 1;7 and 1;11 (with variation among the children, but in all cases before the second birthday), and together with both inflection and compounding. Laalo points out that emergence and productivity of the derivational suffixes is clearly depending on frequency in CDS, as well as on productivity and transparency.

The Turkish chapter by F. Nihan Ketrez and Ayhan Aksu-Koç presents the data of two children (up to 3;1). In this agglutinating language, suffixation is the major word-formation device. In both children's speech, emergence of word formation starts with derivation in the verbal domain, along with inflections, and the first derivation examples are causative verbs. Derivatives emerge in nouns only after compounds, although compounding is much less frequent than derivation in Turkish. Semantically, instrument nouns are produced before agent nouns. The earliest and most frequent verb derivations used by the children are voice-suffixes (comparable to the other agglutinating languages in this volume) and the authors discuss the influence of the closeness of voice derivation to inflection on the overall picture of the acquisition of derivation. The distribution of derived forms in CS closely reflects the distribution in CDS. Although the two children differ in their pace of development, they display similar paths in their acquisition of derivational morphology.

The completion of this cooperative volume was a lengthy process. As a result, we were unable to consider partially relevant recent developments in the theory of derivational morphology in our planning. These developments included the recent rapid progress in the study of derivational paradigms (Hathout & Namer 2019), which goes far beyond earlier proposals, such as Dressler & Ladányi (2000). We hope that our adaptation of the mini-paradigm to derivational morphology (see Section 3.2.3) will contribute to studies on the origins of derivational paradigms in early phases of acquisition.

References

- Ambridge, B., Kidd, E., Rowland, C. F. & Theakston, A. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2): 239–253.
<https://doi.org/10.1017/S030500091400049X>
- Anglin, J. M. 1993. Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development* 58(10): 1–166. <https://doi.org/10.2307/1166112>
- Baayen, R. H. 2009. Corpus linguistics in morphology: Morphological productivity. In *Corpus Linguistics: An International Handbook*, A. Lüdeling & M. Kytö (eds), 899–919. Berlin: Mouton de Gruyter.
- Bauer, L. 1983. *English Word-Formation*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9781139165846>
- Bauer, L. 2004. The function of word-formation and the inflection-derivation distinction. In *Words in their Places. A Festschrift for J. Lachlan Mackenzie*, H. Aersten, M. Hannay & R. Lyall (eds), 283–292. Amsterdam: Vrije Universiteit.
- Becker, J. A. 1994. ‘Sneak-shoes’, ‘sworders’, and ‘nose-beards’: A case study of lexical innovation. *First Language* 14: 195–211. <https://doi.org/10.1177/014272379401404213>
- Behrens, H. 1998. How difficult are complex verbs? Evidence from German, Dutch and English. *Linguistics* 36(4): 679–712. <https://doi.org/10.1515/ling.1998.36.4.679>
- Behrens, H. 2006. The input-output relations in first language acquisition. *Language and Cognitive Processes* 21: 2–24. <https://doi.org/10.1080/01690960400001721>
- Berman, R. A. 1999. Children’s innovative verbs vs. nouns: Structured elicitations and spontaneous coinages. In *Methods for Studying Language Production*, L. Menn & N. Bernstein-Ratner (eds), 69–93. Mahwah NJ: Lawrence Erlbaum Associates.
- Berman, R. A. 2004. Between emergence and mastery: The long developmental route of language acquisition. In *Language Development across Childhood and Adolescence* [Trends in Language Acquisition Research 3], R. A. Berman (ed.), 9–34. Amsterdam: John Benjamins.
<https://doi.org/10.1075/tilar.3.05ber>
- Bittner, D., Dressler, W. U. & Kilani-Schoch, M. (eds) 2003. *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110899832>
- Bybee, J. L. 1985. *Morphology: A Study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.9>
- Bybee, J. L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455. <https://doi.org/10.1080/01690969508407111>
- Bybee, J. L. 2010. *Language, Usage, and Cognition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511750526>
- Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Sanderl, L. & Weir, J. 1995. A cross-linguistic study of early lexical development. *Cognitive Development* 10: 159–199.
[https://doi.org/10.1016/0885-2014\(95\)90008-X](https://doi.org/10.1016/0885-2014(95)90008-X)
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511554377>
- Clark, E. V. 2010. Adult offer, word-class, and child uptake in early lexical acquisition. *First Language* 30: 250–269. <https://doi.org/10.1177/0142723710370537>
- Clark, E. V. 2014. Acquisition of derivational morphology. In *The Oxford Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 424–439. Oxford: OUP.

- Clark, E. V. 2016. *The Lexicon in Acquisition*, 3rd edn. Cambridge: CUP.
- Clark, E. V. & Hecht, B. F. 1982. Learning to coin agent and instrument nouns. *Cognition* 12: 1–24. [https://doi.org/10.1016/0010-0277\(82\)90027-0](https://doi.org/10.1016/0010-0277(82)90027-0)
- Dhillon, R. 2010. Examining the ‘noun bias’: A structural approach. *University of Pennsylvania Working Papers in Linguistics* 16: 51–60.
- Doleschal, U. & Thornton, A. (eds) 2001. *Extragrammatical and Marginal Morphology*. Munich: Lincom.
- Dressler, W. U. 2007. Productivity in word formation. In *The Mental Lexicon: Core Perspectives*, G. Jarema & G. Libben (eds), 159–183. Amsterdam: Elsevier. https://doi.org/10.1163/9780080548692_009
- Dressler, W. U. 2010. A typological approach to first language acquisition. In *Language Acquisition Across Linguistic and Cognitive Systems* [Language Acquisition and Language Disorders 52], M. Kail & M. Hickmann (eds), 109–124. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.52.09dre>
- Dressler, W. U., Dziubalska-Kolaczyk, K., Gagarina, N. & Kilani-Schoch, M. 2005. Reduplication in child language. In *Studies on Reduplication*, B. Hurch (ed.), 455–474. Berlin: Mouton de Gruyter.
- Dressler, W. U., Dziubalska-Kolaczyk, K., Gagarina, N. & Kilani-Schoch, M. 2014. Reduplication, repetition, hypercharacterisation and other affix-doubling in child language. In *Affix Ordering across Languages and Frameworks*, S. Manova (ed.), 259–276. Oxford: OUP.
- Dressler, W. U., Ketrez, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61>
- Dressler, W. U. & Kilani-Schoch, M. 2017. Natural morphology. In *The Cambridge Handbook of Morphology*, A. Hippisley & G. Stump (eds), 356–389. Cambridge: CUP.
- Dressler, W. U., Kilani-Schoch, M. & Klampfer, S. 2003. How does a child detect morphology? Evidence from production. In *Morphological Structure in Language Processing* [Trends in Linguistics. Studies and Monographs 151], R. H. Baayen & R. Schreuder (eds), 391–425. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110910186_391
- Dressler, W. U. & Ladányi, M. 2000. Productivity in word formation (WF): A morphological approach. *Acta Linguistica Hungarica* 47: 103–144. <https://doi.org/10.1023/A:1014010530824>
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2010. First language acquisition of compounds. With special emphasis on early German child language. In *Cross-Disciplinary Issues in Compounding* [Current Issues in Linguistic Theory 311], S. Scalise & I. Vogel (eds), 323–344. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.311.24dre>
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2012. Acquisition of German diminutive formation and compounding in a comparative perspective: Evidence for typology and the role of frequency. In *Current Issues in Morphological Theory. (Ir)regularity, Analogy and Frequency* [Current Issues in Linguistic Theory 322], F. Kiefer, M. Ladányi & P. Siptár (eds), 237–264. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.322.11dre>
- Dressler, W. U., Libben, G. & Korecky-Kröll, K. 2014. Conflicting vs. convergent vs. interdependent motivations in morphology. In *Competing Motivations in Grammar & Usage*, B. MacWhinney, A. Malchukov & E. Moravcsik (eds), 181–196. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780198709848.003.0011>
- Dressler, W. U., Mayerthaler, W., Panagl, O. & Wurzel, W. U. 1987. *Leitmotifs in Morphology* [Studies in Language Companion Series 10]. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.10>

- Edelman, S. & Waterfall, H. 2007. Behavioral and computational aspects of language and its acquisition. *Physics of Life Reviews* 4: 253–277. <https://doi.org/10.1016/j.plprev.2007.10.001>
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A. & Trueswell, J. C. 2005. Hard words. *Language Learning and Development* 1(1): 23–64. https://doi.org/10.1207/s15473341l1do101_4
- Griffin, Z. M. 2002. Recency effects for meaning and form in word selection. *Brain and Language* 80: 465–487. <https://doi.org/10.1006/brln.2001.2609>
- Gülzow, I. & Gagarina, N. (eds) 2007. *Frequency Effects in Language Acquisition*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110977905>
- Hall, C. J. 2000. Prefixation, suffixation and circumfixation. In *Morphologie / Morphology: Ein internationales Handbuch zur Flexion und Wortbildung / An international Handbook on Inflection and Word Formation*, G. E. Booij, C. Lehmann, C. J. Hall & J. Mugdan (eds), 535–545. Berlin: De Gruyter. <https://doi.org/10.1515/978311011286.1.8.535>
- Harris, M. 1992. *Language Experience and Early Language Development: From Input to Uptake*. Hove: Psychology Press.
- Haspelmath, M. 1996. Word-class changing inflection in morphological theory. In *Yearbook of Morphology 1995*, G. E. Booij & J. van Marle (eds), 43–66. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-3716-6_3
- Haspelmath, M. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25–70. <https://doi.org/10.1017/S0022226705003683>
- Hathout, N. & Namer, F. 2019. Paradigms in word formation: What are we up to? *Morphology* 29: 153–165. <https://doi.org/10.1007/s11525-019-09344-3>
- Hempel, C. & Oppenheim, P. 1936. *Der Typenbegriff im Lichte der neuen Logik* (The Concept of Type in the Light of the New Logic). Leiden: Sijthoff.
- Jakobson, R. 1939/1968. *Child Language, Aphasia, and Phonological Universals*. The Hague: Mouton.
- Kilani-Schoch, M., Balčiūnienė, I., Korecky-Kröll, K., Laaha, S. & Dressler, W. U. 2009. On the role of pragmatics in child-directed speech for the acquisition of verb morphology. *Journal of Pragmatics* 41: 219–239. <https://doi.org/10.1016/j.pragma.2008.10.001>
- Klamer, M. 2001. Expressives and iconicity in the lexicon. In *Ideophones*, E. F. K. Voeltz & C. Kilian-Hatz (eds), 165–181. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.44.14kla>
- Langer, J. 2001. The mosaic evolution of cognitive and linguistic ontogeny. In *Language Acquisition and Conceptual Development*, M. Bowerman & S. C. Levinson (eds), 19–44. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511620669.003>
- Laws, J. & Ryder, C. 2018. Register variation in spoken British English: The case of verb-forming suffixation. *International Journal of Corpus Linguistics* 23(1): 1–27. <https://doi.org/10.1075/ijcl.16036.law>
- Lo Duca, M. G. 1990. *Creatività e regole. Studio sull'acquisizione della morfologia derivativa dell'Italiano* (Creativity and rules. Study on the acquisition of derivational morphology of Italian). Bologna: Il Mulino.
- MacWhinney, B. 2004. A multiple process solution to the logical problem of language acquisition. *Journal of Child Language* 31: 883–914. <https://doi.org/10.1017/S0305000904006336>
- Mattes, V. 2019. What do children know about German verb prefixes? A study on the development of verb derivation from preschool age to school age. *The Mental Lexicon* 14(2): 274–297. <https://doi.org/10.1075/ml.00007.mat>
- Mattiello, E. 2013. *Extra-Grammatical Morphology in English. Abbreviations, Blends, Reduplicatives and Related Phenomena*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110295399>

- Mitchell, R., Myles, F. & Marsden, E. 2013. *Second Language Learning Theories*. New York NY: Routledge. <https://doi.org/10.4324/9780203770795>
- Norcliffe, E., Harris, A. & Jaeger, F. 2015. Linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9): 1009–1032. <https://doi.org/10.1080/23273798.2015.1080373>
- Peters, A. M. 1997. Language typology, prosody, and the acquisition of grammatical morphemes. In *The Crosslinguistic Study of Language Acquisition*, D. I. Slobin (ed.), 135–197. Hillsdale NJ: Lawrence Erlbaum Associates.
- Posner, R. 1996. *The Romance Languages*. Cambridge: CUP.
- Pounder, A. 2000. *Processes and Paradigms in Word-Formation Morphology*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110814378>
- Rainer, F. 2010. *Carmens Erwerb der deutschen Wortbildung* (Carmen's acquisition of German word formation). Wien: Verlag der Österreichischen Akademie der Wissenschaften. <https://doi.org/10.1553/0x00237d8b>
- Ravid, D. & Avidor, A. 1998. Acquisition of derived nominals in Hebrew: Developmental and linguistic principles. *Journal of Child Language* 25: 229–266. <https://doi.org/10.1017/S0305000998003419>
- Ravid, D., Dressler, W. U., Nir-Sagiv, B., Korecky-Kröll, K., Souman, A., Rehfeldt, K., Laaha, S., Bertl, J., Basbøll, H. & Gillis, S. 2008. Core morphology in child directed speech: Crosslinguistic corpus analyses of noun plurals. In *Corpora in Language Acquisition Research* [Trends in Language Acquisition Research 6], H. Behrens (ed.), 25–60. Amsterdam: John Benjamins. <https://doi.org/10.1075/tilar.6.05rav>
- Salerni, N., Assanelli, A., D'Odorico, L. & Rossi, G. 2007. Qualitative aspects of productive vocabulary at the 200- and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language* 27: 75–87. <https://doi.org/10.1177/0142723707067545>
- Savickienė, I. & Dressler, W. U. (eds) 2007. *The Acquisition of Diminutives: A Cross-Linguistic Perspective* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Say, S. 2005. Antipassive *sja*-verbs in Russian: Between inflection and derivation. In *Morphology and its Demarcations* [Current Issues in Linguistic Theory 264], W. U. Dressler, D. Kastovsky, O. Pfeiffer & F. Rainer (eds), 253–275. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.264.16say>
- Soderstrom, M. 2007. Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review* 27(4): 501–532. <https://doi.org/10.1016/j.dr.2007.06.002>
- Spencer, A. & Luís, A. R. 2012. *Clitics. An Introduction*. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139033763>
- Stephany, U. 2012. Selecting a theoretical framework fitting acquisition data is no easy matter. In *Selected Papers of the 10th International Conference of Greek Linguistics*, Z. Gavriilidou, A. Efthymiou, E. Thomadaki & P. Kambakis-Vougiouklis (eds), 89–100. Komotini: Democritus University of Thrace.
- Stephany, U. 2017. Contrastive lexical typology of German and Greek child speech and child-directed speech. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketzrez & M. Kilani-Schoch (eds), 275–286. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.13ste>

- Stephany, U. & Thomadaki, E. 2017. Compounding in early Greek language acquisition. In Dressler, Ketrez & Kilani-Schoch (eds), 119–143. <https://doi.org/10.1075/lald.61.06ste>
- Stephany, U. & Voeikova, M. (eds) 2009. *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110217117>
- Swan, D. 2000. How to build a lexicon: A case study of lexical errors and innovations. *First Language* 20: 187–204. <https://doi.org/10.1177/014272370002005904>
- Tomasello, M. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge MA: Harvard University Press.
- Tribushinina, E., Voeikova, M. D. & Noccetti, S. (eds) 2015. *Semantics and Morphology of Early Adjectives in First Language Acquisition*. Newcastle upon Tyne: Cambridge Scholars.
- Vapnarsky, V. & Veneziano, E. (eds) 2017. *Lexical Polycategoriality. Cross-Linguistic, Cross-Theoretical and Language Acquisition Approaches* [Studies in Language Companion Series 182]. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.182>
- Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., Gagarina, N., Hržica, G., Ketrez, F. N., Kilani-Schoch, M., Korecky-Kröll, K., Kovačević, M., Laalo, K., Palmović, M., Pfeiler, B., Voeikova, M. D. & Dressler, W. U. 2011. On the role of morphological richness in the early development of noun and verb inflection. *First Language* 31: 461–479. <https://doi.org/10.1177/0142723711409976>

The development of derivation in early Greek first language acquisition

Ursula Stephany
University of Cologne

This study is concerned with the development of derivational prefixation and suffixation in early Greek language acquisition and is based on the speech of a monolingual Greek girl from 1;8 to 3;0 years in interaction with her mother (Katis Corpus). The most productive derivational affixes are verbal prefixes while derivational suffixes prevail in adjectives and nouns. Special attention is paid to the development of word families (lexemes sharing a base) and word series (sets of derivatives sharing a derivational affix) in the child's lexicon and to a comparison with child-directed speech. The study is committed to usage-based approaches to language acquisition according to which "language structure emerges from language use" (Tomasello 2003: 327) and productivity arises within item-based schemas.

Keywords: Greek language acquisition, acquisition of derivation, derivational prefixation, derivational suffixation

1. Introduction

Modern Greek (MG) is a typical strongly inflecting-fusional Indo-European language and possesses manifold word-formation devices, both derivational and compounding, in addition to rich inflectional processes. Although compounding is a wide-spread phenomenon in MG (Ralli 2013; see also Christofidou et al. 2013; Ralli 2014; Stephany & Thomadaki 2016, 2017), derivational patterns permeate the entire lexical stock in multiple intricate ways. As pointed out by Mackridge (1985: 308), one of the factors "underlying the large size of the MG vocabulary is the great ease with which derivatives may be formed" (see also Clairis & Babiniotis 2005; Ralli 2014; Holton, Mackridge & Philippaki-Warbuton 2016).

Word formation in Greek language acquisition has not been extensively studied so far (Stephany & Thomadaki 2016: 6). There is an experimental study on

compounding by Stephany (1980; see also Stephany 1997a: 260) and observational studies are to be found in Thomadaki (1986; see Stephany 1997a: 255–261) and Stephany and Thomadaki (2016, 2017).¹ The most important derivational process in early Greek language acquisition and child-centred situations is diminutive formation (Stephany 1997b; Thomadaki & Stephany 2007). Overgeneralizations of derivational patterns and neologisms have been found in one child at 3;9 years and another one from 6;4 to 6;11 years acquiring Greek as their native language (Stephany 1997a: 257–261 quoting Katis 1984 and Thomadaki 1986). Stephany (1997a: 261) concludes that “morphological and semantic transparency facilitate lexical analysis” and “children [...] extend adult patterns and apply them to new lexical items, thus creating names with a more specific or more expressive meaning and fitting syntactic constructions they choose to use.”

No evidence has been found so far for the emergence of derivational morphology, particularly diminutives, prior to inflectional development in Greek language acquisition (Stephany & Thomadaki 2016: 6; see also Thomadaki & Stephany 2007: 117).² In an experimental study of inflectional and derivational awareness in 236 monolingual Greek children aged 4;0 to 7;0 aiming at identifying critical linguistic skills related to the development of reading comprehension and including judgment and production tasks, Diamanti et al. (2018) found that “full metalinguistic awareness of derivational morphology [tested by production tasks] trails behind that of inflectional morphology” (Diamanti et al. 2018: 545–546).

As far as derivation and compounding are concerned, both of these word-formation processes begin to develop before the end of the second year in Anna’s speech and gain momentum in the second half of her third year.³ The present study of the development of prefixing and suffixing devices in this girl acquiring Greek as her native language is based on longitudinal production data recorded in natural speech situations between the child’s ages of 1;8 to 3;0.

In order to come to grips with the emergence of lexical structure, attention will not only be paid to the development of productive prefixing and suffixing devices

1. For references concerning further studies of compounds in Greek language acquisition see Stephany & Thomadaki (2016: 7).

2. Inflectional development consists in the use of several grammatical forms of a given lemma. Such evidence has been found from the beginning of observation of five subjects at 1;9 and 1;11 by Stephany (1985; see also Christofidou & Stephany 2003 and Stephany & Christofidou 2009). Although the boy Christos uses a single grammatical form per verb until 1;11, different forms of different verbs serve adequate functions (e.g. imperative, non-past, past) (Christofidou & Stephany 2003).

3. On the acquisition of compounding in early Greek language acquisition see Stephany and Thomadaki (2017).

but also of paradigmatic relations among members of word families and word series occurring in the child's lexicon. Evidence for the role played by child-directed speech (CDS) in the acquisition of derivational processes and the enlargement of the child's vocabulary will be obtained from studying the usage of derivational affixes in this register.

The present study is framed within the constructivist, usage-based approach to language acquisition (Tomasello 2003, 2010; Bybee 2010; Ambridge & Lieven 2011; Stephany 2012; Diessel 2013). It is assumed that the internal structure of words "is derivable from sets of connections made between words that have related parts" (Bybee 2007[1998]: 280) and are accordingly semantically and phonologically similar (Bybee 2010: 22). Thus, the internal structure of the Greek derivational formation *para-berdévome* 'to get very confused' may be gathered from its relation to other derivational formations such as *para-léo* 'to say more than needed' (*léo* 'to say') and *para-miláo* 'to speak in an exaggerated way, talk too much' (*miláo* 'to speak') in which the intensifying prefix *para-* denoting exaggeration recurs. Generalizations on verbal derivations such as these may lead to the schema *para*___ whose slot can be filled by a number of different verbs (e.g. *para-dulévo* 'to overwork oneself', *para-tróo* 'to overeat').⁴ Booij (2013: 260) points out that "a construction may have holistic properties that are not derivable from the properties of its constituents and/or its structure." This is especially true for lexical constructions such as compounds or derivations, whose meaning is often specialized and does not equal the sum of meanings of their constituents (e.g. *para-dulévo* 'to overwork oneself' (*dulévo* 'to work'),⁵ but *para-ynorízo* 'to confound, misjudge' (*ynorízo* 'to know'), *para-miló* 'to fantasize, talk in one's dreams' (*miló* 'to speak'⁶)).

The chapter is organized as follows: An overview of the main characteristics of derivation in Standard Modern Greek (SMG), with special attention to the derivational processes playing a role in the data analysed in this paper, is given in Section 2. The data are presented in Section 3. Main Section 4.1 is devoted to the

4. Babiniotis (1998) quotes a list of 130 verbs prefixed by *para-*.

5. In spite of the fact that the verb *paradulévo* 'to overwork oneself' and the noun *paraduléftra* 'maid' have the same root, the verb is a member of the word series *para-dulévo* 'to overwork oneself', *para-tróo* 'to overeat', *para-miláo* 'to speak more than normal' (and many others, see Babiniotis 1998), but there is no derivational relation between the verb *paradulévo* and the noun *paraduléftra* 'maid'. The meaning of *para-* in *paraduléftra* does not express exaggeration but derives from AG *para* 'near' so that *paraduléftra* is the one providing support.

6. In spite of the fact that *paramiló* and *paramiláo* differ in meaning, with *paramiló* meaning 'to fantasize, talk in one's dreams' and *paramiláo* 'to speak in an exaggerated way, talk too much' (Babiniotis 2002), many speakers make variable use of these forms neutralizing their semantic differentiation (D. Theophanopoulou-Kontou, p.c.).

results of the analysis and is subdivided into the development of selected prefixing and suffixing processes in child speech (CS), taking the emergence of productivity, word series, and word families into consideration. The usage of such derivational devices in CDS is studied in Section 4.2 and comparisons are drawn to CS. The concluding section contains a summary of the results and more general conclusions.

2. Derivational devices of Standard Modern Greek

The most important Greek derivational processes are suffixation and prefixation. Other, less frequent processes are *ablaut* (e.g. *tréxo/troxós* ‘to run/wheel’ (Ralli 2014: 140)), syntactic conversion (e.g. *kritikós/Kritikós* ‘Cretan.ADJ/Cretan.N’ (Ralli 2014: 142)), and accent shift (together with morphological conversion; e.g. *kaláthi/kálaθος* ‘basket/basket (in basketball)’ (Ralli 2014: 143)). As pointed out above, the present study is limited to the most frequent and overt derivational processes of affixation in SMG.

Although derivational prefixes and suffixes are not free forms, certain Modern Greek prepositions such as *apó* ‘from’ or *antí* ‘instead, in place of’ may also function as prefixes (e.g. *apo-limantikó* ‘disinfectant’; *anti-stréfo* ‘to put sth. upside down, convert’, *anti-komunistikós* ‘anticommunist’) (Holton, Mackridge & Philippaki-Warburton 2016: 245).⁷ Another example is the adverb *ksaná* ‘again’ which may serve as a preverb⁸ and forms a very productive word series (see below). According to the lists of derivational affixes presented in Clairis and Babiniotis (2005: 79–80, 82–91, 254–260, 708–716), there are more than three times as many derivational suffixes than prefixes (116 vs. 32).⁹ Their respective frequencies differ according to parts of speech. While derivational suffixes outnumber prefixes in nouns (85 vs. 6) and adjectives (23 vs. 9), prefixes predominate in verbs (17 vs. 8).

All major parts of speech, nouns, verbs, adjectives, and adverbs, may be formed by derivation in SMG. Besides the place at which they are attached to the base (prefixes vs. suffixes), derivational affixes may be classified according to the part of speech to which the derivative belongs as well as the class of its base (Table 1).

7. These prepositions include some AG ones “which no longer exist as independent prepositions” in SMG (Holton, Mackridge & Philippaki-Warburton 2016: 244).

8. See Ralli (2004).

9. Authors differ to a certain extent in the kinds affixes considered as derivational prefixes or suffixes (cf. Mackridge 1985; Clairis & Babiniotis 2005; Ralli 2014; Holton, Mackridge & Philippaki-Warburton 2016).

Only derivational suffixes may change the grammatical category of words. Thus, the denominal verbalizing suffix *-iz-* derives verbs from nouns (e.g. *zoyraf-íz-o* ‘to paint’ ← *zoyraf-ía* ‘picture’).

Table 1. Bases of nominal, verbal, and adjectival derivatives (see Ralli 2014: 139–140)

Derivative	Base	Example	Gloss
Noun	Noun	<i>spíti</i> → <i>spit-áki</i>	house → house-DIM
	Adjective	<i>mikrós</i> → <i>mikr-ótita</i>	small → smallness
	Verb	<i>δέno</i> → <i>δέ-simo</i>	bind.1SG → binding
Adjective	Noun	<i>ksílo</i> → <i>ksíl-inos</i>	wood → wooden
	Adjective	<i>γnostós</i> → <i>á-γnostos</i>	known → unknown
	Verb	<i>αγαπό</i> → <i>αγαπι-tós</i>	love.1SG → beloved
Verb	Noun	<i>xorós</i> → <i>xor-évo</i>	dance → dance.1SG
	Adjective	<i>mávros</i> → <i>mavr-ízo</i>	black → blacken.1SG
	Verb	<i>γράφο</i> → <i>δια-γράφο</i>	write.1SG → delete.1SG

The linear and hierarchical structure of derivatives may be quite complex. An example comprising three derivational suffixes and an inflectional ending is the adjective [[[[[*xor*][*ef*]*t*]*ik*]*ós*] ‘dance-like, dancing’ (Ralli 2014: 157). As also shown in Table 2, derivational affixes are placed closer to the stem than inflectional ones (see also Ralli 2014: 54). A given stem may be combined with several derivational prefixes and suffixes leading to quite complex stems carrying an inflectional ending. The inflectional endings of the first two examples in Table 2 serve to distinguish words whose grammatical gender indicates social gender (‘brother’ vs. ‘sister’).

Table 2. Greek derivational processes (based on Clairis & Babiniotis 2005: 77)

Der.Pref	Stem	Der.Suff	Infl.Suff	Word.Class	Gloss	
	<i>áδelf</i>		<i>ós</i>	NOUN	brother	
	<i>áδelf</i>		<i>í</i>	NOUN	sister	
	<i>áδelf</i>	<i>ik</i>	<i>ós</i>	ADJ	brotherly	
	<i>áδelf</i>	<i>ik</i>	<i>ótít</i>	<i>a</i>	NOUN	brotherhood
	<i>áδelf</i>	<i>osín</i>	<i>i</i>	NOUN	fraternity	
<i>sin</i>	<i>áδelf</i>		<i>os</i>	NOUN	colleague	
<i>sin</i>	<i>áδelf</i>	<i>ik</i>	<i>ótít</i>	<i>a</i>	NOUN	staff of colleagues
<i>anti</i>	<i>sin</i>	<i>áδelf</i>	<i>ik</i>	<i>ós</i>	ADJ	uncooperative

The productivity of derivational processes may be blocked by the presence of other lemmas with the same stem and meaning (e.g. *ráftis* ‘tailor’ ← *rávo* ‘to sew’ but **kureftís* ← *kurévo* ‘to cut s.o.’s hair’ because of *kuréas* ‘barber’; Ralli 2014: 156). In contrast to the semantic relation between inflected forms, which remains constant across lemmas, the semantic relation of derivatives to their bases may vary considerably (Table 3; derivatives from Ralli 2014: 48–49). While the meaning of *çísimo* ‘pouring’ and *xásimo* ‘loss’ is close to that of their verbal bases *çin-* and *xan-*, only the literal meaning of *lísimo* ‘dissolution’ is close to that of *líno* ‘to solve, untie’ but the metaphorical meaning ‘cancellation’ is not. Furthermore, *xamós* is close to the meaning of its verbal base *xán-* only in the sense of ‘loss’, but not ‘disorder’ or ‘hell’. Finally, the meaning of *çimós* ‘juice’ is quite idiosyncratic.

Table 3. Semantic relation of inflections vs. derivatives

		Inflection	
<i>lín-o</i>	<i>é-lís-a</i>	<i>é-lín-a</i>	‘I untie/untied/used to untie’
untie.IPFV-NPST.1SG	AUG-untie. PFV-PST.1SG	AUG-untie. IPFV-PST.1SG	
<i>xán-o</i>	<i>é-xas-a</i>	<i>é-xan-a</i>	‘I lose/lost/used to lose’
lose.IPFV-NPST.1SG	AUG-lose. PFV-PST.1SG	AUG-lose. IPFV-PST.1SG	
<i>févy-o</i>	<i>é-fiy-a</i>	<i>é-fevy-a</i>	‘I leave/left/used to leave’
leave.IPFV-NPST.1SG	AUG-leave. PFV-PST.1SG	AUG-leave. IPFV-PST.1SG	
		Derivation	
<i>çíno</i> ‘to pour’	<i>çísimo</i> ‘pouring, spillage’	<i>çimós</i> ‘juice’	
<i>xáno</i> ‘to lose’	<i>xásimo</i> ‘loss’	<i>xamós</i> ‘loss, disorder, hell’	
<i>líno</i> ‘to untie, solve’	<i>lísimo</i> ‘cancellation’	* <i>lisós</i>	<i>lísi</i> ‘solution’

A typical characteristic of derivational processes is that they structure the lexicon by forming word series and word families. Word series are sets of derivatives sharing a derivational affix and word families are classes of lexemes with a common stem being formally and semantically closely related. Due to its large inventory of derivational affixes, the SMG lexicon is rich in both of these structuring devices.

A word series containing the productive verbal (and nominal) prefix *kse-* with a reversative or intensive meaning is given in Table (4a). The derivatives quoted are morphotactically and morphosemantically transparent. Some members of the word family sharing the base *γraf-* ‘write’ (Table 4b) are less transparent morphotactically but more so semantically.

Table 4a. Word series with the prefix *kse-*

Verbal derivatives	Gloss	Base	Gloss
<i>kse-kurázome</i>	'to rest'	<i>kurázome</i>	'to get tired'
<i>kse-pléno</i>	'to wash off, rinse'	<i>pléno</i>	'to wash'
<i>kse-kinó</i>	'to set out, start'	<i>kinó</i>	'to move'
<i>kse-piðó</i>	'to shoot out, be ejected'	<i>piðó</i>	'to jump'

Table 4b. Word family based on the stem *graf-* 'to write' (Ralli 214: 45, 205)

Lemma	Gloss	Lemma	Gloss
<i>gráfo</i>	'to write'	<i>grafí</i>	'writing, writing system'
<i>gráma</i>	'letter'	<i>grafío</i>	'desk, office'
<i>grápsimo</i>	'way of writing'	<i>grafikós</i>	'picturesque, graphical'

3. Data

The computerized transcription of the audio-taped data of CS and CDS studied in the present chapter come from the rich Katis Corpus which Demetra Katis gathered from her elder daughter Anna growing up in Athens, Greece, typically interacting with her mother in a wide variety of natural speech situations.¹⁰ The section of the corpus chosen for the present analysis covers the child's age from 1;8 to 3;0 years. For the purpose of tracing the development of derivation in CS, the child's data have been subdivided into three subperiods (Table 5).

Table 5. Greek CS and CDS data of Anna in the Katis Corpus (word tokens)

Age	CS	CDS
1;8–1;11	15,886	52,988
2;0–2;6	31,352	70,355
2;7–3;0	22,966	28,843
Total	70,204	152,186

10. Since word stress is not indicated in the CHAT transcription of the Katis Corpus and there is no sound track, stress has only been indicated in examples taken from CDS but not from CS in the present article.

4. Results

4.1 The development of derivational prefixes and suffixes in Greek language acquisition

For the purposes of early language acquisition, emergent productive usage of a word series will be assessed if a given derivational affix is documented with at least three stems as in *kse-kubóno* ‘to unbutton’, *kse-pléno* ‘to rinse’, and *kse-skónizo* ‘to dust’ (see Examples 3a below). This will be taken to hold even in the absence of simple lexemes (e.g. *kub-í* ‘button’) or other derivatives with the same stem (e.g. *kub-ón-o* ‘to button’) in the data studied. One reason is that many bases of MG derivatives are inherited from Ancient Greek (AG) and either do not occur in simple lexemes in MG (e.g. MG *diavázo* ‘to read’ < AG *dia-viváz-o*, MG **vivázo*) or are rarely used and have a special meaning inaccessible to young children (e.g. *vál-o* ‘to shoot, attack’ in *iper-válo* ‘to excel, exaggerate’). The other reason is theoretical. Morphological relations among words are commonly taken to be based on their semantic and phonological similarity. In contrast to structuralist approaches, the usage-based, constructivist approach to language structure on which the present article is based, is goal-directed rather than source-directed, considering “complex forms as wholes or *gestalts* whose internal structure is discovered by relating them to similar morphological forms with similar grammatical function in the language. Packages of such similar form–function pairs constitute schemas and guarantee productivity” (Stephany 2012: 92). As pointed out by Bybee (2010: 23), “one advantage of this approach to morphological analysis is that it does not require that a word be exhaustively analysed into morphemes.” Thus, the analysis of derivatives into affix and base is no precondition for the productive use of derivatives containing a given derivational prefix or suffix and the emergence of a corresponding word series.

Although neologisms may emerge from the productive use of derivational affixes, they are no prerequisite for productiveness.¹¹ What is more important is that children learn to master the derivational patterns of the language being acquired by developing schemas such as *ksana* ____, consisting of the preverb *ksana* ‘again’ and an empty slot which can be filled by almost any dynamic verb (e.g. *ksana-pléno* ‘to wash again’, *ksana-páo* ‘to go again’). Schemas with a high type frequency, such as *ksana* ____ or *kse* ____, are stronger than those with a low type frequency (Bybee 1995: 430, 2001: 13). A schema for diminutive formation is ____*aki* (Thomadaki & Stephany 2007: 110).

11. As pointed out below (Section 5), derivative neologisms emerge much later in Greek language acquisition than in the age range studied in this paper.

4.1.1 *The development of derivational prefixes in CS*

In the first subperiod of observation from 1;8 to 1;11, there is only slight evidence for the child's awareness of the morphological complexity of prefixed derivatives. Reasons are that most of the derivatives carrying one of the prefixes *ana-*, *apo-*, *epi-*, *para-*, or *kse-* in the adult language are either partially deformed by the child (Examples 1c, 1d), semantically opaque (Examples 1b, 1d) or both (*parakaló* 'please' in 1d). The only example of a morphotactically and morphosemantically transparent prefixed verb form is *kse-kubóthike* 's/he unbuttoned' (Example 1e). Along with the denominal derivative *kubóno* 'to button' and its base *kubí* 'button', it forms a word family (for another word family see Example 1a).

(1) Prefixes in Anna's speech, 1;8–1;11

- a. *ana-*
ana-podes 'upside.down.F.NOM/ACC.PL' (*anápodos*, 1;10, 1 token; *pódi* 'foot', from 1;9, 41 tokens; *podaráki* 'foot.DIM', from 1;10, 6 tokens)
- b. *apo-*
apo-thiki 'storeroom, closet' (← *thiki* 'box, case, shell') (*apothiki*, 1;11, 1 token; base undocumented)
- c. *epi-*
(*e*)*pi-kijina* (for *epi-kíndina*) 'dangerous.NEUT.NOM/ACC.PL' (← *kíndinos* 'danger') (*epikíndinos*, 1;11, 1 token; base undocumented)
- d. *para-*
miskalo, *meskalo*, *pakalo*, *paka* (for *para-kaló*) 'to ask, please' (← *kaló* 'to ask, beg') (*parakaló* from 1;9, 14 tokens; base undocumented)
para-thiro, *para-thir-aki* 'window, window-DIM' (← etym. *para-thíra* 'side door', Babiniotis 1998: 1336) (also rendered as *paθathia*, *peraθiro*, *paθiraki*) (*paráthiro*, from 1;9, 11 tokens; *paraθiráki*, 1;11, 2 tokens)
- e. *kse-*
ke-kubiθike (for *kse-kubóthike* un-button.MEDIOPASS.PFV.PST.3SG) 'it got unbuttoned' (← *kub-ón-o* (button-DER-INFL) 'to button' ← *kubí* 'button') (*ksekubóno*, 1;11, 1 token; *kubóno*, 1;11, 7 tokens; *kubí*, from 1;8, 40 tokens)

In the first half of the third year (2;0–2;6), the numbers of derivational prefixes and prefixed lemmas increase. Still, evidence for word series and word families remains limited. With the exception of *pro-xtés* 'day before yesterday' (see below), most of the derivatives carrying one of the prefixes *a-*, *apo-*, *en-/em-* or *pro-* (Examples 2) are likely to be morphotactically and morphosemantically opaque for the child because they are either distorted (e.g. *vafanistike* for *em-fanístike* in 2c) or their base is not documented in either CS or CDS and may even come from AG (Examples 2a). Furthermore, the semantic relation between the derivative and its base may be indirect (e.g. *pro-xoráo* 'to move on' ← *xoráo* 'to fit' ← *xóros* 'place' in 2d). Although

there is only one token each of *pro-xtés* ‘day before yesterday, the other day’ and its base *xtes* ‘yesterday’, the base is documented two months before the derivative and the relation between the derivative and its base is not only morphotactically but also morphosemantically transparent so that the child may have recognized the structure of the derivative. A cautious interpretation of the data is that there is little more than formal evidence for a minimal word series *pro-xoráo* and *pro-xtés*.

(2) Opaque prefixed derivatives in Anna’s speech, 2;0–2;6

a. *a-*

a-taksia ‘disorder’ (← *táksi* ‘order’) (2;6, 1 token, base undocumented)
noisia, anisies (for *a-noisía, a-noisés* ‘nonsense’ (SG, PL) (← *nóisi* ‘intellect’) (2;6, 1 token, base undocumented)
apurimeni (for *aporiméni* ‘helpless.F.NOM/ACC.SG’) (< AG *aporo*) (2;6, 1 token)

b. *apo-*

(*a*)*po-fasise* ‘s/he decided’ (← *apó-fasi* ‘decision’) (*apofasízo*, 2;6, 2 tokens; *apófasi* ‘decision’ undocumented)
apo-kries (lit. ‘without meat’ (*kréas*)) ‘Lent’ (*apókries*, 2;6, 1 token, base undocumented)

c. *en-/em-*

vafanistike (for *em-fanístike* ‘she appeared’) (< AG *emfanís*) (*emfanízome*, 2;5, 1 token; *fénome* ‘to appear, show’, from 2;0, 33 tokens)

d. *pro-*

na plolavume (for *na pro-lávume*) ‘so that we will manage’ (*pro-lavéno* ← *lam-váno/lavéno* ‘to get, receive’) (*prolavéno*, 2;2, 1 token, base undocumented)
pro-xoráo ‘to move on, continue’ (← *xor-áo* ‘to fit’ ← *xór-os* ‘place’) (*proxoráo*, from 2;2, 4 tokens; *xoráo* from 2;0, 37 tokens; *xóros*, 2;5, 2 tokens)
pro-xtes (← *xtes* ‘yesterday’) ‘day before yesterday, the other day, recently’ (*proxtés* 2;6, 1 token; *xtes*, 2;4, 1 token)

Stronger evidence for the development of a relation between derivatives and their bases during the period from 2;0 to 2;6 comes from the prefixes *kse-*, with a reverse or, more rarely, an intensive meaning (Ralli 2004: 3), and *para-*, which may express the meaning of exaggeration in comparison with the base (Ralli 2014: 243). The prefix *kse-* is usually combined with accomplishment verbs (e.g. *kse-skonízo* ‘to dust’, lit. ‘to remove dust’) but not with activity or stative verbs (Ralli 2004). This condition is fulfilled by the three derivatives occurring in Anna’s speech from 1;11 to 2;5 (Examples 3a), which can be taken to be morphotactically and morphosemantically transparent for the child. Since the bases and roots of *kse-kubóno* ‘to unbutton’ and *kse-pléno* ‘to rinse’ are found in Anna’s data and *kse-* signals reversing or undoing some action or event, the child may have recognized the meaning

of the prefix and the morphological complexity of these verbs. A comparison of *kse-skonízo* ‘to dust’ with these verbs may also lead to the analysis of the former. In the subperiod 2;0–2;6, the word family *kse-kub-ón-o* ‘to unbutton’, *kub-ón-o* ‘to button’, *kub-í* ‘button’ already found in the developmental period 1;8–1;11 is complemented by the word series *kse-kubóno* ‘to unbutton’, *kse-pléno* ‘to rinse, wash out’, *kse-skonízo* ‘to dust’. By contrast, the structure of the verb *ksipnó* ‘to wake up’ and its opaque relation to the root of *ípnos* ‘sleep’ is unlikely to be recognized by the child (see Examples 3a).

Since both *para-berdétikan* ‘they got very confused’ and its base *berdétikan* ‘they got confused’ occur in immediate vicinity to each other in a file from 2;4, the morphotactic and morphosemantic transparency of the derivative may have been recognized by the child (see Examples 3b). Other instances of the word series constituted by the prefix *para-*, namely *para-míthi* ‘fairytale’, *para-mána* ‘foster mother’, and *pará-thiro* ‘window’ are only morphotactically transparent since the prefix does not carry the meaning of intenseness or exaggeration (Examples 3b). There is evidence of the morphotactic transparency of the derivative *para-mána* ‘foster mother’ since its base already occurs in the preceding developmental period; however, just as the preceding examples, it is morphosemantically opaque due to the diversity of uses of *pará*, which functions as a prefix and preposition in AG.¹²

The word families *an-e-véno* ‘to go up’, *kat-e-véno* ‘to go down’, and *an-e-vázo* ‘to put up’, *kat-e-vázo* ‘to put down’ first encountered in the first half of Anna’s third year are interesting because both members contain the same prefixes carrying a contrastive meaning and forming the word series *anevéno*, *anevázo* and *katevéno*, *katevázo* (Examples 3c). It should be mentioned that the forms *an-e-* and *kat-e-* derive from past tense forms of these verbs including an augment (e.g. *an-é-vaz-a* (up-AUG-put.IPFV-PST.1SG) ‘I put up’, *kat-é-vaz-a* (down-AUG-put.IPFV-PST.1SG) ‘I put down’). Since there are more than twenty tokens covering different grammatical forms of each of the two verbs *anevéno* ‘to go up’ and *katevéno* ‘to go down’ to be found in Anna’s data from 2;0 on, it seems quite likely that the semantic and formal relations between the two verbs have been recognized in spite of the fact that their base *véno* ‘to develop’, which is of restricted use in SMG, does not occur in the data. Notwithstanding the fact that there is only a single formally distorted token of *anevázo* ‘to put up’ contrasted with two tokens of *katevázo* ‘to put down’

12. The relationship of the derivatives *paramána* ‘foster mother’, *paramíthi* ‘fairytale’, and *paráthiro* ‘window’ to the meaning of their two elements is obscure due to the diversity of uses of *para* in AG denoting ‘near’, opposition, and relation among others (Jannaris 1968 [1897]: 388–391; Humbert & Kourmoulis 2002 [1957]: 302–303, 326–328; see also Babiniotis 2011). This not only applies to children but also sometimes to adults (D. Theophanopoulou-Kontou, p.c.).

documented at 2;0, it is not to be excluded that the formal and semantic relation of these two verbs to their common base *vázo* ‘to put’ (264 tokens from 1;9) as well as to the pair *anevéno* ‘to go up’ and *katevéno* ‘to go down’ with the same prefixes may have been recognized. It is unlikely, however, that the derivative *kata-lavéno* ‘to understand’ (Example 3c) is morphotactically or morphosemantically transparent for the child since the semantic relation of the prefix to the derivative is not apparent and the base is undocumented. Other verbs with the prefix *kata-* will only appear in the second half of the third year (see Examples 4b below).

- (3) The prefixes *kse-*, *para-*, and *ana-* vs. *kata-* in Anna’s speech, 2;0–2;6
- a. *kse-*
kse-kuboso ‘un-button.PFV.NPST.1SG’ (← *kub-ón-o* ‘to button’ ← *kub-í* ‘button’) (*ksekubóno*, from 1;11, 3 tokens; *kubóno*, from 1;11, 26 tokens; *kubí*, from 1;8, 53 tokens)
kse-plinis ‘rinse.PFV.NPST.2SG’ (← *pléno* ‘to wash’) (*ksepléno*, 2;5, 1 token; *pléno* from 2;0, 28 tokens)
kse-skonise ‘dust.PFV.PST.3SG’ (← *skón-íz-o* ‘to cover with dust’ ← *skón-i* ‘dust’) (*ksekonízo*, 2;2, 1 token; bases undocumented)
ksipnisa ‘wake.up.PFV.PST.1SG’ (and other grammatical forms) (← *ípnos* ‘sleep’) (*ksipnó*, from 2;2, 12 tokens; base undocumented)
- b. *para-*
para-bedefiti (for *para-berdéftikan* ‘they got very confused’ (← *berδ-év-ome* ‘to get confused’)) (*paraberévome*, 2;4, 1 token; *berévome*, 2;4, 2 tokens)
para-miθi ‘fairytale’ (← *miθ-os* ‘myth, legend’) (*paramiθi*, 2;2, 3 tokens; base undocumented)
para-mána ‘foster mother’ (← *mána* ‘mother’) (*paramána*, 2;2, 3 tokens; *mána*, 1;11, 1 token, 2;4, 3 tokens)
para-kalo ‘please’ (2;4, 1 token; base undocumented) (see Examples 1d above)
para-θiro ‘window’ (from 1;9, 7 tokens from 2;0–2;6) (see Examples 1d above)
- c. *ana-/an-* vs. *kata-/kat-*
aneveno ‘to go up’ vs. *kateveno* ‘to go down’ (← *véno* ‘to develop’) (*anevéno*, from 2;0, 22 tokens; *katevéno*, from 2;0, 22 tokens, base undocumented)
(a)nevasi ‘put.up.PFV.NPST.3SG’ vs. *katevas/katevasane* ‘put.down.PFV.PST.2SG/3PL’ (← *vázo* ‘to put’) (*anevázo*, 2;0, 1 token; *katevázo*, 2;0, 2 tokens; *vázo*, from 1;9, 264 tokens)
kata-lavi ‘understand.PFV.NPST.3SG’ (← *lavéno/lamváno* ‘to receive’) (*kat-alavéno*, from 2;4, 2 tokens; base undocumented)

Comparing Anna's development of derivation in the first subperiod of observation from 1;8 to 1;11 to that in the first half of her third year, the most important development is that there is now evidence for four word series (Examples 3) in spite of the fact that some of their members are only morphotactically transparent. These series contain from two to even five members. In addition to the three-member word family already present since the first period, four further ones consisting of two members have developed (Examples 3).

In the last subperiod of observation (2;7–3;0), the inventory of prefixes increases by one item, *eks-*, and a few further derived lemmas with the old prefixes are added to Anna's lexicon, enlarging the respective word series. There is now evidence for the productive formation of derivatives carrying the prefix *kse-* but still only limited proof of the productivity of *kata-*, *para-*, and *pro-*.

The verbs *eks-afanízome* 'to disappear', first documented in one token at 2;6, and its antonym *em-fanízome* 'to appear', one token of which is recorded at 2;5 and 2;7, form a word family together with *fanerónome* 'to appear' (1 token, 3;0) and morphotactically opaque *fénome* 'to appear, seem, to be visible' (11 tokens from 2;7 on).¹³ The prefix *apo-* expands the respective word series by the lemmas *apó-jevma* 'afternoon' (← *jevma* '(midday) meal') and *apo-stimeno* (for *aposteriméno*) 'robbed' (*apo-steró* ← *steró* 'to rob') added to earlier *apo-fasízo* 'to decide' and *apó-kries* 'Lent' (see Examples 2b above). Although the morphotactic structure of at least some of these derivatives (e.g. *apójevma*) may be recognized by the child (in spite of the fact that the base *jevma* 'meal' is not documented in either CS or CDS), they will most likely remain semantically opaque since their base forms are either learned words (*apo-steró*) or they are formed by opaque stems (*apó-kries*). Also, *apo-* does not have the same meaning in these structures. The prefix means 'after' in *apó-jevma* (lit. after meal) 'afternoon' but 'without' in *apo-stiriméno* 'robbed' and *apó-kries* (lit. without meat) 'Lent' and finally 'entirely' in *apo-fasízo* 'to decide' (Charalambakis 2014).

The prefixes *kse-*, *kata-*, *para-*, and *pro-* yield three- to seven-member word series making the most important contributions to the child's lexicon in this domain (Examples 4). However, as far as the derivatives *kata-píno* 'to swallow', *kata-skevázó* 'to make, produce' (Examples 4b) in addition to *kata-lavéno* 'to understand' (Examples 3c above) are concerned, there is hardly any evidence for their semantic transparency since the meaning of these verbs is not semantically associated with that of their prefix. The same is true of *para-kaló* 'to ask, beg' (Examples 4c).¹⁴

13. The opacity results from AG *ablaut*.

14. There is no apparent semantic relation between the prefix and the base of *para-kaló* 'to beg, ask, please' due to multiple meanings of this prefix in AG.

(4) Productive prefixes in Anna's speech, 2;0–2;6 and 2;7–3;0

a. *kse-*

2;0–2;6 (see Examples 3a above)

kse-kubóno 'to unbutton', *kse-pléno* 'to rinse', *kse-skónízo* 'to dust', *ksipnáo* 'to wake up'

2;7–3;0

kse-pléno 'to rinse' (2;7, 2 tokens) (← *pléno* 'to wash', from 2;7, 4 tokens)*kse-kinó* 'to set out, start' (2;7, 3 tokens) (← *kinó* 'to move', undocumented)*kse-kurázome* 'to rest' (2;10, 1 token) (← *kurázome* 'to get tired', from 2;7, 6 tokens)*kse-pidó* 'to splash' (2;8, 1 token) (← *pidó* 'to jump', from 2;7, 22 tokens)*ksi-prízo* (for *kse-prízo*) 'to recede (swelling)' (2;11, 1 token) (← *prízo* 'to make sth. swell', undocumented)*ksi-lóno* 'to undo, remove (nail)' (2;10, 1 token) (← *ílos* 'nail' < AG *ílos* 'nail', undocumented)*ksipnáo* 'to wake up' (from 2;7, 13 tokens) (← *ípnos* 'sleep', from 2;8, 3 tokens)b. *kata/kat-*

2;0–2;6 (see Examples 3c above)

kat-e-vázo 'to put down', *kat-e-véno* 'to go down', *kata-lavéno* 'to understand'

2;7–3;0

kat-e-véno 'to go down' (from 2;7, 14 tokens, base undocumented)*kata-píno* 'to swallow' (2;7, 2 tokens) (← *píno* 'to drink', from 2;9, 3 tokens)*kata-skev-áz-o* 'to make, produce' (2;11, 1 token) (← *skév-os* 'device, instrument', *skév-i* 'dishes, crockery', undocumented bases)c. *para*¹⁵

2;0–2;6 (see Examples 3b above)

para-kaló 'please', *para-berdévome* 'to get very confused', *para-mána* 'foster mother', *para-míthi* 'fairytale', *pará-θiro* 'window'

2;7–3;0

para-kaló 'to ask, beg' (from 2;8, 6 tokens) (← *kaló* 'to invite', undocumented)*para-léo* 'to exaggerate' (3;0, 1 token) (← *léo* 'to say', from 2;7, 337 tokens)*para-míthi* 'fairytale' (from 2;7, 11 tokens) (← *míthos* 'myth, legend', undocumented)*para-míθ-aki* 'fairytale-DIM' (from 2;7, 2 tokens)*pará-θiro* 'window' (from 2;10, 2 tokens)*para-θir-áki* 'window-DIM' (3;0, 2 tokens)

15. In the adult language, the prefix *para-* has two meanings, exaggeration (e.g. *para-léo* 'to exaggerate') and proximity (e.g. *para-psixología* 'parapsychology', *para-stratíotikós* 'paramilitary').

- d. *pro-*
 2;0–2;6 (see Examples 2d above)
pro-lavéno ‘to manage’, *pro-xoráo* ‘to move on, continue’, *pro-xtés* ‘day before yesterday, recently’
 2;7–3;0
pro-lavéno ‘to manage’ (from 2;9, 4 tokens) (base undocumented)
pro-xtés ‘day before yesterday, recently’ (from 2;9, 2 tokens) (← *xtes* ‘yesterday’, 2;11, 1 token)

There are 5 at least morphotactically transparent word series of prefixed derivatives to be found from 2;0 to 2;6 and 5 partially overlapping ones from 2;7 to 3;0. They consist of 2 to 5 members in the first subperiod and of 2 to 7 members in the second (Examples 2 and 3 above). Series of prefixed derivatives with *kse-* and *para-* have most members (*kse-*: 4 and 7 members in the two subperiods, respectively; *para-*: 5 and 6 members, respectively). At the end of Anna’s third year, the number of prefixed word series amounts to 9 cumulatively.

As far as word families with a prefixed derivative as one of their members are concerned, their number amounts to 9 from 2;0 to 2;6 and 7 from 2;7 to 3;0. Cumulatively, there are 14 different word families of this type to be found at the end of the entire period of observation. While the number of word families with a prefixed member exceeds that of prefixed word series, most of the word families are limited to two members and only four of them comprise three members (Examples 5).

- (5) Word families with prefixed members in Anna’s speech, 2;0–2;6 and 2;7–3;0
- a. 2;0–2;6 (see Examples 2 and 3 above)
anevázo ‘to put up’ (2;0, 1 token), *katevázo* ‘to put down’ (2;0, 2 tokens), *vázo* ‘to put’ (from 1;9, 264 tokens)
ksekubóno ‘to unbutton’ (from 1;11, 3 tokens), *kubóno* ‘to button’ (from 1;11, 26 tokens), *kubí* ‘button’ (from 1;8, 53 tokens)
proxoráo ‘to move on, continue’ (from 2;2, 4 tokens), *xoráo* ‘to fit’ (from 2;0, 37 tokens), *xóros* ‘place’ (2;5, 2 tokens)
emfanízome ‘to appear’ (from 2;5, 2 tokens), *eksafanízome* ‘to disappear’ (2;6, 1 token), *fénome* ‘to appear, show’ (from 2;0, 33 tokens)
anevéno ‘to go up’ (from 2;0, 22 tokens), *katevéno* ‘to go down’ (from 2;0, 22 tokens)
proxtés ‘day before yesterday, recently’ (2;6, 1 token), *xtes* ‘yesterday’ (2;4, 1 token)
ksepléno ‘to rinse’ (2;5, 1 token), *pléno* ‘to wash’ (from 2;0, 28 tokens)
paraberdévome ‘to get very confused’ (2;4, 1 token), *berdévome* ‘to get confused’ (2;4, 2 tokens)
paramána ‘foster mother’ (2;2, 3 tokens), *mána* ‘mother’ (1;11, 1 token; 2;4, 3 tokens)

- b. 2;7–3;0 (see Examples 4 above)
- ksepléno* ‘to rinse’ (2;7, 2 tokens), *pléno* ‘to wash’ (from 2;7, 4 tokens)
- ksekurázome* ‘to rest’ (2;10, 1 token), *kurázome* ‘to get tired’ (from 2;7, 6 tokens)
- ksepidó* ‘to splash’ (2;8, 1 token), *pidó* ‘to jump’ (from 2;7, 22 tokens)
- ksipnáo* ‘to wake up’ (from 2;7, 13 tokens), *ípnos* ‘sleep’ (from 2;8, 3 tokens)
- katapíno* ‘to swallow’ (2;7, 2 tokens), *píno* ‘to drink’ (from 2;9, 3 tokens)
- paraléo* ‘to exaggerate’ (3;0, 1 token), *léo* ‘to say’ (from 2;7, 337 tokens)
- proxtés* ‘day before yesterday, recently’ (from 2;9, 2 tokens), *xtes* ‘yesterday’ (2;11, 1 token)

As mentioned above, the adverb *ksaná* ‘again’ may function as a preverb and forms a very productive word series in SMG. This also applies to CS (and CDS, see below). For the adult language, the lexicon of the MG language by Babiniotis (1998: 1218–1219) lists a series of 59 morphotactically and morphosemantically transparent verbs (and 2 nouns) in a footnote in addition to the lexical entries containing this preverb. The reason for taking *ksana* into consideration in this study of derivation is that the properties of preverbs, be they prefixes such as *kse-* or words such as *ksaná*, “display properties that cut across the word-formation processes of prefixation and compounding” (Ralli 2004: [1]). *Ksaná* is an adverb that can appear as a phrasal element or as a first constituent of compounds, “with more or less the same adverbial meaning” (Ralli 2004: [3]). An important difference between the preverbs *kse-* and *ksana-* is that according to Ralli (2004: [17]) the former is attached to word stems while the latter combines with an inflected word form (e.g. [*kse-dín*]-*o* ‘to undress’ vs. *ksana*-[*dín*-*o*] ‘to dress again’).

Word formations with the preverb *ksana* ‘again’ have been included in the present study because the adverb *ksaná* does not occur in CS during the observational period.¹⁶ Besides, it is by far the most productive element for forming complex verbs (Examples 6). Since *ksaná* does not occur as a free form in CS and word series are typical of derivational formations, this element may be taken to function as a prefix in early child Greek (as it does in CDS; see below). Its productiveness may be due to the looser relation between the preverb and its base, which is a word rather than a stem, resulting in morphotactic transparency. In addition, the meaning of the base remains unchanged so that complex verbs carrying the preverb *ksana* are also morphosemantically transparent. The word series with *ksana* formations occurring in the second half of Anna’s third year comprises 13 members and is thus nearly twice as large as the *kse-* series in the same period (7 members). With one exception, all bases are documented in the same period of development, most of them in many more tokens than the *ksana-* formations.

16. Interestingly, Holton, Mackridge and Philippaki-Warburton (2016: 246) list *ksana-* as a verbal prefix without further discussion.

- (6) The *ksana* ‘again’ word series in Anna’s speech, 2;0–2;6 and 2;7–3;0
- a. 2;0–2;6
ksana-érxome ‘to come again’ (from 2;2, 2 tokens) ← *érxome* ‘to come’ (from 2;0, 28 tokens)
 - b. 2;7–3;0
ksana-kitázo ‘to look again’ (2;8, 1 token) ← *kitázo* ‘to look’ (from 2;7, 47 tokens)
ksana-vlépo ‘to see again’ (2;9, 1 token) ← *vlépo* ‘to see’ (from 2;7, 92 tokens)
ksana-éxo ‘to have again’ (2;7, 2 tokens) ← *éxo* ‘to have’ (from 2;7, 79 tokens)
ksana-févyo ‘to leave again’ (from 2;9, 2 tokens) ← *févyo* ‘to leave’ (from 2;7, 37 tokens)
ksana-ftjáxno ‘to make again’ (2;10, 1 token) ← *ftjáxno* ‘to make’ (from 2;8, 27 tokens)
ksana-jirízo ‘to come back again’ (2;11, 4 tokens) ← *jirízo* ‘to come back’ (from 2;7, 15 tokens)
ksana-kléo ‘to cry again’ (2;10, 1 token) ← *kléo* ‘to cry’ (from 2;7, 30 tokens)
ksan-anevéno ‘to go up again’ (2;11, 1 token) ← *anevéno* ‘to go up’ (from 2;7, 35 tokens)
ksana-páo ‘to go again’ (2;10, 1 token) ← *páo* ‘to go’ (from 2;7, 290 tokens)
ksan-archízo ‘to start again’ (2;7, 1 token) ← *archízo* ‘to start’ (undocumented)
ksana-thélo ‘to want again’ (2;11, 1 token) ← *thélo* ‘to want’ (from 2;7, 135 tokens)
ksana-vázo ‘to put again’ (3;0, 1 token) ← *vázo* ‘to put’ (from 2;7, 165 tokens)
ksana-xáno ‘to lose again’ (2;7, 1 token) ← *xáno* ‘to lose’ (2;9, 1 token)

4.1.2 The development of derivational suffixes in CS

Ralli (2014: 147–157) divides 67 derivational suffixes of SMG according to their nominal, verbal or adjectival bases, the parts of speech to which the resulting derivatives belong, and the semantic categories of derivative nouns. Derivational suffixes producing nouns are the most numerous and mostly convert nouns into other nouns (Mackridge 1985: 319), such as agent nouns referring to persons (e.g. *tiropit-ás* ‘cheese-pie maker/seller’ ← *tirópita* ‘cheese pie’), to objects (e.g. *psom-jéra* ‘bread-basket’ ← *psomí* ‘bread’) or places (e.g. *pítsa-ría* ‘pizzeria’ ← *pítsa* ‘pizza’; *tiropitá-ðiko* ‘cheese pie shop’ ← *tiropitás* ‘cheese pie maker’), nouns deriving feminines from masculine nouns by motion (e.g. *vasílisa* ‘queen’ ← *vasiljás* ‘king’, *maθíttria* ‘schoolgirl’ ← *maθitís* ‘schoolboy’) (Mackridge 1985: 160–161; see also Thomadaki 1988: 55–58). Some highly productive suffixes convert nouns or adjectives into verbs (Ralli 2014: 147; see also Mackridge 1985: 322–323) (e.g. *-íz-o*, *-ón-o*, *-áz-o*, *-év-o*, *-én-o* as in *psar-év-o* ‘to fish’ ← *psár-i* ‘fish’, *kaθar-íz-o* ‘to clean’ ← *kaθar-ós* ‘clean’).

Due to limitations of space and partially also for lack of data, only a few derivational suffixes occurring in early Greek CS can be dealt with in this article. These

are the verbalizing suffixes *-íz-o*, *-ón-o*, *-áz-o*, *-év-o*, and *-én-o*, the adjectival suffixes *-ik-os* and *-in-os*, and the more productive nominal suffix *-tis* deriving agentive nouns (Thomadaki 1988: 37) as well as feminines in *-isa* derived by motion from masculines. As mentioned above, diminutive suffixes, which are the most productive nominal derivational suffixes occurring in child-centered situations, are excluded from the present article since they have been studied by Thomadaki and Stephany (2007).

Verbal derivatives containing one of the suffixes *-íz-o*, *-ón-o*, *-áz-o*, *-év-o*, or *-én-o* occur from the beginning through the end of the observation period (1;8–3;0) and increase considerably in the course of development (Examples 7). Verbs with the by far most productive verbalizing suffix *-íz-o*¹⁷ as well as the suffix *-ón-o* are most numerous and those with *-én-o* least frequent in each of the three subperiods. The nominal or adjectival bases of these derivative verbs are found in CS in only 30% to 40% of cases. Furthermore, bases often emerge later than derivatives and are in most cases much less frequently used than the latter. It may therefore be hypothesized that the suffixed verbal lemmas are semantically opaque for the child and their derivational structure has not yet been detected.

- (7) Examples of verbs with the derivational suffixes *-íz-o*, *-ón-o*, *-áz-o*, *-év-o*, and *-én-o* in Anna's speech
- a. *-íz-o* (1;8–1;11, 16 lemmas; 2;0–2;6, 30 lemmas; 2;7–3;0, 23 lemmas)
arch-íz-o 'to begin' (← *arch-i* 'beginning'), *kaθar-íz-o* 'to clean' (← *kathar-ós* 'clean'),
skup-íz-o 'to sweep' (← *skúp-a* 'broom'), *xten-íz-o* 'to comb' (← *xtén-a* 'comb')
 - b. *-ón-o* (1;8–1;11, 13 lemmas; 2;0–2;6, 15 lemmas; 2;7–3;0, 25 lemmas)
klið-ón-o 'to lock' (← *klið-i* 'key'), *kri-ón-o* 'to be cold' (← *krió.N* 'cold'),
meyal-ón-o 'to grow' (← *meyál-os* 'big')
 - c. *-áz-o* (1;8–1;11, 3 lemmas; 2;0–2;6, 14 lemmas; 2;7–3;0, 14 lemmas)
fon-áz-o 'to shout' (← *fon-i* 'voice'), *nist-áz-o* 'to feel sleepy' (← *níst-a* 'sleepiness')
 - d. *-év-o* (1;8–1;11, 4 lemmas; 2;0–2;6, 7 lemmas; 2;7–3;0, 11 lemmas)
majir-év-o 'to cook' (< AG base), *xor-év-o* 'to dance' (← *xor-ós* 'dance')
 - e. *-én-o* (1;8–1;11, 1 lemma; 2;0–2;6, 4 lemmas; 2;7–3;0, 4 lemmas)
maθ-én-o 'to learn' (← *máθima* 'lesson', *maθitís* 'pupil'), *zest-én-o* 'to heat' (← *zest-ós* 'warm', *zést-i* 'heat')

The assumption that verbal derivatives of this group are opaque for the child is additionally supported by the circumstance that they occur quite frequently

17. Thomadaki (1988: 77) explains the high productivity of this verbalizing suffix by the fact that it is combinable with bases belonging to any part of speech.

already before the end of the child's second year (1;8–1;11, 37 lemmas; 2;0–2;6, 70 lemmas; 2;7–3;0, 77 lemmas), quite in contrast to prefixed verbs, which begin to develop only during the third year (see Section 4.1.1 above). This may be explained by the fact that, in contrast to verbal prefixes, verbal derivational suffixes are semantically empty¹⁸ and merely function as a verbalizing device (Joseph & Philippaki-Warbuton 1987: 219). The resemblance between the members of word series constituted by verbs containing one of the suffixes *-íz-o*, *-ón-o*, *-áz-o*, *-év-o*, or *-én-o* is therefore mainly morphotactic.¹⁹ No productive enlargement of the lexical inventory is possible with the help of these suffixes because most stems are combinable with only one particular suffix (Examples 8).²⁰

- (8) a. *kaθar-íz-o* 'to clean-DER-INFL' but **kaθar-ón/áz/év/én-o*
 b. *mal-ón-o* 'to scold-DER-INFL' but **mal-íz/áz/én/év-o*

Morphotactic transparency of these verbal derivatives may, however, be obscured by their inflectional characteristics. In particular, the form of the derivational suffix changes in diverse ways in the perfective aspectual form (Examples 9).²¹ These changes do not only occur across word series (Examples 9a vs. b) but even within one and the same series (Examples 9c vs. d).

- (9) Imperfective and perfective forms of suffixed verbal derivatives
 a. *arch-íz/ís-o* 'begin-DER.IPFV/PFV-NPST.1SG'
 b. *xor-év/éps-o* 'dance-DER.IPFV/PFV-NPST.1SG'
 c. *etim-áz/ás-o* 'prepare-DER.IPFV/PFV-NPST.1SG'
 d. *fon-áz/áks-o* 'call-DER.IPFV/PFV-NPST.1SG'

The fact that many simple verbs form their perfective stems on *-is-* in the same way as verbs derived by *-íz-o* may furthermore interfere with the identification of members of the word series of derived verbs (Example 10a (derivative) vs. 10b and 10c (simple verbs)).

18. However, Thomadaki (1988: 79–80) distinguishes several semantic classes of verbs derived by *-ízo* by taking the meaning of the bases into consideration.

19. For details concerning the bases of these verbal derivatives and their meanings in SMG see Thomadaki (1988: 77–88).

20. There are however some exceptions such as *jir-íz-o* (around-DER.IPFV-NPST.1SG) 'to turn (sth.), to return' and *jir-év-o* (around-DER.IPFV-NPST.1SG) 'to look for sth.' (Thomadaki 1988: 76).

21. On perfective and imperfective verb stem formation in SMG see Holton, Mackridge and Philippaki-Warbuton (2016: 141–147) and in SMG as well as Greek CS see Stephany (1985: 67–82).

- (10) Perfective forms of derived vs. simple verbs
- a. *arch-íz/ís-o* ‘begin-DER.IPFV/PFV-NPST.1SG’
 - b. *xtíz/xtís-o* ‘build.IPFV/PFV-NPST.1SG’
 - c. *xtip-ó* ‘hit.IPFV-NPST.1SG’ vs. *xtip-ís-o* ‘hit-PFV-NPST.1SG’

Turning to suffixed adjectival derivatives, only *-ik-os* and *-in-os* have been selected for the present study because no more than 6 of the other 14 adjectival suffixes (Ralli 2014: 152–153) are documented by 1 or at most 2 lemmas in the second half of Anna’s third year. Adjectival stems ending in *-ik-* or *-in-* are typically derived from nouns (Examples 11). Their number increases from 1;8 to 3;0, but the *-ik-* word series is more frequent than the *-in-* series (2;7–3;0, *-ik-* 22 lemmas, *-in-* 12 lemmas). It is rather doubtful that the two word series will be transparent for the child. Although the suffixes do not change their form in different inflections of the adjectives, they are not easily separable because they do not occupy a prominent position in the word, being situated between the stem and the inflectional ending (Examples 11). In addition, the bases of the derivatives are often undocumented or occur rarely in CS. Most importantly, the two suffixes do not have a proper meaning and are limited to the function of deriving adjectives. Since every nominal stem is compatible with only one of the adjectival derivational suffixes, most of the respective word families in CS have only two members, the derivative and its base. An exception is the word family *vrómikos*, *vromerós* ‘dirty’, *vromjá* ‘dirt’, *vromó* ‘to stink’ all members of which are found in CS.

- (11) Adjectives derived by the suffixes *-ik-os* or *-in-os* (-DER-INFL)
- a. *vróm-ik-os* ‘dirty’ (← *vróma* ‘dirt’), *nist-ik-ós* ‘sober’ (← *nistía* ‘sobriety, fasting’)
 - b. *kók-in-os* ‘red’ (< AG *kókos* ‘(coffee) bean’), *ksíl-in-os* ‘wooden’ (← *ksílo* ‘wood’)

In CS, nouns derived by suffixes from other nouns or occasionally verbs are even more scarce than derived adjectives. There is only one example of a derivative noun referring to places found at 2;10 (Example 12a). Derived nouns referring to objects occur in 1 or at most 2 exemplars in each of the three subperiods (Examples 12b). Nominal derivatives registered a little more frequently are certain agent nouns and feminines derived from masculines by motion. Agent nouns with the suffix *-tis* only emerge in Anna’s third year amounting to 9 lemmas altogether (Examples 12c). Although feminines derived from masculines carrying the suffix *-is-a* emerge before the end of the second year and occur throughout the observation period, they are only documented by 4 lemmas cumulatively (Examples 12d). Feminines with the suffix *-tria* emerge in the second half of the child’s third year and total 3 lemmas (Examples 12d).

(12) Derived nouns

a. Nouns referring to places

2;10: *paputsi-dik-o* [for *paputsá-dik-o*] ‘shoe store’ (← *papútsi* ‘shoe’)

b. Nouns referring to objects

1;10: *kafet-jer-a* ‘coffee pot’ (← *kafés* ‘coffee’)²²2;5: *ban-jer-a* ‘bathtub’ (← *bánjo* ‘bath, bathroom’)

c. Agent nouns

2;7: *trayud-ist-is* ‘singer’ (← *trayudó* ‘to sing’, *trayúdi* ‘song’)2;2: *astr-o-naf-ti-s* ‘astronaut’ (← *náftis* ‘sailor’)2;5: *klef-ti-s* ‘thief’ (← *klév-o* ‘to steal’)

d. Feminines derived from masculines by motion

1;11: *vasil-is-a* ‘queen’ (← *vasilj-ás* ‘king’)2;11: *manav-is-a* ‘female greengrocer’ (← *manáv-is* ‘male greengrocer’)2;7: *xoref-tri-a* ‘female dancer’ (← *xoref-tí-s* ‘male dancer’)2;10: *komo-tri-a* ‘female hairdresser’ (← *komó-ti-s* ‘male hairdresser’)

Due to the scanty data of derived nouns, it is doubtful whether the respective word series or word families have been recognized by the child. The only evidence for possible “nuclear” word families are lexemes whose bases emerge earlier or in the same month as the derivative (e.g. *vasílisa* ‘queen’ and *vasiljás* ‘king’, both at 1;11; *kléftis* ‘thief’ and *klévo* ‘to steal’, both at 2;5). The recognition of word series of suffixed nominal derivatives may be somewhat facilitated by the fact that their inflectional endings rarely change because most of them are used in only one or two case–number forms. The overall ending of the forms including the derivational and the inflectional suffix therefore remains more or less constant (e.g. *kléf-tis* ‘thief’ (NOM.SG) and *kléf-ti* (ACC.SG) but *kléf-tes* (NOM/ACC.PL)).

4.2 The usage of derivational prefixes and suffixes in CDS in comparison with CS

Since it can be hypothesized that derived lemmas occurring in CDS may only influence CS if they are met with frequently enough, this study will focus on derivational processes which have developed to some extent in CS, namely certain prefixes as well as verbal, adjectival, and nominal suffixes (see Section 4.1).

In contrast to CS, prefixed derivatives, most of them verbs and nouns, are amply documented in CDS from the first subperiod on (1;8–1;11). While lemmas with one of the prefixes *kata-*, *para-*, or *pro-* total 25 in the entire period of observation, verbs

22. Since none of the examples *psist-jér-a* ‘grill’, *tost-jér-a* ‘toaster’, or *frut-jér-a* ‘fruit bowl’ is documented in CS or CDS, the analysis of *kafet-jér-a* ‘coffee pot’ with a stem-final /t/ is problematic.

derived by the preverbs *kse-* or *ksana* demonstrate much more lexical diversity, amounting to 17 lemmas each. However, as far as the number of tokens of prefixed derivatives is concerned, formations with *kata-*, *para-*, or *pro-* by far exceed those with *kse-* or *ksana*. Tokens are not evenly distributed among derivatives. 77% of the overall number of the first group (*kata-*, *para-*, *pro-*, $n = 1,161$) come from the verbs *katalavéno* ‘to understand’ ($n = 276$), *parakaló* ‘to ask, beg, please’ ($n = 285$), or *proséxo* ‘to be careful, watch out’ ($n = 333$). By contrast, the overall number of tokens of the *kse-/ksana* group only amounts to 128 and the usage of individual verbs is more evenly distributed than in the other group. While the greater lexical diversity of the *kse-/ksana* group and the ensuing larger word series in CDS will promote productivity in CS, the greater frequency of tokens of the word series of the other group will support entrenchment. Thus, there is a total of 27 lemmas to be found in the *kse-/ksana* group in CS in contrast to only 11 lemmas in the *kata-/para-/pro-* word series. Productivity of the *kse-/ksana* group is also furthered by the morphotactic and morphosemantic transparency of most of these derivatives used in CDS. This situation contrasts with the morphotactic and morphosemantic opacity of many derivatives of the *kata-/para-/pro-* group.

As far as the grammatical status of *ksaná* ‘again’ in CDS is concerned, it is interesting to note that this form, just as in CS, exclusively occurs as a preverb rather than an adverb (38 tokens). This supports the analysis of this form as a prefix in CS as well as CDS.²³

Turning to suffixes, between 80 and 113 different verbal lemmas derived by one of the five suffixes *-íz-o*, *-áz-o*, *-ón-o*, *-éz-o*, and *-én-o* occur in CDS in each of the three subperiods of observation from 1;8 to 3;0. There is no tendency of an increased use of such verbs by the mother in the child’s course of development. Derived verbs of these groups amount to 162 lemmas cumulatively in the entire period of observation. The groups are not evenly distributed, however. Verbal lemmas with the suffix *-íz-o* are much more frequent than those with either of the suffixes *-áz-o* or *-ón-o*, and verbs derived by *-év-o* and especially *-én-o* are fewest (*-íz-o*: 59 lemmas; *-áz-o*: 37; *-ón-o*: 35; *-év-o*: 21; *-én-o*: 10). Not only is the sheer number of these verbal lemmas impressive but also the variety of their inflection forms. Many of them occur in 7 or 8 different TAM forms and some of them in even 14 forms.

The frequency of occurrence of these verbs in the input does not, however, seem to enhance their transparency or result in a productive use of the derivational suffixes in question by the child. As pointed out above, one reason will be that these suffixes only have a grammatical verbalizing function but no semantic content so that the derivative verbs are morphosemantically opaque. In addition, their

23. As mentioned above, Holton, Mackridge and Philippaki-Warburton (2016: 246) propose the same analysis for contemporary SMG.

morphotactic transparency may be obscured by inflection. As mentioned above (see Examples 9 and 10), the derivational suffix of many of these verbs changes its form more or less radically in the perfective aspect (Example 13b vs. 13a) or even disappears (Examples 13c). Such changes affecting consonantism, vocalism, or accent are likely to interfere with the morphotactic transparency of the word series constituted by the respective derived verbs. The formal homogeneity of these word series is additionally inhibited by a coincidence of certain of their TAM forms with those of simple verbs (Examples 13d). Hence, the child will not only have difficulties in recognising the morphological relations among verbs carrying a certain derivational suffix but also in distinguishing derived from simple verbs.

(13) Inflection of derived verbs

- a. *kaθar-iz/ís-o* (clean-IPFV/PFV-NPST.1SG)
aδj-áz/ás-o (empty-IPFV/PFV-NPST.1SG)
- b. *fon-áz/áks-o* (call-IPFV/PFV-NPST.1SG)
kri-ón/ós-o (be.cold-IPFV/PFV-NPST.1SG)
majir-év/éps-o (cook-IPFV/PFV-NPST.1SG)
- c. *maθ-én-o* (learn-IPFV-NPST.1SG) vs. *máθ-o* (learn.PFV-NPST.1SG)
zest-én-o (heat-IPFV-NPST.1SG) vs. *zest-án-o* (heat-PFV-NPST.1SG)
zest-én-ete (heat-IPFV-MEDIOPASS.NPST.3SG) ‘it is heating’ vs.
θα zest-aθ-i (FUT.PTL heat-PFV.MEDIOPASS-NPST.3SG) ‘it will heat’
- d. simple verb: *klín-o* (close-IPFV-NPST.1SG), *klís-o* (close.PFV-NPST.1SG),
é-klís-a (AUG-close.PFV-PST.1SG)
derived verb: *jiríz-o* (return-IPFV-NPST.1SG), *jirís-o* (return.PFV-NPST.1SG),
jírís-a (return.PFV-PST.1SG)

Consequently, the morphotactic complexity and resulting opacity of suffixed verbal derivatives as well as their morphosemantic opacity will make it difficult for the child acquiring Greek to unite them into word series. This differs strongly from word series of prefixed verbal derivatives, in which the prefix not only remains unchanged in different inflections, but is also semantically meaningful.²⁴ Since the establishment of word families presupposes the recognition of the base of derivatives, the formal difficulties met with many forms of verbs derived by suffixes will also interfere with this aspect of lexical structuring. Still, word families whose members have become firmly entrenched by their frequent use in CDS may also appear in CS. An interesting example is the use of the stem of *kub-i* ‘button’ in CDS (Table 6). The mother seems to be scaffolding the child’s acquisition of both derivation and inflection by laying the groundwork for the entrenchment of the

24. The only exception is an accent shift in certain forms (e.g. *kse-yráf-o* DER-write-IPFV-NPST.1SG vs. *ksé-yraps-a* DER-write-PFV-PST.1SG ‘I cross out’ vs. ‘I crossed out’).

base in the child's second year and at the same time introducing certain inflections of the derived verbs *kub-ón-o* 'to button' and *kse-kub-ón-o* 'to unbutton', the latter of which carries both a derivational prefix and a derivational suffix in front of the inflection. Inflections of the derived verbs virtually explode in the first half of the third year but have disappeared from the second half of that year by coincidence. Usage of the noun *kubí* 'button' and the two derived verbs *kubóno* 'to button' and *ksekubóno* 'to unbutton' in CS reflects CDS: While all occur in CS in the first two subperiods, some of them a little less frequently than in CDS, only the noun is found in the third period.

Table 6. Usage of the stem of *kub-í* 'button' in CDS (tokens)

	1;8–1;11	2;0–2;6	2;7–3;0
<i>kubí</i> 'button'	28 <i>kubí</i> SG 31 <i>kubjá</i> PL	4 <i>kubí</i> SG 12 <i>kubjá</i> PL	5 <i>kubí</i> 1 <i>kubjá</i>
<i>kub-áki</i> 'button-DIM'	4 <i>kubáki</i> SG 4 <i>kubákja</i> PL	3 <i>kubáki</i> SG 2 <i>kubákja</i> PL	–
<i>kúb-oma</i> 'buttoning'	–	1 <i>kúboma</i>	–
<i>kub-ón-o</i> 'to button'	3 <i>kubónis</i> 'button.IPFV. NPST.2SG' 1 <i>kubónume</i> 'button.IPFV. NPST.1PL' 1 <i>kubósi</i> 'button.PFV.NPST.3SG'	1 <i>kubónis</i> 'button.IPFV.NPST.2SG' 1 <i>kubóni</i> 'button.IPFV.NPST.3SG' 1 <i>kubónun</i> 'button.IPFV.NPST.3PL' 5 <i>kubóso</i> 'button.PFV.NPST.1SG' 1 <i>kubósis</i> 'button.PFV.NPST.2SG' 3 <i>kubósi</i> 'button.PFV.NPST.3SG' 4 <i>kubósume</i> 'button.PFV.NPST.1PL' 1 <i>kúbosa</i> 'button.PFV.PST.1SG' 1 <i>kúbose</i> 'button.PFV.PST.3SG' 1 <i>kuboθi</i> 'button.PFV.MEDIOPASS. NPST.3SG'	–
<i>kse-kub-ón-o</i> 'to unbutton'	3 <i>ksekúbose</i> 'unbutton.PFV. PST.3SG' 4 <i>ksekubóθike</i> 'unbutton.PFV. MEDIOPASS.PST.3SG'	1 <i>ksekubónis</i> 'unbutton.IPFV.NPST.2SG' – 1 <i>ksekubóni</i> 'unbutton.IPFV.NPST.3SG' 1 <i>ksekubósis</i> 'unbutton.PFV.NPST.2SG' 1 <i>ksekubúse</i> 'unbutton.IPFV.PST.3SG' 2 <i>ksekuboméno</i> 'unbuttoned.PST.PTCP'	–

Turning to the use of the derivational suffixes *-ik-* and *-in-* in CDS deriving adjectives or nouns, the word series with the suffix *-ik-* includes 47 lemmas and 150 tokens (Examples 14a) and the one with the suffix *-in-* 16 lemmas but 187 tokens²⁵ (Examples 14b). Certain forms of a number of lemmas in both series are nominalized

25. More than half of these tokens are attributable to *kókinos* 'red' and *epikíndinos* 'dangerous'.

and partially lexicalized (Examples 14c). There are a few formations in CDS combining both suffixes (Example 14d).

- (14) Nominal derivatives on *-ik-os* and *-in-os* (DER-INFL) in CDS
- a. *-ik-os*
vróm-ik-os ‘dirty’ (← *vróm-a* ‘dirt, nasty smell’)
xristujenját-ik-os ‘seasonal, Christmas style’ (← *Xristújena* ‘Christmas’)
 - b. *-in-os*
skot-in-ós ‘dark’ (AG base)
kók-in-os ‘red’ (< AG *kókos* ‘(coffee) bean’),²⁶ *kítr-in-os* ‘yellow’ (← *kítro* ‘citron’), *prás-in-os* ‘green’ (AG base)
 - c. Nominalized forms
níxt-ik-ó ‘nightdress’ (← *níxt-a* ‘night’) (lexicalized)
*kakúrj-in-a*²⁷ ‘bad girl’ (← *kakúry-os* ‘criminal’) (nominalized)
 - d. *kakurj-ín-ik-os*²⁸ ‘bad’

The frequency relation of the derivational suffixes *-ik-os* and *-in-os* found in CS mirrors CDS. Although the correspondence between the derivatives and their bases in CDS has not been studied, it may be hypothesized that the use of these suffixes by the child’s mother will have little impact on productivity and only contribute to the entrenchment of the respective derivatives in CS. The reasons preventing transparency of these derivatives for the child mentioned above (Section 4.1.2) also apply to those used in CDS: They are not segmentable, the suffixes do not occupy a prominent position in the word and are semantically empty. Since many bases of these formations come from AG, they are not documented in CDS.

With the exception of diminutives (see Thomadaki & Stephany 2007: 99–103), nominal derivatives are quite limited in CDS throughout the observation period. There is a single noun referring to places ending with *-ria* (e.g. *trapezaría* ‘dining room’; cf. *trapézi* ‘table’) and there are four nouns ending with *-jera* denoting objects (e.g. *banjéra* ‘bathtub’; cf. *bánjo* ‘bath’). The only type better represented among derived nouns are agent nouns ending with *-tis* (8 lemmas, 64 tokens) or *-stis* (1 lemma, 2 tokens) and feminines with *-isa* (4 lemmas, 63 tokens, 82% of the latter made up by *májisa* ‘witch, sorceress’ or *vasílisa* ‘queen’) or *-tria* (3 lemmas, 10 tokens, 8 of the latter covering *komótria* ‘female hairdresser’). The examples largely coincide with those used by the child (see Examples 12 above). Given the scarcity

26. The meaning ‘red’ comes from the use of kermes oak beans for obtaining red color (Babiniotis 1998).

27. *Kakúrj-in-a* and *kakurj-ín-ik-os* are ad hoc formations by the mother.

28. See the preceding fn.

of lexical diversity and the small number of tokens of most nominal derivatives, it may be concluded that CDS will do no more than entrench a few of these lemmas, such as the ones just quoted.

5. Summary and conclusions

The main results of this study on the development of derivation by prefixation and suffixation in early Greek language acquisition can be summarized as follows:

1. Before the end of the second year, there is hardly any evidence for the child's potential awareness of the morphological complexity of prefixed derivatives and virtually no evidence for the development of the derivational process of prefixation because prefixed adult forms occurring in CS are rare, the ones found are semantically opaque, and their renditions often distorted.
2. The derivational process of prefixation begins to develop in the first half of the third year and gathers momentum in its second half. By the end of the child's third year, the number of prefixed word series amounts to 9 with up to 7 members and 14 word families with mostly 2 and exceptionally 4 members.
3. There is an important difference between the functions of the preverb *ksaná* 'again' in SMG on the one hand and CDS as well as CS on the other. While *ksaná* is used as an adverb or the first constituent of compound verbs in the standard language, it exclusively occurs as a preverb in CS and CDS. It has therefore been considered to function as a prefix in both of these varieties, which shows that not only the grammatical structure of CS but also the one of CDS may differ from the standard language.²⁹ *Ksaná* is by far the most productive element for forming complex verbs in CS and its word series comprises 13 members in the second half of the child's third year. The productiveness of *ksaná* can be explained by the circumstance that its base is a word rather than a stem (Ralli 2004: [17]), resulting in morphotactic transparency. Furthermore, in *ksaná* formations, the meaning of the bases, which are all documented in CS already before the end of the second year, remains unchanged so that it may be assumed that *ksaná* formations are also morphosemantically transparent for the child.
4. Since diminutives have been excluded from the present study, the most numerous derivational formations occurring in CS are suffixed verbs. In contrast to prefixed formations, suffixed verbal derivatives are amply represented in CS from the beginning of observation in the last part of the second year. There is however evidence that these verbs remain largely morphotactically and

29. But see Holton, Mackridge & Philippaki-Warburton (2016: 246) referred to above.

morphosemantically opaque for the child throughout the observation period. Not only do their bases occur in at most 40% of exemplars, but they may also emerge later than the derivatives. Furthermore, in contrast to verbal prefixes, derivational suffixes only function as a verbalizing device and are semantically empty. No productive enlargement of the lexical inventory is possible with the help of these suffixes because each stem is combinable with only one particular suffix, a choice which has to be learned by heart. Finally, the morphotactic resemblance of the members of word series constituted by suffixed verbal derivatives is obscured by certain inflectional processes, some of which change the form of the derivational suffix across as well as within word series and inflections of such derivatives may even interfere with those of simple verbs. It is therefore likely that the child will treat structurally largely obscure suffixed verbal derivatives in the same way as underived, simple verbs and that their frequent occurrence from early on will at most support their entrenchment. Some of the reasons for the opacity of suffixed verbal derivatives also apply to suffixed adjectival derivatives. Most word families to which these adjectives belong only consist of the derivative and its base, in the rare cases in which the latter is documented in CS. Among nouns derived by suffixes, only agent nouns and feminines derived from masculines by motion occur fairly frequently (9 and 7 lemmas, respectively). The only evidence for possible “nuclear” word families are lexemes whose bases emerge earlier or in the same month as the derivative and whose tokens occur rather frequently (e.g. *vasiljás* ‘king’, 32 tokens; *vasílisa* ‘queen’, 23 tokens, both occurring in CS from 1;8 on).

5. Use of the preverbs *kse-* (reversative/intensive) and *ksaná* ‘again’ in CDS compared to other verbal prefixes provides a good example of the importance of type frequency, i.e. lexical diversity, as opposed to token frequency in promoting productivity in CS. In CDS, the *kse-/ksana* group of verbs comprises 34 lemmas and 128 tokens as opposed to 25 lemmas and 1,162 tokens of verbs derived by other prefixes. The type frequency found in CDS is reflected by 27 verbs of the *kse-/ksana* group as opposed to merely 11 of the other group in CS. Besides type frequency, an additional factor facilitating the development of productivity of the *kse-/ksana* group in CS is the morphotactic and morphosemantic transparency of most of these derivatives occurring in CDS, quite in contrast to the morphotactic and morphosemantic opacity of many verbal derivatives prefixed by *kata-*, *para-* or *pro-*. The importance of transparency for the development of derivational productivity in view of type frequency is demonstrated by the fact that the great lexical diversity of suffixed verbs occurring in CDS (162 suffixed verbal derivatives as opposed to 59 prefixed ones) does not incite the child to use such verbs productively (for reasons causing the opacity of suffixed verbal derivatives see Sections 4.1.2 and 4.2). The low

frequency of suffixed adjectival and nominal derivatives in CDS corresponds to what is found in CS. It may therefore be hypothesized that, with the exception of diminutive suffixes, the usage of adjectival and nominal derivational suffixes in CDS will have little impact on productivity in CS and will merely contribute to entrenchment.

Extending the view to SMG, the developmental level of derivation by prefixation and suffixation reached by the child at age 3;0 shows that the relative number of suffixes and prefixes occurring in her speech roughly corresponds to the situation found in the standard language (SMG: 116 suffixes vs. 32 prefixes; CS: 15 suffixes vs. 10 prefixes, including 6 diminutive suffixes (Thomadaki & Stephany 2007: 98–99)). Since the structural complexity of derivatives in CS is limited to maximally two derivational affixes, usually one prefix and one suffix, CS does not quite exploit the potentialities of SMG, where up to three derivational affixes may be combined with a stem (see Table 2 above).³⁰ The meaning of derivatives may be more or less close to that of their bases in SMG as well as in CS.

Item-based schemas, which are an important device for the development of lexical productivity, are limited to morphotactically and morphosemantically transparent derivatives closely related to the meaning of the bases functioning as slot fillers (e.g. *para*___, *kse*___, *ksana*___ with verbal fillers). In contrast to this, semantically opaque derivatives will be acquired piecemeal as *gestalts* depending on their frequency of use in daily interactions (e.g. frequent *apo-θiki* ‘closet’ (← *θiki* ‘case’) and *apó-jevma* ‘afternoon’ (← *jévma* ‘meal, dinner’) vs. infrequent *apó-kries* ‘Lent’). Evidence for the development of schemas in CS up to 3;0 is limited to a few derivational prefixes but has not been found with suffixes. The reason is that many word series of prefixed derivatives are morphotactically and morphosemantically transparent, but series of suffixed derivatives may be both morphotactically and morphosemantically opaque. The fact that the scope of word families registered in CS is much smaller than that of many word series and often consists of no more than a derivative and its base, shows that derivational devices allowing the combination of a given base with different derivational affixes are not yet taken full advantage of.

An interesting phenomenon in the speech of a 6-year-old boy is “the creation of derivatives for the purpose of avoiding difficult inflectional forms or syntactic constructions” (Stephany 1997a: 204). Overgeneralizations of e.g. derivative suffixes also appear much later than the age period studied in the present paper (Jannis,

30. It remains to be seen, however, in how far derivatives with such a degree of complexity are actually used in everyday language among adults.

6;2–6;4, *elefant-ína* (elephant-DER.F) ‘female elephant’, *yat-ul-ína* (cat-DIM-DER.F) ‘female little cat’, *mamá i pul-ína* (mommy the.F bird-DER.F) ‘mother bird’).³¹

To conclude, let me present a charming example of metalinguistic awareness drawn from a conversation between a linguist mother and her 6-year-old son (Thomadaki 1986, quoted by Stephany 1997a: 256).³²

- Mother: What does *yat-áki* (cat-DIM.NEUT) mean?
 Child: The little one; *yát-a* (cat-F) is the mother, *yát-os* (cat-M) is the father, and *yat-áki* (cat-DIM.NEUT) is the little one. I would rather call the mother *yat-ína* (cat-DER.F).
 Mother: And what is *yat-úla* (cat-DIM.F)?
 Child: One calls it *yatúla* because one loves it.

Acknowledgements

I would like to thank Demetra Katis, National and Kapodistrian University of Athens, for allowing me to use the corpus of child speech and child-directed speech which she collected from her elder daughter Anna. My gratitude also goes to Dimitra Theophanopoulou-Kontou, National and Kapodistrian University of Athens, for providing important comments on earlier versions of the manuscript and to two anonymous reviewers for their suggestions.

References

- Ambridge, B. & Lieven, E. V. M. 2011. *Child Language Acquisition: Contrasting Theoretical Approaches*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511975073>
- Babiniotis, G. 1998. *Lexiko tis Neas Ellinikis Glossas* (Lexicon of the Modern Greek language). Athens: Kentro Lexikologias.
- Babiniotis, G. 2002. *Lexiko tis Ellinikis Glossas: Me scholia ja ti sosti chrisi ton lexeon* (Lexicon of the Greek language: With comments on how to use words properly), 2nd edn. of Babiniotis 1998. Athens: Kentro Lexikologias.
- Babiniotis, G. 2011. *Etymologiko lexiko tis Neas Ellinikis Glossas: Istoría ton lexeon* (Etymological lexicon of the Modern Greek language: The history of words). Athens: Kentro Lexikologias.
- Booij, G. 2013. Morphology in construction grammar. In *The Oxford Handbook of Construction Grammar*, T. Hoffmann & G. Trousdale (eds), 255–273. Oxford: OUP.
- Bybee, J. L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455. <https://doi.org/10.1080/0169096950840711>

31. Examples taken by Theophanopoulou-Kontou (p.c.) from her personal diary.

32. This example also comes from the (unpublished) metalinguistic conversations held by Dimitra Theophanopoulou-Kontou, University of Athens, with her son Jannis.

- Bybee, J. L. 2001. *Phonology and Language Use*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511612886>
- Bybee, J. L. 2007 [1998]. The emergent lexicon. In *Frequency of Use and the Organization of Language*, J. L. Bybee (ed.), 279–293. Oxford: OUP. (Reprinted from CLS 34: 421–435, 1998)
- Bybee, J. L. 2010. *Language, Usage and Cognition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511750526>
- Charalambakis, C. 2014. *Christiko lexiko tis Neoellinikis Glossas* (Practical dictionary of the Modern Greek language). Athens: Academy of Athens.
- Christofidou, A., Afentoulidou, V., Karasimos, A. & Dimitropoulou, I. 2013. Ilektroniko programma neologismon kai orologias NeoDEMIA: Prokliseis kai dhiktio-liseis (The electronic programme for neologisms and terminology NeoDEMIA: Challenges and internet solutions). In *Dimiourgia kai morfi sti glossa: Deltio Epistimonikis Orogias kai Neologismon* (Creation and form in language: Reports on scientific terminology and neologisms), Vol. 12, A. Christofidou (ed.), 198–243. Athens: Academy of Athens.
- Christofidou, A. & Stephany, U. 2003. Early phases in the development of Greek verb inflection. In *Development of Verb Inflection in First Language Acquisition: A Cross-linguistic Perspective* [Studies on Language Acquisition 21], D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 89–129. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110899832.89>
- Clairis, C. & Babinotis, G. (in cooperation with Moser, A., Bakakou-Orfanou, E. & Skopeteas, S.). 2005. *Grammatiki tis Neas Ellinikis: Domoleitourgiki-epikoinoniaki* (A grammar of Modern Greek: Structural-functional and communicative). Athens: Ellinika Grammata.
- Diamanti, V., Benaki, A., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S. & Protopapas, A. 2018. Development of early morphological awareness in Greek: Epilinguistic versus metalinguistic and inflectional versus derivational awareness. *Applied Psycholinguistics* 39: 545–567. <https://doi.org/10.1017/S0142716417000522>
- Diessel, H. 2013. Construction grammar and first language acquisition. In *The Oxford Handbook of Construction Grammar*, T. Hoffmann & G. Trousdale (eds), 347–364. Oxford: OUP.
- Holton, D., Mackridge, P. & Philippaki-Warbuton, I. 2016. *Greek: An Essential Grammar*, 2nd edn. London: Routledge. <https://doi.org/10.4324/9781315680309>
- Humbert, J. & Kourmoulis, G. 2002[1957]. *Syntaktikon tis Archaias Ellinikis Glossis* (Syntax of the Ancient Greek language). Athens.
- Jannaris, A. N. 1968[1897]. *An Historical Greek Grammar, Chiefly of the Attic Dialect*. London: MacMillan. Reprinted 1987, Hildesheim: Georg Olms.
- Joseph, B. D. & Philippaki-Warbuton, I. 1987. *Modern Greek*. London: Croom Helm.
- Mackridge, P. 1985. *The Modern Greek Language: A Descriptive Analysis of Standard Modern Greek*. Oxford: Clarendon Press.
- Ralli, A. 2004. Stem-based versus word-based morphological configurations: The case of Modern Greek preverbs. *Lingue e Linguaggio* 2: 269–302. <https://doi.org/10.1418/16116>
- Ralli, A. 2013. *Compounding in Modern Greek*. Dordrecht: Springer.
<https://doi.org/10.1007/978-94-007-4960-3>
- Ralli, A. 2014. *Morphologia* (Morphology), 6th edn. Athens: Patakis.
- Stephany, U. 1980. Zur psychischen Realität der Dimension der Deskriptivität (On the psychological reality of the dimension of description). In *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler* (Approaches to research into universals: Linguistic contributions on the occasion of the 60th birthday of Hansjakob Seiler) [Tübinger Beiträge zur Linguistik 145], G. Brettschneider & C. Lehmann (eds), 549–555. Tübingen: Gunter Narr.

- Stephany, U. 1985. *Aspekt, Tempus und Modalität: Zur Entwicklung der Verbalgrammatik in der neugriechischen Kindersprache* (Aspect, tense and modality: On the development of verbal grammar in Modern Greek child language) [Language Universals Series 4]. Tübingen: Gunter Narr.
- Stephany, U. 1997a. The acquisition of Greek. In *The Crosslinguistic Study of Language Acquisition*, Vol. 4, D. I. Slobin (ed.), 183–333. Mahwah NJ: Lawrence Erlbaum Associates.
- Stephany, U. 1997b. Diminutives in early child Greek: A preliminary investigation. In *Studies in Pre- and Protomorphology*, W. U. Dressler (ed.), 153–163. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Stephany, U. 2012. Selecting a theoretical framework fitting acquisition data is no easy matter. In *Selected Papers of the 10th International Conference of Greek Linguistics*, Z. Gavriilidou, A. Efthymiou, E. Thomadaki & P. Kambakis-Vougiouklis (eds), 89–100. Komotini, Greece: Democritus University of Thrace. Retrieved from <www.icgl.gr> (10 January 2019)
- Stephany, U. & Christofidou, A. 2009. The emergence of nominal inflection in Greek. In *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective* [Studies on Language Acquisition 30], U. Stephany & M. D. Voeikova (eds), 217–264. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110217117.217>
- Stephany, U. & Thomadaki, E. 2016. *Compounds in Early Greek First Language Acquisition – Including an Onomasiological Approach to Lexical Typology of Greek and German* [Working Paper No. 60]. Cologne: General Linguistics, Institute of Linguistics, University of Cologne. <<http://www.ub.uni-frankfurt.de/online/home.html>>, <<https://ifl.phil-fak.uni-koeln.de/27221.html>> (15 January 2019).
- Stephany, U. & Thomadaki, E. 2017. Compounding in early Greek language acquisition. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketrez & M. Kilani-Schoch (eds), 119–143. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.06ste>
- Thomadaki, E. 1986. *Wortbildungsstrategien und Mutterspracherwerb* (Strategies of word formation and acquisition of the mother tongue). Ms., Institute of Linguistics, University of Cologne.
- Thomadaki, E. 1988. *Neugriechische Wortbildung* (Modern Greek word formation) [Working Paper No. 8 (N.F.)]. Cologne: Institute of Linguistics, University of Cologne. <<http://www.ub.uni-frankfurt.de/online/home.html>>, <<https://ifl.phil-fak.uni-koeln.de/27221.html>> (20 January 2019)
- Thomadaki, E. & Stephany, U. 2007. Diminutives in Greek child language. In *The Acquisition of Diminutives. A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43], I. Savickienė & W. U. Dressler (eds), 89–123. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.05tho>
- Tomasello, M. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge MA: Harvard University Press.
- Tomasello, M. 2010. *Origins of Human Communication*. Cambridge MA: The MIT Press.

Derivational patterns in spontaneous data of French-speaking parent-child interactions before age three

Marianne Kilani-Schoch and Aris Xanthos

University of Lausanne

This chapter is a first step towards the characterisation of the morphological structure of the French lexicon in early first language acquisition, i.e. before children coin novel word formations. Focusing on nouns and verbs, it analyses the variety of derivational means used by toddlers and caregivers in two corpora of French-speaking children (1;4/1;6–2;11). A comparison is done with a sample of adult-directed (ADS) speech. Findings on derivation are compared with previous observations on compounds in the same data. The results display the development of tight-knit morphological relationships within the lexicon and a clear prevalence of suffixation over other derivational means and compounding. Along with errors in affixed words, these relationships provide cues of early detection of derivational morphology in child speech.

Keywords: affixed words, converted words, morphological errors, compounds

1. Introduction

In her analysis of classic diary studies and other acquisition studies on French, Clark (1993: 161, 209) found that, in contrast to e.g. Germanic language-speaking children, French-speaking children do not coin new nouns or verbs before age 4. As conversion (zero-derivation or “identity operation”, Bauer, Lieber & Plag 2013: 546), has some productivity in the language but produces activity nouns rather than object nouns, she hypothesises that children “might wait until they acquire some affixes”.

This chapter is interested in characterising what is happening regarding derivation, and more specifically affixation, during this waiting period. It strives to show the role of affixed words in the early lexicon and to see whether there are traces of an emerging knowledge of derivation before neologisms are coined. For

that purpose it analyses the proportion of affixed words, the distribution of various affixation patterns in child speech (CS) and child-directed speech (CDS), the degree of diversity displayed by affixes, as well as the developmental trajectories that can be found in the data, including semantic development.

Conversion, which involves the same word classes, will also be examined. However, a thorough characterisation of this device in children's speech deserves a dedicated study in particular because of the theoretical problems it raises (Lippeveld & Oshima-Takane 2015a; b).

The analysis is based on the corpora of two French-speaking children recorded between 1;6 and 2;11. As data from the acquisition literature are scarce (see e.g. Clark 1985, 1993, 1998; Lippeveld & Oshima-Takane 2015a; b) and mainly consist in elicited data, the chapter is meant to be a first step towards the characterisation of the morphological structure of the French lexicon in spontaneous data of first language acquisition before age three. It also aims at finding first signs of emerging knowledge of derivational morphology in the course of development.

A final question is how specific the findings on derivational means in CS and CDS data are with regard to derivational means in ADS. A sample of ADS from the parents of the children recorded will be used as a basis of comparison with the CS and CDS data. In particular, the differences in affix types and diversity will be examined.

The structure of the chapter is the following: after a short description of the main derivational means in the target language in Section 2, data and method are presented in Section 3. Section 4 develops the results of the analysis on affixation and conversion, respectively suffixed nouns, suffixed adjectives, prefixed verbs and noun-verb pairs (conversion). Section 5 on the one hand compares the results on suffixation in CS and CDS with previous findings on compounding, and with observations made in a sample of ADS, on the other hand. In Section 6 we conclude on morphological relations found within the CS's lexicon, pace of development of simplicia and derived words as well as errors in complex words as constituting cues of detection of derivational morphology.

2. French derivational morphology in the target system

French derivational patterns are typically represented by affixation. There is a large inventory of suffixes deriving nouns, verbs, adjectives from bases consisting of nouns, verbs or adjectives. Adverbs can be built on adjectives by suffixation as well (*joli-ment* 'nicely').

Compared with Germanic or Finnic languages, French is generally said to be a poor-compounding language, and the productivity of NN formations by juxtaposition which are very much used in written French, e.g. in advertising as already noticed by Clark (1993: 160), does not change the picture since their status as morphological compounds is not widely accepted (see Fradin 2009; Villoing 2012; Bauer 2011; Gaeta 2015).

As is typical for Romance languages, nominal derivation is more widespread and productive in French than verbal derivation (Bauer 2011: 533). Suffixation prevails over prefixation in terms of diversity of affixes and lemmas. Nonetheless prefixation comprises a number of productive morphemes combining with verbs and nouns, with a possible dominance of verb formations (Bauer 2011: 532). Neologisms are thus mainly formed with affixes and particularly suffixes, e.g. *ubér-is-ation* ‘uberisation’, *développ-eur* ‘developer’, *bruit-iste* ‘noise maker’ from *bruit* ‘noise’, *viral-ité* ‘virality’ (de-adjectival), also in combination with prefixes: *dé-référence-ment* ‘dereferencing’, *dé-faç-age* ‘defacement’ (based on the English word).

Another productive type of operation is conversion (zero-derivation or identity operation, see above) commonly used to coin new denominal verbs (belonging almost exclusively to the first conjugation class), e.g. *mécenn-er* ‘patronise’ from *mécène* ‘patron’, or deverbal adjectives (derived from past or present participle), e.g. *crypt-é* ‘encrypted’, *cliv-ant* ‘cliving’. Conversions from adjectives to verbs (of the 2nd conjugation class) are also common: *sal-ir* ‘dirty’, *roug-ir* ‘blush’, etc. Whether deverbal adjectives based on past participles (*fatigué* ‘tired’, *ralenti* ‘slowed down’, *disparu* ‘disappeared’) have to be categorised as conversions or rather as suffixations is a matter of debate (see Tribout 2010: 22). This is even more true for adjectives derived from present participles.

Several nominal diminutive suffixes (e.g. *-et/-ette* in *chemis-ette* ‘short sleeved shirt’, *-on* in *ours-on* ‘bear cub’, *-elle* in *coup-elle* ‘cup’) occur in French. Feminine *-ette* is the most frequent and, to a certain extent, the most productive of those (Fradin, Hathout & Meunier 2003: 73, 76). The productivity of the diminutive suffix *-ette* is limited to a few transparent patterns (e.g. *maisonn-ette* ‘small house’, *camionn-ette* ‘small truck’). Indeed, a great number of suffixed words in *-et/-ette* words are morphosemantically more or less opaque, i.e. lexicalised (Fradin, Hathout & Meunier 2003: 56). Hence, in contrast to other Romance languages, the French diminutives are basically morphosemantic, and pragmatic effects typical of diminutive-rich languages are taken over by the adjective *petit* ‘small, little’ (Kilani-Schoch & Xanthos 2013).

Although diminutive suffixes most generally do not change the category of the word and show a preference for nominal bases (Dressler & Merlini Barbaresi 1994: 94), some diminutive suffixes in *-ette* with instrument meaning (*sonn-ette*

'bell' from *sonn-er* 'ring'), *servi-ette* 'napkin' from *serv-ir* 'serve') or object meaning (*suc-ette* 'lolly pop' from *suc-er* 'suck') raise the issue of possible verbal bases. Corbin (1987) and Dal (1997), among others, strongly argue in favour of (truncated) nominal bases or conversions.¹ We will not enter into this discussion here.

As diminutives constitute an important part of the basic French lexicon, in contrast to the languages that are presented in the other chapters of the volume, and since the acquisition of French diminutives has not been dealt with in Savickiené and Dressler (2007), the present chapter includes diminutive suffixes.

2.1 Noun formation

Within the set of affixes available in the grammatical sources on French derivation (Arrivé, Gadet & Galmiche 1986; Béchade 1992; Huot 2001; Apothéoz 2002, *TLFi* 2021), nominal suffixes (more than 40) clearly outnumber adjectival and verbal suffixes.

Although accurate data on suffix productivity are still scarce (see, e.g. Fradin, Hathout & Meunier 2003; Grabar et al. 2006), it is commonly considered that frequent and productive nominal suffixes, i.e. suffixes used to coin new words, are deverbal *-eur*, *-ateur* /fem. *-euse*, *-atrice* deriving agent and instrument nouns as well as denominal *-iste* (*dent-iste* 'dentist' from *dent* 'tooth') for agent nouns, the pattern of deverbal action nouns in *-age* (*rempliss-age* 'filling'), to a lesser extent that in *-ment* (*blanchi-ment* 'blanching' from *blanc*, fem. *blanche* 'white'), and of instrument nouns in *-oir(e)* (*arros-oir* 'watering can'). N in *-erie* for location and activity (*épicerie* 'grocery' from *épice* 'spice') are frequent (and have a clear productivity nowadays, e.g. *sandwich-erie* 'sandwich store', *animal-erie* 'pet shop').

Deadjectival property nominal suffixes in *-(i)té* (*beau-té* 'beauty'), denominal agent nouns in *-(i)er/-i)ère* (*pomp-ier* 'fireman' from *pompe* 'pump') and deverbal action, result and object nouns in *-ure* (*bless-ure* 'wound') also belong to frequent suffixes (see Guilbert 1975; Clark 1993; Lehmann & Martin-Berthet 1998; Huot 2001; Grabar et al. 2006; *TLFi* 2021).²

1. Specifically, the agent noun *sonn-eur* (from *sonn-er* 'ring') and the potential (°) instrumental noun °*sonn-oir(e)* (Dal 1997: 179) on the one hand, the object noun *suc-on* from *suc-er* 'suck' (Corbin 1987: 691), on the other.

2. *-tion/-ation* and allomorphs (*-ification*, *-isation*): *absorp-tion*, *distan-ci-ation* 'detachment', *diversific-ation*, *féminis-ation* 'feminisation' deriving action/result nouns, or nouns in *-itude* (*trist-itude* 'sad attitude', Koehl 2012: 300) which are highly productive are not mentioned here as they are not likely to occur in early child language. The same is true of nouns in *-isme* denoting a property, a result (*alcool-isme* 'alcoholism') or a doctrine (*jeun-isme* 'ageism') or of nouns in *-iste* for adherents to a doctrine.

Some prefixes have recently gained a high productivity in casual French, i.e. *sur-* (*sur-booké* ‘overbooked’), *super-* (*super-héros* ‘superhero’), *hyper-* (*hypermarché* ‘hypermarket’) but they are not specific to nouns and much used with verb and adjective bases. Other prefixes combining with nouns include prepositions like *sous-* (*sous-verre* ‘clip-frame’), *contre-* (*contre-sens* ‘misinterpretation’) (see Fradin 2003: 196) and *anti-* (*anti-héros* ‘antihero’). Most of them also apply to verb bases. The question of the theoretical status of these prefixes which are also autonomous morphemes (compound constituent, hybrid element?) is a matter of debate (Amiot 2004). However, they are irrelevant for child language and so are the prefixes *re-* and *in-* which are used with abstract nouns and adjectives (*re-naissance* ‘rebirth’, *in-existence* ‘non-existence’) (but see below for deverbial *re-* and de-adjectival *in-*).

2.2 Verb formation

As said above, verb suffixation patterns are limited to a few suffixes such as denominal and de-adjectival *-is-* (*panthéon-iser* ‘pantheonise’, *précar-iser* ‘make precarious’) and *-ifi-* (*personn-ifier* ‘personify’, *divers-ifier* ‘diversify’) with a factitive meaning, which are very productive (Huot 2001: 80–81).

The most common and productive verb prefixes are iterative and reversative *re-* (*re-dessiner* ‘redraw’, *re-venir* ‘come back’) (Apothéloz 2007) and privative *dé(s)-* (*dé-faire* ‘undo’, *dé-loyal* ‘dishonest’)³ which combine with verbal and adjectival bases. Other prefixes are denominal *en-* (*en-dormir* ‘get to sleep’) and denominal and de-adjectival *a-* (*a-tterrir* ‘land’, *a-grandir* ‘extend’) and *é-/ex-* (*é-courter* ‘shorten’, *ex-matriculer* ‘exmatriculate’).

Conversion between nouns and verbs is a widespread means of word formation in French and plays a significant role in the conventional lexicon. According to Tribout (2010: 143, 148, 170, 172, 176), noun to verb conversions (e.g. N *peigne* [peɲ] – inf. *peign-er*, sg. pres. ind. *peigne* [peɲ] ‘comb’) outnumber verb to noun conversions (e.g. inf. *réveiller*, sg. pres. ind. *réveille* [revɛj] ‘wake up’ – N *réveil* [revɛj] ‘alarm clock’). However, there are cases where the directionality of conversion is undecidable (Tribout 2010: 196). Conversion is even more pervasive in French as one includes cases such as, e.g. inf. *saut-er* [sote], sg. pres. ind. *saute* [sot] ‘jump’ – N *saut* [so], inf. *refus-er* ‘refuse’, sg. pres. ind. *refuse* [rəfyz] – N *refus* [rəfy] ‘refusal’, inf. *dessin-er* ‘draw’, sg. pres. ind. *dessine* [dəsin] – N *dessin* [dəsɛ̃], *sel* [sɛl] – *sal-er* sg. pres. ind. *sale* [sal] ‘salt’ (Melčuk 1996: 130–131; Tribout 2010: 102, 114, 135). In these cases, there is a word-class and paradigmatic change of a base with a

3. The allomorph *dis-* occurs with verbal and nominal bases (*dis-créditer* ‘discredit’, *dis-harmonie* ‘disharmony’).

semantically regular relationship (Manova 2011: 60 ff.) but no strict morphotactic identity between the forms. The first three examples display an alternation between the realisation of a final consonant in the verb and its non-realisation in the noun. In various theoretical models the consonant, which may occur in liaison, inflection or derivation before a vowel (e.g. *dessin-atrice* [desinatris] ‘designer’, is considered as latent (see Tribout 2010 for an overview). The last two pairs show vocalic allomorphy of the base, combined, in *dessine – dessin*, with consonant alternation. As alternations between a short and a long form of a lexeme is ubiquitous in French morphology (see the model of thematic spaces in, e.g. Bonami & Boyé 2005; Plénat 2008; Bonami, Boyé & Kerleroux 2009), Tribout (2010) includes them in conversions based on the argument that the change cannot be equated with derivational affixation. In the present chapter we will nevertheless treat them in the separate category of modification, for the relation between identical forms obviously should be handled more easily by toddlers than the relation between forms blurred by allomorphy.

Tribout (2010: 2) counted 3241 noun – verb pairs in a relation of conversion, the majority of which are noun to verb conversions as just mentioned. It follows from the above that conversion also raises important theoretical questions relating notably to the delimitation of what can be subsumed under the operation.⁴

2.3 Adjective formation

The most productive adjectival suffix is arguably the deverbal possibilitive *-able/-ible* (e.g. with Latin bases) (*mange-able* ‘edible’, *destruct-ible* ‘destructible’). Denominal productive suffixes are e.g. *-eux* (*joy-eux* ‘joyful’), relational *-ier* (*piétonn-ier* ‘pedestrian’), or ethnic *-ain* (*améric-ain* ‘American’).

3. Data and method

3.1 Data

This study is based on the corpora of two children from Lausanne (Switzerland): Sophie and Emma, born in upper-middle class families (see Table 1). The children were audio recorded by the parents about twice a month, for about half an hour.

4. Here we will neither handle the question of the relevance of verbal stem space in conversion (Bonami, Boyé & Kerleroux 2009) nor discuss the distinction between syntactic and morphological conversion (Kerleroux 1996; Tribout 2010). But see Section 3.2.

Recording situations vary between free and everyday activities such as eating, washing, book reading, and having a bath.

Emma's data are more limited than Sophie's. Some of the recordings of Emma are very short (especially those at 1;6, 1;7, and 2;0; at 1;7 diary notes complement the recordings). Probably, this limitation partly accounts for the greater heterogeneity of some of the findings concerning her language development.

Transcription and coding have been done according to the CHAT conventions of CHILDES (MacWhinney 2000).

Table 1. CS and CDS data

Corpus	Period	No. of utterances		No. of recordings	No. of hours
		CS	CDS		
Sophie (SOP)	1;6-2;11	15,564	12,279	45	28
Emma (EMM)	1;4-2;11	7,335	6,639	40	18

3.2 Method

The main and crucial methodological issue is the identification of morphologically complex words. The question of what is a morphologically complex word is even more difficult when dealing with CS data.

The differentiation between actual suffixes and lexicalised or pseudo suffixes (formally similar to existing suffixes but semantically generally empty and without a simplex base) e.g. *tableau* 'board' (CS Sophie), *gâteau* 'cake', *château* 'castle' (Sophie's CDS), or between affix polysemy and affix homophony, between double suffixation and allomorphy, etc. (see below) was done following the morphological literature when available. The morphological dictionary *Le Robert Brio* (2012) as well as the main electronic dictionaries (*TLFi* 2021, *Grand Robert* 2018) were also consulted in order to assess derivational patterns. For instance, following Fradin, Hathout & Meunier (2003: 56), we considered that the (nominal) suffixes masculine *-et* (e.g. *poul-et* 'chicken') and feminine *-ette* (e.g. *chauss-ette* 'sock') are different suffixes. Indeed, the type frequency of *-ette* is higher than the type frequency of *-et*; moreover, opaque lexemes in *-et* are five times more frequent than opaque lexemes in *-ette*. The suffix *-erie* is synchronically a single suffix⁵ (Roché 2009: 164, Booij 2012). The lexeme *can-eton* 'duckling' was not analysed as *cane-et-on* but as *can-eton* as all names of young/baby animals are constructed on the species name (see Roché 2009: 145). The suffix *-eron* (Emma-CS) is considered as an allomorph of *-on* (*TLFi*

5. Reanalysis and coalescence of *-ier* with addition of *-ie*.

2021), *-ieux* in *délic-ieux* ‘delicious’ an allomorph of *-eux* in *courag-eux* ‘brave’ (Sophie’s CDS, Emma’s CS) and *-er* (*roch-er* ‘rock’) an allomorph of *-ier* (*ros-ier* ‘rose bush’). On the other hand feminine *-(i)ère* [jɛr] which differs formally from the masculine was counted separately. We included gender suffixes as derivational morphemes even in the case of nominalised adjectives such as *coqui-ne* (arguing, e.g. for the possible addition of a diminutive suffix: *coqui-n-ette* ‘mischievous’).

In the selection step, we adopted a target-like or system-centered perspective to avoid any unsound assumption on children’s knowledge. All lemmas with a morphotactically and morphosemantically identifiable base and suffix were extracted from the corpora and manually analysed. This criterion entails that examples such as *buanderie* ‘laundry room’ (*buandier* is not used anymore), *brouette* ‘wheelbarrow’, *moquette* ‘carpet’, *lolette* ‘dummy’ or *galette* ‘shortbread biscuit’ in which the base is not synchronically identifiable have not been kept (for morphosemantics, see below). By contrast examples such as *assi-ette* (base *assi-*, past participle from *asse-oir* ‘sit’) and *serviette* (base *servi-*, past participle from *serv-ir*) have been included among suffixed nouns. As far as nouns in *-ette* are concerned, verbal bases have thus been assumed in some examples.⁶ Moreover, in a small number of cases, e.g. *bavette* ‘bib’ or *dînette* ‘playing dinner party’, either a nominal or a verbal base can be assumed.

We have also excluded words which are morphologically complex but not constructed (Corbin 1987: 188), i.e. words which have some internal structure but contain one semantically empty constituent formally identical to an existing morpheme, such as *pommade* ‘ointment’ (the suffix *-ade* ‘made of’ is combined with a semantically empty basis, as *pomme* ‘apple’ denotes a fruit).

Taking all lemmas with a morphotactically and morphosemantically identifiable base and suffix into account leaves open the problem of how to deal with allomorphy, which may alter the identity of base or suffix. The two following criteria have been applied:

1. A lexeme with base or affix allomorphy is treated as a derived word if it is part of a morphological series in the target-language, i.e. if there are several derived words which synchronically attest to its analysability, e.g. *pan-ier* ‘basket’ derived from *pain* [pɛ̃] ‘bread’ (*salad-ier* ‘salad bowl’, *beurr-ier* ‘butter dish’, *cendr-ier* ‘ashtray’), *ann-ée* ‘year’ derived from *an* [ã] ‘year’ (*journ-ée* ‘day’, *soir-ée* ‘evening’), *man-ette* ‘lever’ from *main* [mɛ̃] ‘hand’ or *mi-ette* [mjɛt] ‘crumb’ from *mie* ‘crumb’ (with morphonological gliding of /i/). In this way, a word like *docteur* ‘doctor’ is considered as a simplex word, although its

6. However, this surface analysis does not prejudge the final status of these bases, which is the subject of theoretical debate (see Section 2.).

ending seems identical to agentive *-eur*, as agent nouns are derived from verb or noun bases, not adjectival ones. The criterion of belonging to a series also applies to base and affix allomorphy which does not change the base or affix but adds material, e.g. *suppos-it-oire* 'suppository', *ri-z-ière* 'rice field', *chic-l-ette* 'chewing-gum', *chan-son* 'song' from *chant* [ã] 'song (*boi-sson* 'drink', *cui-sson* 'cooking') are clearly derived words.

2. A lexeme with base allomorphy is a morphologically complex word if it is part of a morphological family: we have already seen *pan-* in *panier* 'basket' (*pan-ure* 'breadcrumbs', *pan-ade*, etc.) which also belongs to a series. It is also the case of *cass-ette* [kaset] 'tape' (derived from *caisse* [kæs] 'crate') which belongs to the series of lexicalised diminutives in *-ette* and to a family of containers: *cass-erole* 'pan', *cass-ol-ette* 'dish', etc.

Let us add that the actual realisation of the lexemes that have been selected was systematically checked in the corpora and instances in which the modification of the word was substantial, i.e. the affix was not easily identifiable, were discarded, e.g. Emma 1;6 [vai] for *voil-ier* 'sailboat' - Sophie 1;10 *tapam* for *tromp-ette* 'trumpet'.

On the semantic level, identification of derived words is especially complex because there are different degrees of lexicalisation (in the sense of morphosemantic opacification). This is why any constituent which was morphotactically identifiable and had a semantic meaning was considered as relevant for assessing the morphological complexity of a lemma, independently of the semantic relation with the derived word: e.g. *pinc-eau* 'paintbrush', *toil-ettes* 'bathroom' or *cass-ette* 'tape', in which *pince* 'pliers', *toile* 'hessian' and *ca(i)sse* 'crate' have no transparent relation with the final lexeme, i.e. are semantically too opaque, were still considered as morphologically complex. Other examples include *chauss-ette* 'sock', *chic-lette* 'chewing gum', *mir-oir* 'mirror' or *tract-eur* 'tractor', in which the nominal or verbal bases *chique/chique(r)* 'tobacco/tobacco chew' *chausse/chauss-e(r)* 'hose/to put on (shoes)', *se mi-re(r)* 'gaze at one's reflection', *tract-e(r)* 'tow' have a weak degree of transparency in the target language. Nevertheless, they are morphosemantically identifiable. They were thus included even though most probably the children never heard these verbs in CDS, because they are very formal/literary or very specific.⁷ From the perspective of acquisition, the argument for keeping *chic-lette* is that a child may identify just one salient morpheme of a word, in this case the productive or frequent suffix *-ette*. In other cases a familiar basis (e.g. *peint(e)* 'paints/painted' in *peint-ure* 'painting') may have been identified.

7. This does not preclude the strong relationship between morphological decompositionability and morphosemantic transparency in processing that has been pointed out by a reviewer. The rather arbitrary criterion of morphosemantic identifiability has been chosen for this corpus study since the available experimental data on degrees of morphosemantic transparency are still patchy.

4. Results

Table 2 displays the total number of verbs, nouns and adjectives found in the two corpora.

Table 2. Total number of both derived and underived verbs, nouns and adjectives in Sophie's and Emma's corpora (lemmas and tokens)

	V		N		ADJ	
	Lemmas	Tokens	Lemmas	Tokens	Lemmas	Tokens
Sophie-CS 1;6-2;11	188	6,459	494	4,544	71	1,241
Emma-CS 1;4-2;11	205	4,757	572	3,060	70	879
Sophie-CDS 1;6-2;11	278	13,595	631	5,939	97	1,866
Emma-CDS 1;4-2;11	294	8,455	780	4,592	93	1,179

4.1 Suffixation

Suffixed words, especially suffixed nouns, occur from early on in the data and provide the largest morphological series, in comparison with prefixation, in accordance with the role and the distribution of suffixation in the target language (see Section 2). Suffixed adjectives are limited while suffixed verbs are non-existent. Generally, suffixed words contain no more than one suffix. But some exceptions do occur, e.g. in words with a gender suffix (Emma's CS *ind-ien-ne* 'Indian' -*ien* ETHNIC, -*ne* FEMININE, *proch-ai-ne* 'next', *Japon-ai-se* 'Japanese').

4.1.1 Nominal suffixes

In Tables 3a and b the number of different nominal suffixes (suffix types) as well as the proportion of suffixed nouns in relation to the total number of noun lemmas are presented. Suffixes combining with at least two bases (Table 3a) are distinguished from isolated suffixes, i.e. occurring with one base, thus in a single lexeme (Table 3b).

At first glance, the results do not yield striking differences between the corpora except for the relations between CS and CDS with Emma's CDS having a higher number of suffixes and proportions of suffixed nouns. Regarding in particular the proportion of suffixed noun lemmas of series (i.e. combined with at least two bases), a chi-square test leads us to reject the hypothesis that they are equal ($\chi^2 = 14.87$, $df = 3$, $p < 0.01$); post-hoc tests using the Marascuilo procedure (Marascuilo & McSweeney 1967) show that the only pairwise difference that is significant is the difference between Emma's CS and Emma's CDS (6.1% vs 12.2%). Proportions of

Table 3a. Suffixed nouns (with at least 2 bases)

	No. suffix types		No. suffixed N		% of N	No. N total	
	With at least 2 bases	Isolated	Lemmas	Tokens		Lemmas	Tokens
Sophie's CS	8	12	42	296	8.5	494	4,544
Emma's CS	8	13	35	165	6.1	572	3,060
Sophie's CDS	12	12	59	425	9.4	631	5,939
Emma's CDS	16	15	95	383	12.2	780	4,592

Table 3b. Suffixed nouns (1 base)

	No. suffixed N		% of N
Sophie's CS	12	76	2.4
Emma's CS	13	74	2.3
Sophie's CDS	12	80	1.9
Emma's CDS	15	86	1.9

suffixed nouns of series are more similar in Sophie's corpus than in Emma's corpus. That there are more differences between Emma's CS and CDS than between Sophie's CS and CDS will be seen in all results.⁸

It is no surprise that the number of isolated suffixes (i.e. showing up with a single lemma) is greater than the number of suffixes combined with at least two bases in the limited lexicon of CS before 3;0. Nevertheless, even though confidence intervals (CI) are broad, a sizeable proportion of the suffixed nouns which form morphological series are related to other words within embryonic morphological families (see Table 4). We call them embryonic for the reason that they have 2 members only (see Introduction to the volume), with the exception of Sophie's CS (and CDS) triplet *manch-on*, *manch-ette*, *manch-oir* 'swimming wings', simplex *manche* 'sleeve' and Emma's CS *bain*, (*se*) *baigne* 'bath', *baign-oire* 'bathtub' and *jouer* 'play', *jouet* 'toy', *joueur* 'player' (but see Section 4.3.1.2).

Table 4. Embryonic families

	Families	No. suffixed N (at least 2 bases)	%	95% CI
Sophie's CS	21	42	50	[34.2, 65.8]
Emma's CS	15	35	42.9	[26.3, 60.6]

8. It is probably due to a large extent to the behaviour of Sophie's mother during the recordings consisting in talking little during the recordings in order to leave the floor to the child, as well as to the more limited data of Emma's corpus (see Section 3.1).

These family relations suggest that suffixed words start getting some morphological motivation in the linguistic representation of the children. The comparison of age of emergence of simplicia or bases and corresponding derived words supports this hypothesis: in the two corpora the simplex word (e.g. *manche* 'sleeve', *peigne* 'comb', sg. pres. ind. *lave* 'washes' or *balance* 'swings') tends to emerge earlier than the derived corresponding word (e.g. *manch-ette*, *manch-on* 'swimming wings', *peign-oir* 'robe', *lav-ette* 'washcloth', *balanç-oire* 'swing') (13/19 in Sophie's corpus, 12/15 in Emma's corpus). Other pieces of evidence for emerging morphology are provided by examples such as (1a, b, c) which shows the child's play with morphologically related words, as well as by errors which display decomposition of base and suffix (see Section 4.1.3).

(1) Emma 2;1:

- a. *l'* *ai* *lava* [: *lavé*] *avec la lavette*;
 PRON.OBJ.3SG avoir.AUX.PRS.1SG laver-PP with the washcloth;
 'washed it/him with the washcloth'
- b. *je* *lave* *la lavette*;
 PRON.SBJ.1SG laver.PRS.SG the washcloth;
 'I am washing the washcloth'
- c. *j'* *ai* *lavé* *avec la lavette*.
 PRON.SBJ.1SG avoir.AUX.PRS.1SG laver-PP with the washcloth.
 'I washed with the washcloth'

The following Tables 5a and 5b focus on suffix diversity by presenting the series of suffixed nouns occurring in the data. In the four subcorpora, the largest series is by far the series in *-ette* which goes up to 20 lemmas (Emma's CDS). The other series contain no more than 10 different lemmas.

Table 5a. Suffixed nouns (at least 2 bases): Corpus Sophie

SOP CS (8 SUFFIXES)						SOP CDS (12 SUFFIXES)					
Lem.	Tok.	Semant. categ.	Base (lem.)*		Age em.**	Lem.	Tok.	Base (lem.)*			
			N	V				N	V		
<i>-ette</i>	17	169	OBJ, INSTR	11	9	1;9	<i>-ette</i>	19	187	14	8
<i>-oir(e)</i>	7	15	INSTR	1	6	2;2	<i>-oir(e)</i>	9	47	1	8
<i>-on</i>	5	41	INSTR	3	2	2;3	<i>-on</i>	7	51	5	2
<i>-ure</i>	4	22	OBJ, RESULT	0	4	2;3	<i>-ure</i>	4	23	0	4
<i>-age</i>	3	14	ACTION, LOC	1	2	2;9	<i>-age</i>	3	6	0	3
<i>-ard</i>	2	20	OBJ, INSTR	1	1	2;9	<i>-ard</i>	3	10	2	1
<i>-ade</i>	2	2	ACTION	0	2	2;9	<i>-(i)er</i> AGENT, OBJ	3	26	2	1
<i>-ne</i>	2	13	GENDER	2	0	2;10	<i>-eau</i> OBJ	3	5	3	0

Table 5a. (continued)

SOP CS (8 SUFFIXES)					SOP CDS (12 SUFFIXES)					
Lem.	Tok.	Semant. categ.	Base (lem.)*		Age em.**	Lem.	Tok.	Base (lem.)*		
			N	V				N	V	
						<i>-ne</i>	2	31	2	0
						<i>-eur</i> INSTR	2	23	0	2
						<i>-et</i> DIM	2	12	2	0
						<i>-esse</i> GENDER	2	4	2	0
Total	42	296		19	26		59	425	33	29

* As mentioned in 3.2.1, some bases of suffixed words in *-ette* are undecidable (N or V).

** At least 2 lemmas.

Table 5b. Series of suffixed nouns (at least 2 bases): Corpus Emma

EMM CS (8 SUFFIXES)						EMM CDS (17 SUFFIXES)				
Lem.	Tok.	Semant. categ.	Base (lem.)		Age em.*	Lem.	Tok.	Base (lem.)		
			N	V				N	V	
<i>-ette</i>	11	46 INSTR, OBJ	6	9	1;8	<i>-ette</i>	20	91	14	11
<i>-on</i>	7	38 DIM, INSTR, AUG	4	3	1;8	<i>-on</i>	10	97	7	3
<i>-oir(e)</i>	4	31 INSTR	2	2	2;1	<i>-ure</i>	8	29	0	5
<i>-(i)er</i>	4	15 OBJ, AGENT	1	3	2;6	<i>-(i)er</i>	8	24	8	0
						<i>-eau</i>	7	11	3	4
<i>-et</i>	3	25 OBJ, DIM	3	0	1;10	<i>-et</i>	6	28	4	2
<i>-age</i>	2	6 RESULT ACTION	0	2	2;5	<i>-age</i>	6	11	0	6
<i>-(i)ère</i>	2	2 LOC-INSTR	1	1	2;9	<i>-oir(e)</i>	5	25	0	5
<i>-eur</i>	2	2 INSTR-AGENT	0	2	2;4	<i>-eur</i>	4	10	0	4
						<i>-ne</i> GENDER	4	10	0	4
						<i>-(i)ère</i>	3	5	2	1
						<i>-erie</i>	3	14	2	1
						LOC-ACTIV				
						<i>-ée</i> TIME	3	4	3	0
						ACTION				
						<i>-ère</i> GENDER	2	2	2	0
						<i>-ine</i> DIM	2	14	2	0
						<i>-ment</i> ACTION	2	6	0	2
						<i>-euse</i> GENDER	2	2	1	1
Total	35	165		17	23		95	383	48	49

** At least 2 lemmas.

First of all, it is interesting to notice that while each corpus displays the specific features mentioned above, i.e. greatest suffix diversity in Emma's CDS and to a certain extent in Emma's CS in spite of more limited data, and more homogeneity in Sophie's corpus, the general picture is one of strong resemblance between the two corpora. In Sophie's corpus the parallel between CS and CDS is striking: the five suffixes having the greatest diversity, i.e. forming the largest series, *-ette*, *-oir(e)*, *-on*, *-ure* and *-age* are the same and are ranked in the same order. The triplet *manch-ette*, *manch-on*, *manch-oir* 'swimming-wings' in CS (and CDS) illustrates the relevance of the three suffixes in the production of the child. Differences pertain to suffixes with low diversity (*-(i)er*, *-eau* and *-eur*, *-et*, *-esse*, as well as *-ade*). In terms of vocabulary, the relation between CS and CDS is straightforward too: with the exception just mentioned and a few other words, the same suffixed words of series are also found in CDS. This is also true of Emma's corpus.

In Emma's corpus, however, the picture is slightly different. While all series of Emma's CS are found in CDS, the ranking of suffixes is close but not identical in CS and CDS (*-ier* and *-oir(e)*). Notice that out of the 9 suffixes of Emma's CDS that do not find a correspondence in Emma's CS, 6 occur with a single lemma (*-ure*, *-eau*, *-erie*, *-ée*, *-ine*). In this way, taking the large inventory of suffixes of Emma's CDS into account, CS's suffixes are parallel to CDS. Curiously, no derived word constructed with the productive suffix *-ade* appears in Emma's corpus. Nonetheless, overall the inventory of series/suffixes is very similar in both CDS.

In CS, the three most-highly ranked suffixes *-ette*, *-on* and *-oir(e)* are identical. Most importantly, these correspondences between the two CS are not simply due to vocabulary correspondences since in each CS several lemmas are different: for instance, *manch-ettes* 'swimming-wings', *sonn-ette* 'bell', *plonge-on* 'dive', *mouch-oir* 'handkerchief' and *tir-oir* 'drawer' occur in Sophie's CS only, whereas *roul-ettes* 'caster', *croqu-ette*, *viol-on* 'violin' and *can-eton* 'duckling' belong to Emma's. They are neither just determined by the child-centered situation nor by the reference to the world of children: almost three quarters of the lemmas with diminutive suffix (e.g. *chauss-ettes* 'socks', *servi-ette* 'table napkin', *tromp-ette* 'trumpet', *fouch-ette* 'fork') belong to the ADS lexicon. Nevertheless, given the strong parallelism between the derived nominal vocabulary of CS and CDS, one cannot exclude that the items have been simply memorised without paying attention to their morphological structure and the morphological series they belong to. The size of the largest series in *-ette* could be seen as due to the great number of lexicalised diminutives which are commonly used in everyday language and are thus expected in CDS. Indeed, all lexemes displaying the diminutive suffixes are lexicalised, with one exception in the corpus of Sophie (the imitation of a diminutivised nominalised adjective used by the mother: *fol-ette* 'scatterbrain'). However, they are the first derived words to emerge and very early form a series (Sophie 1;8, *cass-ette* 'tape', *chauss-ettes*

‘socks’, 1;9 *lun-ettes* ‘glasses’, Emma 1;6 *servi-ette* ‘towel’, *bav-ette* ‘bib’, 1;8 *pouss-ette* ‘scroller’, *chauss-ettes* ‘socks’) and families, as shown above. This suggests that the ending, more specifically the suffix, is relevant (see 4.1.3). Moreover, *-ette* and *-oir(e)* meet the criterion of potential productivity based on the occurrence with three different stems and recurrence of the stem in other words along with a few other suffixes (see Introduction to the volume):

Sophie’s CS: *-ette, -oir(e), -on*
 Emma’s CS: *-ette, -on, -oir(e), -et, -ier.*

In terms of development, one notices a six-month gap in Sophie’s CS between the first series in *-ette, -oir(e), -on* and *-ure* (1;9–2;3) and the other series in *-age, -ard, -ade* and *-ne* (from 2;9 on). In Emma’s CS the emergence of series is more evenly distributed.

4.1.2 Adjectival suffixes

It is well-known that adjectives develop later than nouns and verbs and are thus less frequent in early child language (see e.g. Tribushinina et al. 2013; Tribushinina, Voeikova & Noccetti 2015). Suffixed adjectives represent between 1% and 4% of adjective lemmas (Table 6), which is a significantly lesser proportion than suffixed nouns, except in the case of Sophie’s CDS (Sophie’s CS: $\chi^2 = 3.49$, $df = 1$, $p = 0.03$; Emma’s CS: $\chi^2 = 3.42$, $df = 1$, $p = 0.03$; Sophie’s CDS: $\chi^2 = 2.28$, $df = 1$, $p = 0.07$; Emma’s CDS: $\chi^2 = 9.36$, $df = 1$, $p = 0.001$). They appear later than noun suffixes in development (Sophie’s CS from 2;1 on, Emma’s CS from 2;3 on).

Table 6. Suffixed adjectives

	No. suffix types		No. suffixed ADJ (at least 2 bases)		% of ADJ	No. ADJ total	
	At least 2 bases	Isolated bases	Lem.	Tok.	Lem.	Lem.	Tok.
Sophie’s CS	0	1	1	6	1.4	71	1,241
Emma’s CS	0	5	0	0	0	70	879
Sophie’s CDS	2	0	4	10	4.1	97	1,866
Emma’s CDS	1	9	1	2	1.1	93	1,179

* The lemma *japon-ai-se* ‘japanese’ has 2 suffixes

4.1.3 Errors

A small number of errors (10 in Sophie’s CS and 6 in Emma’s CS) affecting the morphological structure of suffixed words have been found in the data. They can be classified into the following types:

- a. six substitutions of the base by nonsense forms within a series producing non-target-like words, e.g. Sophie 2;2 *baqu-ette*, *paqu-ette* for *casqu-ette* ‘cap’; 2;2 *beubav-oir* for *arros-oir* ‘watering can’; Emma 1;10 *jav-ette* for *bavette* ‘bib’.
- b. Three substitutions of the base within a series, e.g. Sophie 2;4 *barr-ette* ‘barrette’ for *bav-ette* ‘bib’, Emma 1;6 *bav-ette* ‘bib’ for *serv-i-ette* ‘table napkin’.
- c. Three substitutions of a simplex by a complex word: Emma 2;1 *bûch-eron* ‘wood cutter’ for *bûche* ‘log’, 2;5 *ind-i-enne* ‘Indian’ for *Inde* ‘India’, or by an approximation of a complex word: Sophie 2;1 *bat-ette*, *a tass-ette* [: *casquette*] ‘cap’ for *casque* ‘helmet’.
- d. One blend of two complex words based on morphotactic similarity: Sophie 2;9 non-target-like *convert-ure* for *confit-ure* ‘jam’ (*couvert-ure* ‘blanket’). The initial syllable of *confit-ure* ‘jam’ is followed by the penultimate and final syllable of *couvert-ure* ‘blanket’ that occurs at 2;2.

What we see in a. and b. is that the errors alter the stem but preserve the suffix. Therefore, they may presuppose decomposition of the target. The question is the nature of the decomposition and whether it may be phonological rather than morphological, e.g. motivated by the salience of the ending. The argument in favour of morphological decomposition and suffix generalisation rather than uptake of a phonologically salient string is that the latter would imply extraction of a syllable and not only the rhyme, e.g. *-et* [ɛ] and *-oir* [war], as in the example of Sophie 2;1 *vette* for *lavette* ‘washcloth’. But in most of the examples, it is only *-ette* which is retained.

In addition, we found two examples of a change of a complex word into a simplex one, in prefixation (see Section 4.2): Sophie 1;9 CDS (*il faudra que je le re-colle* ‘(I will have to) glue (it) back together’ is repeated as *atolle*, with a filler most probably replacing the prefix, or, alternatively, meant for *ça* ‘that’ as in the frequent sequence/item-based construction *ça colle* ‘it sticks’ occurring from 2;6 on. At 1;8 Emma uses *coller* ‘stick’ (*pas coller* ‘not stick’) instead of *décoller* ‘unstick’ as corrected by the mother.

This minimally suggests that the children may have identified the suffixes and take the shape of words and their morphological complexity into account.

4.1.4 *Development of semantic categories*

As expected, in both CS object and instrument nouns (see Namer & Villoing 2008; Ferret & Villoing 2015 for a characterisation of the semantic categories and a review of the literature) are early whereas action and gendered nouns are late. There is no occurrence of agent noun in Sophie’s CS. In Emma’s, however, the category of agent noun emerges earlier than that of action noun.

Affixed N, ADJ and V are more often morphosemantically transparent than opaque in CS, even more so in CDS. Transparent suffixed nouns represent more than 60% of the lemmas forming series in CS. However, quite paradoxically, in the

largest series of suffix *-ette*, about half of derived nouns (9 out of 17 lemmas, 53%) are rather opaque in Sophie's and Emma's CS. However, the low number of items casts a doubt on this result's representativity: 95% CI [28%, 77%]. These lexicalised words in *-ette* (*cass-ette* 'tape', *assi-ette* 'plate', *servi-ette* 'table napkin', *fouch-ette* 'fork', *chauss-ettes* 'socks') are the first derived words to appear in Sophie's CS as they denote very familiar or pragmatically important (e.g. *cass-ette* 'tape') objects or instruments in the child's life. It is not before 2;4/2;5 that transparent derived words are produced and are accurate (2;4 *din-ette* 'playing dinner party', *bav-ette* 'bib', 2;5 *pouss-ette* 'stroller', *suc-ette* 'lolly pop') along with other transparent derived words *mass-age* 'massage', *balanç-oire* 'swing', *sal-eté* 'dirtiness'. In Emma's CS, on the other hand, the development cannot easily be characterised in relation to transparency/opacity: both types of derived words occur from the beginning.

4.2 Prefixation

Prefixation applies almost exclusively to verbs (recall that suffixed verbs are practically not found in the investigated data). A single instance of a prefixed noun occurs in the corpus, i.e. *re-change* 'change' in Emma's CS and CDS. As to prefixed adjectives, a single example from Emma's CDS (*dés-agré-able* 'unpleasant') is attested.

Prefixed verbs have chiefly two prefixes only: iterative or reversative *re-* (e.g. *re-voir* 'see again', *re-mettre* 'put after removing') and privative *dé-* (e.g. *décoller* 'take off', *déshabiller* 'undress') (Tables 7 and 8). They emerge in the third year, at 2;1 in Emma's CS and 2;6 in Sophie's corpus (see also Clark 1993: 234). All prefixed verbs have their corresponding simplex word in the corpora and appear after it. Moreover they are morphosemantically transparent. In other words, we found clear signs of potential productivity in prefixed verbs. Nonetheless, in comparison to suffixed nouns, the proportion of prefixed verbs, especially of tokens, to all verbs is significantly lower (less than 1%) (Sophie's CS: $\chi^2 = 362.94$, $df = 1$, $p < 0.0001$; Emma's CS: $\chi^2 = 162.34$, $df = 1$, $p < 0.0001$; Sophie's CDS: $\chi^2 = 852.48$, $df = 1$, $p < 0.0001$; Emma's CDS: $\chi^2 = 512.26$, $df = 1$, $p < 0.0001$).

Table 7. Proportions of prefixed verbs: Sophie's and Emma's corpus

	No. prefix types		No. prefixed V (at least 2 bases)		% of V	No. V total	
	At least 2 bases	Isolated	Lem.	Tok.	Lem.	Lem.	Tok.
Sophie's CS	2	2	17	21	9.1	187	6,459
Emma's CS	2	1	9	34	4.4	205	4,757
Sophie's CDS	2	2	19	35	6.8	278	13,595
Emma's CDS	2	1	16	65	5.4	294	8,455

Table 8. Series of prefixed verbs

SOP CS (2 PREFIXES)						SOP CDS (2 PREFIXES)			
	Lem.	Tok.	Semant. categ.	Base	Age em.*		Lem.	Tok.	Base
<i>re-</i>	15	19	ITER, REVERS	V	2;5	<i>re-</i>	14	2	V
<i>dé-</i>	2	2	PRIV	V	2;11	<i>dé-</i>	3	3	V
EMM CS (2 PREFIXES)						EMM CDS (2 PREFIXES)			
<i>re-</i>	7	27	ITER	V	2;1	<i>re-</i>	14	24	V
<i>en-</i>	2	6	DIR	V	2;5	<i>dé-</i>	3	8	V

* At least 2 lemmas.

4.3 Conversion

Given the extent of conversion in French (see Section 2), the part of potential conversion relations in the vocabulary of toddlers is predicted to be quite widespread (depending also on the theoretical delimitation of the notion, see Section 2).

4.3.1 *N-V pairs*

4.3.1.1 *Homophonous N-V pairs*

Let us start with semantically related N-V pairs which are homophonous. The analysis of homophonous pairs encounters a very common methodological problem of early acquisition studies, namely the indeterminacy of the category/class of early words (see Bloom 1970; Ambridge & Lieven 2015; Veneziano 2017 among many others). For instance, it is impossible to assess whether Sophie 1;9 [dut] for *douche* 'shower' is a noun or a verb form. N-V pairs can only be identified when contextual cues allow a non-ambiguous reading of the word. Thus, in Sophie's CS, it is not before 2;0 that N-V pairs can be safely documented (see Table 10a).

The number of pairs found in the data is displayed in Table 9.

Table 9. Homophonous N-V pairs

	No. N-V pairs	No. V lemma total	% of V lemmas	95% CI	No. N lemma total	% of N lemmas	95% CI
Sophie's CS	9	188	4.8	[2.2, 8.9]	494	1.8	[0.8, 3.4]
Emma's CS	10	205	4.9	[2.4, 8.8]	572	1.7	[0.8, 3.2]
Sophie's CDS	12	278	4.3	[2.3, 7.4]	631	1.9	[1, 3.3]
Emma's CDS	19	294	6.5	[3.9, 9.9]	780	2.4	[1.5, 3.8]

Neither the proportion of verb lemmas nor the proportion of noun lemmas in N-V pairs varies significantly across corpora (V: $\chi^2 = 1.5$, $df = 3$, $p = 0.68$; N: $\chi^2 = 1.04$, $df = 3$, $p = 0.79$). The slight difference between Emma's CS and CDS on the one hand and Sophie's CS and CDS figures on the other hand, is thus probably irrelevant.

The fact that the number of homophonous N-V pairs found in the data is rather small may appear to question their relevance, in particular when considering the relatively large confidence intervals. In addition their proportion is significantly lower than the proportion of suffixed nouns (Sophie's CS: $\chi^2 = 21.17$, $df = 1$, $p < 0.0001$; Emma's CS: $\chi^2 = 13.32$, $df = 1$, $p = 0.0001$; Sophie's CDS: $\chi^2 = 31.58$, $df = 1$, $p < 0.0001$; Emma's CDS: $\chi^2 = 53.23$, $df = 1$, $p < 0.0001$). However, we can notice that among Ns occurring in CS without the corresponding Vs and of Vs occurring without the corresponding Ns, those which may form plausible pairs at this age represent a lower number than the number of attested pairs. We can enumerate in Sophie's CS on the one hand 4 verbs (*coller* 'stick' 11 tokens, *danser* 'dance' 13 tokens, *griffer* 'scratch' 3 tokens (1 token of sg. pres. ind.), and *taper* 'hit' 6 tokens) for which one could expect the nouns *colle* 'glue', *danse* 'dance', *griffe* 'claw', *tape* 'tap', that do not occur in Sophie's CS.

On the other hand there are 3 nouns (*coche* 'tick', *lange* 'diaper', *scotch* 'scotch tape') for which the corresponding verbs are not found in our CS data.

In Emma's CS, 3 nouns (*mousse* 'froth', *pince* 'pliers' and *scotch* 'tape') do not have a plausible verb counterpart (*mousser* 'froth', *pincer* 'nip', *scotcher* 'tape') in her data, and 2 verbs (*caresser* 'caress', *crier* 'scream') occur without the corresponding noun (*caresse* 'caress', *cri* 'scream').

In CDS only isolated examples of some of the members of the pairs lacking in CS have been found (Sophie's CDS: V *scotcher*, Emma's CDS: V *pincer* and N *cri*).

As a counter-example, one would not expect a young child to use the nouns corresponding to common verbs like *regarder* 'look', *demander* 'ask for', *marcher* 'walk', *aider* 'help' which are abstract nouns, neither the verbs *cuisin-er* 'cook', *tartin-er* 'spread' corresponding to Ns such as *cuisine* 'kitchen', *tartine* 'spread', etc. which are much more specific. Moreover, in French everyday speech, these verbs are often replaced by the structure *faire* + N 'do + N' (*faire la cuisine*, etc., lit. 'do kitchen').

The homophonous pairs are presented in Tables 10a and 10b, in which they are arranged in ascending order according to age of emergence of the earliest member and time distance between the two members. We have included some doubtful pairs in which the homophonous member either contains additional material (i.e. is a reduplication or a multilexical unit, e.g. *canne à pêche* 'fishing rod') or is semantically unrelated (e.g. *goûter* used with the meaning of 'taste' and not 'eat after-school snack').

Table 10a. Age of emergence of homophonous N-V pairs in Sophie's CS

Age of first N	Age of first V	Lemmas	Gloss
1;7	2;9	<i>téléphone – téléphon-er</i>	'phone'
1;10	2;5	<i>habit – habill-er</i>	'item of clothing, dress'
1;11	2;0/2;5	<i>bagarre – se bagarr-er</i>	'fight'
2;7	2;6	<i>caresse – caress-er</i>	'caress'
2;7	2;6	<i>travail – travaill-er</i>	'work'
1;10/2;10	2;6	<i>brosse – bross-er</i>	'brush'
2;8	2;8	<i>foehn – foehn-er</i>	'hairdryer, hairdry'
Doubtful pairs			
2;3	2;3	<i>pousse-pousse – pouss-er</i>	'stroller, push'
2;4	2;4	<i>gôûter – goût-er</i>	'after-school snack, to taste'
Total: 9 pairs (2 doubtful)			

Table 10b. Age of emergence of homophonous N-V pairs in Emma's CS

Age of first N	Age of first V	Lemmas	Gloss
2;4	1;8	<i>habit – habill-er</i>	'item of clothing, dress'
2;9	1;9	<i>griffe – griff-er</i>	'claw'
1;9	1;10	<i>travail – travaill-er</i>	'work'
2;5	1;10	<i>plongée – plong-er*</i>	'diving, dive'
1;10	2;1	<i>déjeuner – déjeuner-er</i>	'breakfast'
2;3	2;3	<i>danse – dans-er</i>	'dance'
2;5	2;5	<i>pêche – pêch-er**</i>	'fishing, catch'
Doubtful pairs			
2;0	1;10	<i>cache(-cache) – (se) cach-er</i>	'hide-and-peek, hide'
1;10	2;1	<i>pousse-pousse – pouss-er</i>	'stroller, push'
2;0	2;1	<i>brosse (à dents) – bross-er</i>	'toothbrush, brush'
2;5	2;5	<i>canne à pêche – pêch-er (see above)**</i>	'fishing rod, 'catch'
Total: 10 pairs (4 doubtful)			

* Final vowel lengthening in the noun, that is typical of French spoken in the area of Switzerland where the child lives, has not been taken into consideration.

** *pêche – pêcher* and *canne à pêche – pêcher* are counted as a single pair.

As far as the relation to CDS is concerned, results of both corpora show once again a strong correspondence between CS and CDS and more specifically between frequency in CDS and age of emergence in CS. Not only are all the CS's pairs related to the CDS's pairs (although, as expected, CDS has a greater number of pairs than CS), but the first member of a pair to emerge is generally the most frequent member of the pair in CDS.

Furthermore, we have found a clear tendency for members of a homophonous pair to co-occur in a same file or sequence of interaction in CDS, and to a lesser extent in CS.⁹

These findings seem to indicate that for CDS and CS the N and V lemmas of the homophonous pairs are related. The picture is corroborated by the results of the N-V pairs with modification presented in the next section.

4.3.1.2 *N-V pairs with modification*

A few nearly homophonous N-V pairs display two types of recurrent modification (final consonant or glide addition/subtraction and vowel alternation in the base, see 2.2). Although they represent at most 1% of the total of nouns and a slightly higher percentage of the verbs, they are worth mentioning here insofar as their development seems to be parallel to the development of homophonous pairs (Table 11).

Table 11. N-V pairs with modification

	No. N-V pairs	No. V lemma total	% of V lemmas	95% CI	No. N lemma total	% of N lemmas	95% CI
Sophie's CS	5	188	2.7	[0.9, 6]	494	1.0	[0.3, 2.3]
Emma's CS	4	205	2	[0.5, 4.9]	572	0.7	[0.2, 1.8]
Sophie's CDS	4	278	1.4	[0.4, 3.6]	631	0.6	[0.2, 1.6]
Emma's CDS	4	294	1.4	[0.4, 3.4]	780	0.5	[0.1, 1.3]

These pairs can be chronologically ranked in the same way as strictly homophonous pairs of Tables 10a and 10b (see Tables 12a and 12b).

Table 12a. Age of emergence of N-V pairs with modification in Sophie's CS

Age of first N	Age of first V	Lemmas	Phonetic transcription	Gloss
1;6	2;8	<i>bain – (se) baign-er</i>	[bɛ̃] – [beɛɛ]	'bath, take a bath'
2;3	1;9	<i>tour – tourn-er</i>	[tur] – [turne]	'turn'
2;2	1;11	<i>dessin – dessin-er</i>	[dɛsɛ̃] – [dɛsine]	'draw'
2;6	2;11	<i>pet – pét-er</i>	[pɛ] – [pɛte]	'fart'
1;10	2;10	<i>savon – savonn-er</i>	[savɔ̃] – [savɔɛɛ]	'soap, wash'

9. Due to space constraints, we leave the detailed analysis of the two issues we have just mentioned for another publication.

Table 12b. Age of emergence of N-V pairs with modification in Emma's CS

Age of first N	Age of first V	Lemmas	Phonetic transcription	Gloss
1;5	1;10	<i>bain</i> – (<i>se</i>) <i>baign-er</i>	[bɛ̃] – [bɛ̃ʁe]	'bath, take a bath'
1;10	1;8	<i>dessin</i> – <i>dessin-er</i>	[dɛsɛ̃] – [dɛsine]	'draw'
1;10	2;1	<i>savon</i> – <i>savonn-er</i>	[savɔ̃] – [savɔ̃ne]	'soap, wash'
2;1	2;0	<i>tour</i> – <i>tourn-er</i>	[tur] – [turne]	'turn'

With the exception of one (playful) pair in Sophie's CS (*pet* – *pét-er* 'fart'), the pairs found in both children are the same. These pairs again correspond for the most part to the pairs that occur in CDS (in Sophie's CDS the playful one plus another one differ). It is interesting to notice that they also display co-occurring members in a same recording session.

The data show that a few bases of N-V pairs are also bases of suffixed words (Sophie's CS and CDS 3, Emma's CS 2 and Emma's CDS 4). It follows that the size of morphological families in the corpora reach 4 members once homophonous and nearly homophonous N-V pairs are included (see Section 4.1.1), e.g. Sophie's CS *bain* 'bath', (*se*) *baign-er* 'take a bath', *baign-oire* 'bathtub', *baign-ade* 'bathing'.

In sum, we may assume that the N-V pairs (homophonous and with modification) that are documented in the data highlight the fact that the lexicon of the children and the parents takes advantage of the connections between N and V. These connections, which potentially represent conversions, contribute to the emerging morphological network which seems to characterise the vocabulary of toddlers by the end of the third year.

4.3.2 *Verbs derived from adjectives*

A very small number of verbs derived from adjectives have been found: in both CS these verbs are *vid-er* 'empty', *sal-ir* 'dirty' and, with modification of the stem *sec* 'dry', *séch-er* 'dry'.

4.3.3 *Adjectives converted from verbs*

(*PP*, *PPres*, *sg. pres. ind.*)/*deverbal adjectives*

We have seen that there is a low amount of suffixed adjectives in the corpora. Some adjectives derived from verbs complement them (see Table 13).

Most of the deverbal adjectives are derived from past participles of the first conjugation class (e.g. *fatigué* 'tired', *mouillé* 'wet', *fâché* 'angry') but a couple of them derives from another class (*tordu* 'twisted', *fichu* 'broken') or from a present

Table 13. Deverbal adjectives

	No. deverbal ADJ		% of ADJ	No. ADJ total	
	Lemmas	Tokens	Lemmas	Lemmas	Tokens
Sophie's CS	7	13	9.9	71	1,241
Emma's CS	6	30	8.6	70	879
Sophie's CDS	11	53	11.3	97	1,866
Emma's CDS	9	48	9.7	93	1,179

participle (*brûlant* 'hot', CDS *intéressant* 'interesting'). None of the corresponding verb bases occurs in the corpora. Their development seems to be purely lexical.

4.4 ADS

The sample of ADS data consists in two recordings of approximately 2 hours of conversation between friends, respectively Sophie's mother and the first author, and Emma's father (both present during the recording sessions of the children), the first author and another person (Table 14).

Table 14. Sample of ADS

Speaker	No. of word lemmas	No. of word tokens	No. of utterances
Sophie's mother	2,115	20,537	1,711
Emma's father	1,974	16,918	1,420

As it appears in Tables 16 and 17, results on proportions of suffixed nouns and on the distribution of ADS suffixes display a striking consistency between the two speakers. The difference that was observed between Sophie's CDS and Emma's CDS is not found in the most frequent suffixes of ADS. In both adults' production, the proportion of suffixed noun lemmas is much higher than in their CDS – more than the double in the speech of Sophie's mother, a little less than the double in Emma's father's (Sophie's ADS: $\chi^2 = 30.09$, $df = 1$, $p < 0.0001$; Emma's ADS: $\chi^2 = 18.26$, $df = 1$, $p < 0.0001$).

Notice that this result cannot be interpreted as suggesting that the parents do not follow a noun bias: the proportion of nouns in CDS is higher than in ADS and is more than twice as large as the proportion of verbs (Table 15).

Table 15. Proportion of nouns and verbs

	N		V		Words
	No. lemmas	%	No. lemmas	%	No. lemmas
Sophie's CS	498	45	189	17.1	1,106
Emma's CS	572	49.5	205	17.7	1,155
Sophie's CDS	631	48.8	278	21.5	1,292
Emma's CDS	780	53.5	294	20.2	1,459
Sophie's ADS	651	40.5	345	21.5	1,606
Emma's ADS	624	31.6	317	16.1	1,974

Table 16. Suffixed nouns in ADS

Speaker	No. suffixes with at least 2 bases	No. suffixes ADS only	Suffixed nouns with at least 2 bases	% of N	No. N total
Sophie's mother	18	13	132	20.5	645
Emma's father	24	13	129	20.7	622

Table 17a. ADS noun suffixes: Sophie's mother

ADS-specific suffixes (with at least 2 bases): 9			Suffixes in ADS & CDS (1 lemma incl): 9		
Suffix	No. lemmas	Semantic category	Suffix	No. lemmas in ADS	Semantic category
<i>-ion</i>	26	RESULT – ACTION	<i>-(i)té</i>	15	QUALITY*
<i>-ance/-ence</i>	17	ACTION – STATE	<i>-ment</i>	13	ACTION
<i>-isme</i>	9	QUALITY – ACTION	<i>-eur</i>	6	AGENT
<i>-ain</i>	4	ETHNIC – QUALITY – AGENT	<i>-age</i>	6	ACTION
<i>-iste</i>	3	AGENT	<i>-ure</i>	6	RESULT – ACTION – COLL
<i>-ais</i>	3	ETHNIC	<i>-ique</i>	4	QUALITY
<i>-esse</i>	3	QUALITY – ETHNIC	<i>-ne</i>	4	GENDER
<i>-aire</i>	2	RESULT – QUALITY	<i>-ette</i>	1	DIM
<i>-ie</i>	2	QUALITY – ACTION	<i>-ade</i>	1	RESULT

* Or PROPERTY which is synonymous

Table 17b. ADS noun suffixes: Emma's father

ADS-specific suffixes (with at least 2 bases): 9			Suffixes in ADS & CDS (1 lemma incl): 17		
Suffix	No. lemmas	Semantic category	Suffix	No. lemmas in ADS	Semantic category
<i>-ion</i>	32	RESULT – ACTION	<i>-ance/ -ence</i>	14	ACTION – STATE
<i>-iste</i>	3	AGENT	<i>-ment</i>	12	RESULT – ACTION
<i>-at</i>	3	RESULT – INSTITUTIONAL FUNCTION	<i>-ateur</i>	8	AGENT – INSTR
<i>-aire</i>	3	QUALITY	<i>-(i)té</i>	8	QUALITY
<i>-ien</i>	3	AGENT	<i>-age</i>	7	ACTION – COLL
<i>-ie</i>	3	QUALITY – LOC	<i>-ique</i>	3	QUALITY
<i>-isme</i>	2	DOCTRINE – COLLECTIVE	<i>-erie</i>	3	COLLECTIVE – EVALUATIVE
<i>-ant</i>	2	AGENT	<i>-ette</i>	3	OBJ – INSTR
<i>-aine</i>	2	NUM	<i>-oire</i>	2	LOC
			<i>-(t)ure</i>	3	RESULT – INSTR
			<i>-ais</i>	2	ETHNIC
			<i>-ant</i>	2	AGENT
			<i>-ise</i>	2	QUALITY
			<i>-ée</i>	2	TEMP – INSTR
			<i>-(i)er</i>	1	COLL
			<i>-re</i>	1	GENDER
			<i>-ne</i>	1	GENDER

Tables 17a and b show not only that the number of ADS noun suffixes is similar in both adults, but that the suffixes yielding the largest series are the same (with the exception of *-ure* in Sophie's mother):

S. *-ion, -ance, -ité, -ment + -isme, -eur, -age, -ure*

E. *-ion, -ance, -ment + -ateur, -ité, -age*

As expected, the suffixes of ADS form abstract nouns. The difference with the suffixes showing the greatest diversity in CDS, i.e. the absence of *-oir(e)* and *-on* and the scarcity of *-ette*, denoting objects and instruments, can also be accounted for by the speech situation and topics of conversation.

Regarding noun suffixes that are shared with CDS, as expected diminutives are much less represented. Indeed there is a single lemma (*boulette* 'blunder') in the speech of Sophie's mother and three lemmas (*cassette* 'tape', *lavette* 'washcloth', *sandalette* 'sandal') in that of Emma's father.

It is noticeable that in this short recording sample of ADS, suffixed verbs are found in both parents while there is none in CS and CDS: 2 lemmas in the speech of Sophie's mother, 4 in that of Emma's father. In all examples it is the productive suffix *-iser* (*util-is-er* 'use', *sensibil-is-er* 'raise awareness among', *économ-is-er* 'save', etc.) which is used.

Another difference with CS and CDS comes from nominal prefixes which again occur in both ADS, i.e. iterative *re-*, privative *dés-/dis-*, temporal *pré-* and excessive *sur-*.

With regard to verbal prefixes, iterative *re-* and privative *dé(s)-* are found as in CS and CDS, to which directional *in-* (*im-merger* 'immerse') and *é-* (*é-courter* 'shorten') and temporal *pré-* (*pré-aviser* 'give advance notice') are added.

5. Discussion

Although we found a greater morphological diversity and more heterogeneity between CS and CDS in Emma's corpus, the intra- and inter-similarity of the results in our two corpora of CS and CDS on noun, adjective and verb affixation, as well as ADS on noun and verb affixation, as well as on homophonous pairs, is striking given the differences in caretakers' style, recording context and family situation (single child vs. third child). Nominal suffixation displays the greatest number of derivational series, the same nominal suffixes produce the largest series and the ranking is very similar (see Tables 5a and 5b); prefixes appear with verbs and are confined to two productive ones (Table 8), except in ADS, and affixed adjectives are isolated (Table 6). Conversion relations between nouns and verbs, between verbs and adjectives, as well as between a few adjectives and verbs converge in number and types. Moreover, the investigated derivational means have a similar order of emergence in the two corpora. Suffixed words and conversion relations are documented before the less frequent prefixed ones.

Adjectival suffixes appear later than noun suffixes in development: in Sophie's CS the single example of a derived adjective (*délic-ieux* 'delicious' from 2;1 on) emerges later than the first suffixed nouns in *-ette* (1;9) but earlier than the prefixed verbs (2;5) (see Section 4.2). In Emma's CS, 9 of the 10 adjectival derived lemmas emerge from 2;3 on, i.e. later than the first three types of nominal suffixes (1;8, 2;1). Suffixed adjective lemmas are significantly less frequent in CDS (Sophie's CDS 4.1%, Emma's CDS 1.1%) than in ADS (Sophie's ADS: 12.6%; Emma's ADS: 23.9%) (Sophie's parents: $\chi^2 = 4.57$, $df = 1$, $p = 0.02$; Emma's parents: $\chi^2 = 22.41$, $df = 1$, $p < 0.0001$). Further research should assess whether this speaks for fine-tuning in CDS.

Can we identify an age which would gather converging signs of detection of derivational morphology? At 2;2 we see that Sophie displays the first N – V pairs

and she has already produced several errors on morphological complex words (5 out of 10). However, her four richest series are formed at 2;3 so that this seems a safer age. In Emma's data, 2;1 seems to be a turning point: the 3 richest series plus one and N – V pairs are attested, most errors appear, and so do prefixed verbs. In both children this corresponds to late protomorphology, more precisely to one month before morphology proper (see Kilani-Schoch 2017).

Turning to the comparison between noun suffixation and nominal compounding (Kilani-Schoch 2017), we see that the data support the hypothesis of a preference for suffixation over compounding in CS and CDS, as well as in ADS (Table 18).

Table 18. Suffixed N vs. N compounds (lemmas)

	No. suffixed N total (1 lemma included)*	% of N	No. N strict compounds & multilexical units	% of N	No. N total
Sophie's CS	54	10.9	18	3.6	494
Emma's CS	48	8.4	33	5.8	572
Sophie's CDS	71	11.3	27	4.3	631
Emma's CDS	110	14.1	50	6.4	780
Sophie's ADS	132	20.3	26	4.0	651
Emma's ADS	129	20.7	38	6.1	624

* Suffixed N include those in which the suffix occurs with at least two lemmas as well as isolated suffixed lemmas

As expected, on average the proportion of lemmas of suffixed nouns is higher than the proportion of compound lemmas in relation to the total of noun lemmas: between 1.5 times higher, in Emma's CS, and 3 times higher, in Sophie's CS, although the former difference is not statistically significant (Sophie-CS: $\chi^2 = 9.38$, $df = 1$, $p = 0.001$; Emma-CS: $\chi^2 = 0.02$, $df = 1$, $p = 0.45$; Sophie-CDS: $\chi^2 = 11.99$, $df = 1$, $p < 0.001$; Emma-CDS: $\chi^2 = 14.72$, $df = 1$, $p < 0.0001$; Sophie-ADS: $\chi^2 = 12.89$, $df = 1$, $p < 0.001$; Emma-ADS: $\chi^2 = 26.39$, $df = 1$, $p < 0.0001$). Sophie's CS has three times more suffixed lemmas, Sophie's CDS a little less but more than twice as much, as Emma's CDS. It is in Emma's CS that the difference between the two types of complex words is the smallest. We see here again in the first case a strong similarity between CS and CDS (Sophie) in favour of suffixed nouns. In the other case (Emma), there is much less similarity between CS and CDS as far as suffixed noun lemmas (and tokens to a lower extent) are concerned while proportions of compounds are close. Emma's CDS shows the highest part of suffixed nouns simultaneously to the highest proportions of compounds, but still her compounds represent less than half of suffixed nouns. Hence, the preference for nominal suffixation is clear in CS and CDS data.

6. Conclusion

In sum, these data provide an insight into the morphological structure of the French lexicon before age 3;0 and seem to allow some generalization regarding the detection of derivational morphology before its productive use consisting in coining neologisms.

We have found various cues of an emerging morphological network characterising the lexicon at the end of the third year in the tight-knit morphological relations within the lexicon as well as in errors on affixed words, which suggests that the children have developed and used the possible morphological associations between words in the construction of their lexicon. The morphological families of suffixed and prefixed nouns that we were able to establish from the data along with emergence of simplex words before complex words, as well as conversion relations or pairing between nouns and verbs provide additional cues. Whether the same picture can be found in a larger sample including a greater number of children is a question for further research. Another issue is the relationship between the morphological structure of the lexicon in this early period and the later development of derivational productivity, e.g. to what extent the suffixes with the largest series remain the same and are predictive of the overgeneralisations in coinages.

Acknowledgements

Many thanks to Laura Schött for her help with CDS and ADS. We also thank two anonymous reviewers and the editors of the volume for insightful suggestions.

References

- Ambridge, B. & Lieven, E. V. M. 2015. A constructivist account of child language acquisition. In *Handbook of Language Emergence*, B. MacWhinney & W. O'Grady (eds), 478–510. Hoboken NJ: Wiley Blackwell. <https://doi.org/10.1002/9781118346136.ch22>
- Amiot, D. 2004. Préfixes ou prépositions? Le cas de sur(-), sans(-), contre(-) et les autres. *Lexique* 16: 67–83.
- Apothéloz, D. 2002. *La construction du lexique français*. Paris: Ophrys.
- Apothéloz, D. 2007. La préfixation en RE-, l'antonymie directionnelle et les phénomènes de polarité sémantique. *French Language Studies* 17: 143–158. <https://doi.org/10.1017/S0959269507002815>
- Arrivé, M., Gadet, F. & Galmiche, M. (eds) 1986. *La grammaire d'aujourd'hui*. Paris: Flammarion.
- Bauer, B. L. M. 2011. Word formation in Romance. In *The Cambridge History of the Romance Languages*, Vol. I, M. Maiden, J. C. Smith & A. Ledgeway (eds), 532–563. Cambridge: CUP.

- Bauer, L., Lieber, R. & Plag, I. 2013. The analysis and limits of conversion. In *The Oxford Reference Guide to English Morphology*, L. Bauer, R. Lieber & I. Plag, 545–567. Oxford: OUP.
<https://doi.org/10.1093/acprof:oso/9780198747062.003.0025>
- Béchade, H.-D. 1992. *Phonétique et morphologie du français contemporain*. Paris: PUF.
- Bloom, L. 1970. *Language Development: Form and Function in Emerging Grammar*. Cambridge MA: The MIT Press.
- Bonami, O. & Boyé, G. 2005. Construire le paradigme d'un adjectif. *Recherches Linguistiques de Vincennes* 34: 77–98. <https://doi.org/10.4000/rlv.1370>
- Bonami, O., Boyé, G. & Kerleroux, F. 2009. L'allomorphie radicale et la relation flexion–construction. In *Aperçus de morphologie du français*, B. Fradin, F. Kerleroux & M. Plénat (eds), 103–126. Paris: Presses Universitaires de Vincennes.
- Booij, G. 2012. *The Grammar of Words*. Oxford: OUP.
- Clark, E. V. 1985. The acquisition of Romance, with special reference to French. In *The Cross-linguistic Study of Language Acquisition*, Vol.1, D. I. Slobin (ed.), 687–782. Hillsdale NJ: Lawrence Erlbaum Associates.
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511554377>
- Clark, E. V. 1998. Lexical creativity in French-speaking children. *Cahiers de Psychologie Cognitive* 17: 513–530.
- Corbin, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.
- Dal, G. 1997. *Grammaire du suffixe -et(te)*. Paris: Didier.
- Dressler, W. U. & Merlini Barbaresi, L. 1994. *Morphopragmatics*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110877052>
- Ferret, K. & Villoing, F. 2015. French N-age instrumentals: Semantic properties of the base verb. *Morphology* 25(4): 473–496. <https://doi.org/10.1007/s11525-015-9273-0>
- Fradin, B. 2003. *Nouvelles approches en morphologie*. Paris: PUF.
<https://doi.org/10.3917/puf.fradi.2003.01>
- Fradin, B. 2009. Compounding in French. In *The Oxford Handbook of Compounding*, R. Lieber & P. Štekauer (eds), 417–435. Oxford: OUP.
- Fradin, B., Hathout, N. & Meunier, F. 2003. La suffixation en -ET et la question de la productivité. *Langue Française* 140: 56–78. <https://doi.org/10.3406/lfr.2003.1066>
- Gaeta, L. 2015. Lexeme formation in a conscious approach to the lexicon. In *Semantics of Complex Words*, L. Bauer, L. Körtvélyessy & P. Štekauer (eds), 115–141. Dordrecht: Springer.
https://doi.org/10.1007/978-3-319-14102-2_7
- Grabar, N., Tribout, D., Dal, G., Fradin, B., Hathout, N., Lignon, S., Namer, F., Plancq, C., Yvon, F. & Zweigenbaum, P. 2006. Production quantitative des suffixations par -ité et -Able dans un corpus journalistique moderne. *TALN* 2006: 167–175.
- Guilbert, L. 1975. *La créativité lexicale*. Paris: Larousse.
- Huot, H. 2001. *Morphologie*. Paris: Armand Colin.
- Kerleroux, F. 1996. *La coupure invisible. Études de syntaxe et de morphologie*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Kilani-Schoch, M. 2017. Early development of compounds in two French children's corpora. In *Nominal Compound Acquisition [Language Acquisition and Language Disorders 61]*, W. U. Dressler, F. N. Ketzet & M. Kilani-Schoch (eds), 91–118. Amsterdam: John Benjamins.
<https://doi.org/10.1075/lald.61.05kil>

- Kilani-Schoch, M. & Xanthos, A. 2013. The adjective *petit* 'small, little' in French acquisition data: An example of the relationship between pragmatics and morphosyntactic development. *Journal of Pragmatics* 56: 113–132. <https://doi.org/10.1016/j.pragma.2012.01.008>
- Koehl, A. 2012. La construction morphologique des noms désadjectivaux suffixés en français. PhD dissertation. Université de Lorraine.
- Le Grand Robert de la langue française*. 2018. <<https://www.lerobert.com>> (2 May 2021).
- Lehmann, A. & Martin-Berthet, F. 1998. *Introduction à la lexicologie*. Paris: Dunod.
- Le Robert brio*. 2012. Paris: Dictionnaires Le Robert.
- Lippeveld, M. & Oshima-Takane, Y. 2015a. Nouns to verbs to and verbs to nouns: When do children acquire class extension rules for deverbal nouns and denominal verbs? *Applied Psycholinguistics* 36(3): 559–588. <https://doi.org/10.1017/S0142716413000349>
- Lippeveld, M. & Oshima-Takane, Y. 2015b. The effect of input on children's cross-categorical use of polysemous noun-verb pairs. *Language Acquisition* 22: 209–239. <https://doi.org/10.1080/10489223.2014.943902>
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Manova, S. 2011. *Understanding Morphological Rules*. Dordrecht: Springer. <https://doi.org/10.1007/978-90-481-9547-3>
- Marascuilo, L. A. & McSweeney, M. 1967. Nonparametric post hoc comparisons for trend. *Psychological Bulletin* 67(6): 401–412. <https://doi.org/10.1037/h0020421>
- Melčuk, I. 1996. *Cours de morphologie générale*, Vol. III. Montréal: CNRS.
- Namer, F. & Villoing, F. 2008. Interpréter les noms déverbaux: Quelle relation avec la structure argumentale du verbe de base? Le cas des noms en -oir(e) du français. In *Actes du 1er Congrès Mondial de Linguistique Française*, J. Durand, B. Habert & B. Laks (eds), 1551–1569. Paris: EDP Sciences. <https://doi.org/10.1051/cmlf08226>
- Plénat, M. 2008. Le thème L de l'adjectif et du nom. In *Actes du Congrès Mondial de Linguistique Française 2008*, J. Durand, B. Habert & B. Laks (eds), 1613–1626. Paris: Institut de Linguistique Française. <https://doi.org/10.1051/cmlf08306>
- Roché, M. 2009. Un ou deux suffixations? Une ou deux suffixations? In *Aperçus de morphologie du français*, B. Fradin, F. Kerleroux & M. Plénat (eds), 143–174. Paris: Presses Universitaires de Vincennes.
- Savickienė, I. & Dressler, W. U. (eds) 2007. *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Trésor de la langue française informatisé* (TLFi). <<http://atilf.atilf.fr/tlf.htm>> (30 April 2021).
- Tribout, D. 2010. Les conversions de nom à verbe et de verbe à nom en français. PhD dissertation. Université Paris Diderot.
- Tribushinina, E., van den Bergh, H., Kilani-Schoch, M., Aksu-Koç, A., Dabašinskienė, I., Hrzica, G., Korecky-Kröll, K., Noccetti, S. & Dressler, W. U. 2013. The role of explicit contrast in adjective acquisition: A cross-linguistic longitudinal study of adjective production in spontaneous child speech and parental input. *First Language* 33(6): 594–616. <https://doi.org/10.1177/0142723713503146>
- Tribushinina, E., Voeikova, M. & Noccetti, S. (eds) 2015. *Semantics and Morphology of Early Adjectives in First Language Acquisition*. Newcastle upon Tyne: Cambridge Scholars.

- Veneziano, E. 2017. Noun and verb categories in acquisition: Evidence from fillers and inflectional morphology in French-acquiring children. In *Polycategoriality and Lexical Categories: Cross-linguistic, Cross-theoretical and Language Acquisition Approaches* [Studies in Language Companion Series 182], V. Vapnarsky & E. Veneziano (eds), 381–411. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.182.13ven>
- Villoing, F. 2012. French compounds. *Probus* 2: 29–60. <https://doi.org/10.1515/probus-2012-0003>

Emergence and early development of derivatives in Danish child language

Laila Kjærbæk and Hans Basbøll

University of Southern Denmark

This chapter is a first attempt to describe and characterize the development of derivational morphology in monolingual Danish-speaking children's early spontaneous speech. It introduces the Danish derivational system and gives an overview of the most common derivational types in Danish, emphasizing the interaction between derivation and prosody. It analyses derivatives in a corpus of parent-child-interactions and discusses factors impacting early development of derivatives. The results show that conventional derivatives appear from age 1;3–1;7. The most frequent derivatives and derivational patterns in child speech are also among the most frequent in child-directed speech. Only a very small inventory of derivational patterns has emerged up to age 3;11, and only vague traces of an emerging knowledge of productive derivation are found.

Keywords: input-output relations, frequency, prosody, conversion, mini-paradigm criterion of potential productivity

1. Introduction

Research has documented Danish-speaking children's acquisition and development of inflectional morphology (e.g. Bleses 2000; Bleses, Vach, Wehberg, Faber & Madsen 2007; Kjærbæk 2013, 2015; Kjærbæk, dePont Christensen & Basbøll 2014; Kjærbæk & Basbøll 2015, 2016, 2019), their productive application of a limited rule set to novel word stems (Paulin & Beckman 2004) and their development of early compounding (Kjærbæk & Basbøll 2017). By contrast, derivation is still a rather unexplored area in Danish-speaking children's first language acquisition, and it has, to our knowledge, never been studied systematically. However, Jespersen (1923) presents anecdotal examples of innovative derivatives in Danish-speaking children

already from the age of 1;11 where the girl Else said **fej(e)*¹ (*fej(e)* ‘sweep’ verb (v)² + *-er* → noun (n)) about a *kost* ‘broom’.

This chapter focuses on the emergence of derivational morphology and its development up to age 3;11. It addresses three central issues for the understanding of the acquisition of derivational morphology: (i) distributional properties of derivational morphology in the target language and in the language input; (ii) developmental patterns and trajectories in spontaneous speech; and (iii) mechanisms of learning from input. Thus, the chapter provides: (i) a brief introduction to the Danish derivational system as well as an overview of some of the most common derivational types in Danish; (ii) analyses of derivatives in a corpus of parent-child-interactions (up to age 3;11); and (iii) a discussion of input factors. The aim of the study is to identify the profiles of derivatives in naturalistic spontaneous speech of four typically developing monolingual Danish-speaking children from the first occurrence up to the ages of 2;5, 2;5, 2;11 and 3;11. Corresponding derivative profiles of caregivers in the same interactions will also be identified. The study furthermore aims at identifying signs of emerging knowledge of derivational morphology in the course of development. The findings on the emergence and development of derivatives are compared with previous studies on nominal compounds (Kjærbaek & Basbøll 2017) and noun plural inflection (e.g. Kjærbaek & Basbøll 2019), partly using the same data.

Our main research questions are:

1. When do derivatives start to emerge?
2. Which derivational patterns emerge first?
3. Is there a correlation between early emerging patterns and input frequency?
4. How does complexity of derivational morphology evolve?
5. To what extent do the children use derivational patterns productively?

1. * is used here to mark non-conventional word forms.

2. The infinitive form of a verb is the most important base for morphology; when the infinitive form ends in a schwa, the verb stem is formed by subtracting schwa from the infinitive form (see Basbøll 2005: 447–460). Unless otherwise specified, verb examples throughout the chapter are given in the infinitive form. If the infinitive ends in schwa, the infinitive marker *-e* is added in parentheses (*e*) as e.g. in *bad(e)* ‘bathe’. The productive base is the infinitive form of the verb; sometimes, however, the verb stem is the base. The second of two consecutive schwas is deleted in the morphology (Basbøll 2005: 309).

2. Danish derivational morphology

Danish is a North Germanic language which is morphologically rather similar to its close Scandinavian relatives Swedish and Norwegian, whereas it is much more different from them in phonological respects, including prosody. Danish is an inflectional-fusional language which may be on its path towards becoming agglutinating (see Herslund 2001, 2002). With respect to inflectional morphology Danish is richer than, for example, English, but poorer than, for example, German. Compounding is extremely productive in Danish word formation, both in terms of providing a large number of established word forms and in terms of being used to create novel formations – especially in nouns, and somewhat in adjectives and verbs (Bauer 2011; Kjærbaek & Basbøll 2017).

Derivation is productive in Danish. The derivational basis is the constituent that is neither prefix nor suffix and that can stand alone, i.e. the part of the word which remains after you remove the derivational affix (e.g. *ven* ‘friend’ N in *u-ven* (lit. ‘unfriend’) ‘foe’ N; *skue+spil* (lit. ‘sight play’) ‘play’ N in *skue+spill-er* ‘actor’ N). Cranberry morphemes (bound roots which occur in only one word) are common in conventional Danish derivatives, for example in the adjective *dej-lig* ‘lovely’ where *-lig* is a well-known and frequently occurring derivational suffix that changes the basis into an adjective, but the basis *dej* is not a known word in modern Danish; the same goes for confix morphemes (bound roots of Greek or Latin origin) in whole word families (e.g. *elektr-iker* ‘electrician’ N, *elektr-isk* ‘electric’ adjective (ADJ), *elektr-ificer(e)* ‘electrify’ V).

Derivational affixes constitute closed but not fixed paradigms and they can be used productively to form both conventional words (e.g. *u-ven*, *skue+spill-er*, *dej-lig*, see above) and innovative formations (e.g. **fej(e)-er*, see above). Both native and borrowed affixes can occur with native as well as borrowed bases (e.g. *bad* ‘bath’ N → *bad(e)-Ø* ‘bathe’ V; *akt* ‘act’ N → *akt-ion* ‘action’ N). Danish has a rich inventory of various types of *explicit* (overt, non-null) derivational affixes for most lexical categories. Many affixes are polysemous and/or polyfunctional, i.e. they attach to bases of various word classes. Prefixation (e.g. *tal(e)* ‘speak’ V → *be-tal(e)* ‘pay’ V) and suffixation (e.g. *forsk(e)* ‘do research’ V → *forsk(e)-er* ‘researcher’ N and *forsk-ning* ‘research’ N), as well as *implicit* derivation, i.e. a type of derivation with root change but no pronounced derivational affix (e.g. *giv(e)* ‘give’ V → *gave-Ø* (i:a) ‘present’ N; *ligg(e)* ‘lie’ intransitive V → *lægg(e)* (i:æ) ‘put’ transitive V). Affixes vary greatly in their type frequency; for example, agentive *-er* has produced several hundred derivatives, whereas agentive *-ator* is far more restricted.

A special type of derivation is *conversion*,³ which can, for example, nominalize a verb basis with a zero derivational suffix (e.g. *løb(e)* ‘run’ V → *løb-Ø* ‘race’ N; *spill(e)*

3. See Lippeveld & Oshima-Takane (2014, 2015) for discussions of problems regarding conversion.

‘play’ $v \rightarrow spil\text{-}\emptyset$ ‘game’ N). In Danish, however, there is not always full phonological equivalence in the derivation from verbs (see note 2); the infinitive typically ends in schwa, and the stem is formed by subtracting the schwa. The infinitive is always the basis for fully productive verb forms (both derivational and inflectional), whereas in non-productive forms either the infinitive or the stem is the basis. An isolated verb stem is imperative; but if it has *stød*-basis, it has *stød* (a contrastive laryngeal syllable rhyme prosody with a complicated grammatical distribution). The corresponding infinitive, if it ends in schwa, has no *stød*, e.g. infinitive *løb(e)* ‘run’ has long [ø:] without *stød*; imperative *løb* has [ø:] with *stød*. This completely follows Basbøll’s (non-)*stød* model (see e.g. Basbøll 2005). However, we consider such forms with *stød* as true conversion (see also Basbøll 2018). Derivational suffixes, implicit derivation, as well as conversion can change the word class of the basis.

2.1 Root changes

A large number of changes of the root can occur in connection to derivation. Since the children acquire language based on spoken language, we will here only consider phonological changes of the root (not purely orthographical ones). They can be summarized as follows:⁴ (i) *prosodic changes*: they can involve stress as well as *stød*; derivational morphemes can be either stressed or unstressed, or have both possibilities, and stress position in the derivative can be different from that of the basis (on the rules, see Basbøll 2005: 464–513); (ii) *phonemic changes*: this can involve either vowels or consonants, or both.

2.2 Derivation versus inflection

Derivation must be distinguished from inflection on the one hand, and from compounding on the other (see Section 2.3). According to the standard grammar of Diderichsen (1946: 20–21), there are four cases where it is rather difficult to draw the borderline between derivation and inflection: (i) certain infinite verb forms, viz. infinitive and participles (e.g. *studer(e)-ende* ‘studying, student’ with the present participle suffix *-ende*); (ii) the relation between *lærer* ‘(male) teacher’ and *lærer-inde* ‘(female) teacher’ can be perceived both as a kind of gender inflection and as a derivation; (iii) adverbs like *ind-e* ‘within’ where the suffix *-e* signifies static (as opposed to the zero form *ind-∅* ‘in(to)’ which is dynamic); (iv) the suffix of ordinal numbers (e.g. *syv* ‘seven’ \rightarrow *syv-ende* ‘seventh’). In addition: (v) most

4. See Basbøll, Kjærbaek & Lambertsen (2011) for a detailed presentation of stem changes in Danish noun plural morphology, all of which can occur also in derivational morphology, and Laaha et al. (2011) for a comparison between stem changes in German and Danish.

adjectives can function as adverbs in which case a *-t* is most often (but not always) added to the basis (e.g. *Han er hurtig* ‘He is fast’ ADJ; *Han løber hurtig-t* ‘He runs fast’ adverb (ADV)). This *-t*-suffix can either be regarded as a derivational suffix or the neuter inflectional suffix of the adjective (e.g. Gregersen & Kristiansen 2009; Christensen & Christensen 2019: 61); (vi) *aller-* (e.g. *aller-klog-est* ‘wisest of all’) can be interpreted as either (a) an inflectional prefix; however, inflectional prefixes are not found in Danish, i.e. inflectional affixes occur at the end of the word, i.e. after the stem (including derivational suffixes) but before a clitic; or as (b) a derivational prefix, but this interpretation violates the fundamental principle that derivation always precedes inflection.

2.3 Derivation versus compounding

Derivation must furthermore be distinguished from *compounding*. However, the high number of “mixed forms” complicates the matter: forms containing both (i) at least one derivational affix, and (ii) at least two roots; (i) points towards derivation and (ii) towards compounding. This necessitates a detailed grammatical analysis in layers (depending on the theoretical framework) (see Dressler et al. 2019; Basbøll forthcoming). Some morphemes which were lexemes in their source language (Greek and Latin) can be considered affixes (bound morphemes) since they do not occur as independent words in Modern Danish. They are restricted to either the initial (e.g. *tele-*) or final position (e.g. *-grafi*) of a combination; however, prefixes and suffixes do not combine with one another, i.e. words as for example *tele-grafi*, may be considered as compounds (neoclassical compounds) or combining forms (Olsen 2014).

2.4 Prosody of Danish derivatives

There are derivatives which are, prosodically, just like prototypical compounds, for example the noun *klog-skab* ‘wisdom’ with the derivational suffix *-skab* pronounced exactly like *skab* ‘cupboard’ in noun compounds like *klæde+skab* ‘wardrobe’ (from *klæde* ‘cloth’) (cf. also German *-schaft* and English *-ship*). The stress pattern here is primary stress on the first morpheme and secondary stress on the last morpheme. Danish is (word-)prosodically more complex than other Germanic languages (that are more complex than e.g. French that has no compound stress) since Danish has the typical Germanic stress pattern for compounds (with stress gradation), and in addition the *stød*. Since *stød* depends on the sonority shape of the syllables, and since there can be several *stød* in one word, Danish is prosodically even more complex than Swedish and Norwegian which have tonal word accents in addition

to the typical Germanic stress (and are thereby prosodically more complex than e.g. German), but lack the particular phonological restrictions of Danish (see Basbøll 2019; forthcoming). Other derivatives are prosodically just like simplex words (e.g. *dyb* ‘deep’ ADJ → *dyb-de* ‘depth’ N). Derivative suffixes like *-inde*, *-at* and *-ist* have primary stress if they are word-final (‘unit accentuation’). Derivational morphemes can be analysed as inherently stressed or unstressed, and they follow the general rules for stress gradation and *stød* (see Basbøll 2005: 464–513).

3. Derivational affixes relevant for Danish child speech (CS) and child-directed speech (CDS)

This section presents derivational patterns that are relevant for early Danish child language acquisition, i.e. the presented derivational patterns occur in a corpus of CS and CDS (see Section 5 and 6). We follow the word descriptions in (a) *Den Danske Ordbog* ‘The Danish Dictionary’; (b) *Ordbog over Det Danske Sprog* ‘Dictionary of the Danish Language’; (c) *Etymologisk ordbog* ‘Etymological Dictionary’ (Katlev 2000), in that order.

3.1 Noun derivation

Nouns show the greatest number of morphologically complex words in Danish. Rajnik (2011) specifies an inventory of approximately 100 productive derivational affixes in Danish noun formation (including a number of synonymous forms). This section presents the derivational patterns in nominal derivation departing from Rajnik’s (2011) very detailed analysis of Danish nominal derivational morphology, which, however, does not take phonology into consideration.

3.1.1 Prefixes

Very few derivational prefixes are productive in Danish noun formation, for example *sam-* (e.g. *fund* ‘find’ → *sam-fund* ‘society’) which is used to express conformity, common presence or interconnectedness, and *gen-* (e.g. *syn* ‘sight’ → *gen-syn* ‘reunion’) which can express repetition or an action whereby a previous condition occurs. The prefix *u-* (which can be pronounced with and without stress), with the meaning of expressing negation, is rather frequent, and its use has widely expanded; but not all nouns can be combined with *u-*. The basis is often a psychological condition (e.g. *ro* ‘silence’ → *u-ro* ‘commotion’; *held* ‘luck’ → *u-held* ‘bad luck’; *lykke* ‘happiness’ → *u-lykke* ‘misfortune, accident’).

3.1.2 Suffixes

A large number of derivational suffixes are productive in Danish noun formation. The suffixes *-ing/-ning* belong to the most productive derivational suffixes in Danish; *-ning* is used more often than *-ing* (Rajnik 2011: 80). The derivatives of *-ing/-ning* constitute a semantically open and little defined category of nouns derived from both verbs, adjectives and nouns (Rajnik 2011: 80ff).

The extremely productive derivational suffix *-er* (unstressed) is polyfunctional (Rajnik 2011: 63): *-er*-derivations are typically divided into two semantic groups: (a) agent (e.g. *bag(e)* 'bake' v → *bag(e)-er* 'baker'); and (b) instrument (e.g. *hold(e)* 'hold' v → *hold(e)-er* 'stand').

The suffix *-dom* (e.g. *syg-dom* 'illness'), which is cognate with English *-dom* and German *-tum*, is rather frequent in everyday Danish, but it is not productive in modern Danish (Rajnik 2011: 48f), contrary to the suffix *-else* which is still very productive (Rajnik 2011: 51); and the derivatives of *-else* constitute a rather open semantic category of nouns which are almost always derived from verbs (e.g. *hør(e)* 'hear' → *hør(e)-else* 'hearing').

The suffix *-ion* (with the orthographic variants *-(t)ion*, *-(at)ion*, *-(it)ion*) has, according to Rajnik (2011: 88), become very productive in Danish; however, it usually occurs on borrowed nouns (e.g. *port-ion* 'portion'), verbs (e.g. *diskuss-ion* 'discussion') and adjectives (e.g. *stat-ion* 'station'); when on verbs, the verb basis and the derivative often appear in slightly different forms (e.g. *diskut-er(e)* 'discuss' → *diskuss-ion* 'discussion').

Danish does not have productive diminutive suffixes, as for example German; however, historically words ending in the suffix *-ling* are diminutives (e.g. *kil-ling* 'kitten' and *kyl-ling* 'chicken'), and these forms are rather frequent in child language.

Other relevant suffixes are: *-ade* (e.g. *hyst-ade* 'hysterical woman'), *-ance/-ence* (e.g. *kompet-ence* 'competence'), *-en* (e.g. *hils(e)* 'greet' v → *hils(e)-en* 'greeting'; *rundt-en+om* 'slice of bread, round (of bread)', from *rund-t* 'round, neuter' ADJ and *om* 'around' ADV, Rajnik 2011: 56 ff.), *-eri* (e.g. *rod* 'mess' N → *rod-eri* 'messiness'), *-hed* (e.g. *svag* 'weak' ADJ → *svag-hed* 'weakness'), *-ik* (e.g. *gymnast* 'gymnast' N → *gymnast-ik* 'gymnastics'), *-itet* (e.g. *univers-itet* 'university'), *-n* with root change (e.g. *sov(e)* 'sleep' v → *søv-n* (â:ö)), *-or* (e.g. *cens-or* 'censor, external examiner'), *-tor* (e.g. *trak-tor* 'tractor'), *-sel* (e.g. *fød-sel* 'birth'), *-skab* (e.g. *sel-skab* 'party'), *-um* (e.g. *punkt-um* 'period') and *-ur/-atur* (e.g. *cens-ur* 'censorship', *temper-atur* 'temperature').

3.1.3 Implicit derivation

Implicit derivation also occurs in noun formation (e.g. *giv(e)* 'give' V → *gave-Ø* (i:a) 'present'). It occurs rather frequently, although it does not seem to be productive in modern Danish.

3.1.4 Conversion

Conversion (zero-derivation) is extremely frequent in Danish noun formation, and it tends to produce activity nouns rather than object nouns (e.g. *bid(e)* ‘bite’ v → *bid-Ø* ‘bite’).

3.2 Verb derivation

This section presents the verb derivation patterns relevant for Danish child language acquisition, i.e. all examples of derivational patterns that occur in a corpus of CS and CDS (see Section 5 and 6).

3.2.1 Prefixes

The most frequent verbal prefix in the Danish derivational system is *be-* which modifies or radically changes the meaning of the verb basis and often also the syntactic features of the verb, for example by transitivity (e.g. *stemm(e)* ‘vote’ intransitive → *be-stemm(e)* ‘decide’ transitive). Also, prefixes such as *und-* and *re-* can be added to a verb basis in order to modify the stem (e.g. *skyld(e)* ‘owe’ → *und-skyld(e)* ‘excuse’; *citer(e)* ‘cite’ → *re-citer(e)* ‘recite’).

3.2.2 Suffixes

One of the most productive verbal suffixes is the denominal *-er(e)* (with stress; with the variations *-r(e)*, *-ter(e)*) (e.g. *analyse* ‘analysis’ → *analys-er(e)* ‘analyse’; *blad* ‘leaf’ → *bladr(e)* ‘turn a leaf’; *hånd* ‘hand’ → *hånd-ter(e)* ‘handle’) which is used to form verbs that denote production or execution of something particular. The suffix *-n(e)* changes one verb basis into another and denotes that something changes (e.g. *våg(e)* ‘watch over’ → *våg-n(e)* ‘wake’, i.e. an inchoative function). Other productive suffixes are denominal *-s* (e.g. *lykke* ‘happiness’ → *lykke-s* ‘succeed’) and deadjectival *-s(e)* (e.g. *ren* ‘clean’ → *ren-s(e)* ‘cleanse’), whereas a less productive suffix is denominal *-sk(e)* (e.g. *hu* ‘mind’ → *hu-sk(e)* ‘remember’).

3.2.3 Implicit derivation

Implicit derivation in verbs is somewhat frequent; however, it does not seem to be productive in modern Danish. Both semantic and/or syntactic change of the verb (e.g. *ligg(e)* ‘lie’ intransitive → *lægg(e)* ‘put’ transitive) and denominal implicit derivation (e.g. *behov* ‘requirement’ → *behøv(e)* ‘require’) are frequent. Deadjectival implicit derivation (e.g. *tom* ‘empty’ → *tømm(e)* ‘empty’) also exists, but it is less frequent.

3.2.4 Conversion

The most frequent verb derivational patterns are cases of denominal conversion (e.g. *bad* ‘bath’ → *bad(e)* ‘bathe’; *duft* ‘scent’ → *duft(e)* ‘smell’); whereas verb conversion from adjectives (e.g. *brun* ‘brown’ → *brun(e)* ‘brown’; *varm* ‘warm’ → *varm(e)* ‘warm’), from onomatopoeic⁵ words (e.g. *pip* ‘chirp’ → *pipp(e)* ‘chirp’; *rap* ‘quack’ → *rapp(e)* ‘quack’) and from adverbs (*op* ‘up’ → *opp(e)* (*sig*) ‘make an effort’) exist but are less frequent.

3.3 Adjective derivation

This section presents adjective derivational patterns relevant for Danish child language acquisition, i.e. all examples of derivational patterns occur in a corpus of CS and CDS (see Section 5 and 6).

3.3.1 Prefixes

The most frequent and highly productive adjective derivational prefix is *u-* which expresses negation (e.g. *lækker* ‘delicious’ → *u-lækker* ‘disgusting’; *ro-lig* ‘quiet’ → *u-ro-lig* ‘unquiet’; *hygge-lig* ‘cozy’ → *u-hygge-lig* ‘scary’). In cases like *u-ro-lig* and *u-hygge-lig* the formation is formally ambiguous since also the nouns *u-ro* ‘restlessness’ and *u-hygge* ‘uneasy feeling’ exist.

The infrequent prefix *aller-* adds the meaning ‘of all’ (e.g. *aller-bedst* ‘best of all’); *multi-* adds the meaning of ‘many’ to the basis (e.g. *multi-kultur-el* ‘multi-cultural’), whereas *ge-* adds the meaning of ‘connected with’ (e.g. *ge-vald-ig* ‘huge’, cf. *væld-ig* ‘huge’).

3.3.2 Suffixes

The derivational suffix *-ig/-lig*, which is one of the most productive and frequent adjective suffixes in Danish, is predominantly denominal (e.g. *tørst* ‘thirst’ → *tørst-ig* ‘thirsty’; *ven* ‘friend’ → *ven-lig* ‘friendly’); furthermore, it occurs sometimes with root change (e.g. *magt* ‘power’ → *mægt-ig* (a:æ) ‘powerful’), but it can also change the basis from one adjective to another adjective (e.g. *brun* ‘brown’ → *brun-lig* ‘brownish’). Sometimes it is deverbal (e.g. *ked(e)* ‘bore’ → *ked(e)-lig* ‘boring’), and occasionally it forms an adjective from a numeral word (e.g. *en* ‘one’ → *en-ig* ‘in agreement’).

5. We follow a Danish tradition where onomatopoeic words (*lydord*, lit. ‘sound words’, in Danish) are considered to be a word class by itself – different from other word classes as for example interjections (Hansen & Heltoft 2011: 1116).

The suffix *-agtig* is very productive in modern Danish. It expresses similarity to, or association with, the basis, and it is often used to coin an adjective from another adjective basis (e.g. *nøje* ‘precisely’ → *nøj-agtig* ‘precise’). The suffixes *-ant*, *-el*, *-et* and *-som* are denominal (e.g. *interesse* ‘interest’ → *interess-ant* ‘interesting’; *kultur* ‘culture’ → *kultur-el* ‘cultural’; *støv* ‘dust’ → *støv-et* ‘dusty’; *vold* ‘violence’ → *vold-som* ‘violent’), whereas the suffixes *-bar* and *-en* are deverbal (e.g. *kost(e)* ‘cost’ → *kost-bar* ‘valuable’; *sult(e)* ‘starve’ → *sult(e)-en* ‘hungry’; *voks(e)* ‘grow’ → *voks(e)-en* ‘adult’).

The suffix *-sk* (with the variant *-isk*) is denominal and typically denotes an association or affiliation with a particular country, city, institution or similar; it is either derived from a common noun (e.g. *hysteri* ‘hysteria’ → *hysteri-sk* ‘hysterical’; *fakta* ‘fact(s)’ → *fakt-isk* ‘actually’) or from a proper name (e.g. *Danmark* ‘Denmark’ (*Dan* ‘mythical first king of Denmark’) → *dan-sk* ‘Danish’; *Tyrki-et* ‘Turkey’ → *tyrk-isk* ‘Turkish’).

The suffixes *-ende* (originally the present participle) and *-t* (originally the past participle) are deverbal (e.g. *irriter(e)* ‘annoy’ → *irriter(e)-ende* ‘annoying’; *lev(e)* ‘live’ → *lev(e)-ende* ‘alive’). The suffix *-s* is denominal and originally a genitive clitic (e.g. *stakkel* ‘poor creature’ → *stakkel-s* ‘poor’). The suffix (or pseudo-suffix) *-p* is deverbal and is combined with stem change (e.g. *skær(e)* ‘cut’ → *skar-p* (æ:a) ‘sharp’).

3.3.3 Conversion

A rather frequent adjective derivational pattern is denominal conversion (e.g. *lys* ‘light/candle’ → *lys* ‘light’). Implicit derivation does not occur in adjectives.

3.4 Operationalization

The first step in word analysis in Danish is to cut off inflectional suffixes from the end of the word. The remainder is the stem. Stem formation is recursive, i.e. a stem can occur within a stem. A stem consisting of only one morpheme is a root. The last root of the stem is the semantic head, whereas the last morpheme (either a root or a derivational suffix) of the stem determines the word class (Kjørbaek & Basbøll 2017), i.e. derivational suffixes determine the word class, and thus also inflection. In the present study, in order to achieve as rich a data set as possible, we consider all words containing a derivational affix as derivatives, i.e. derivatives contain at least one prefix (e.g. *be-*, *mis-*, *u-*) or one derivational suffix (e.g. *-agtig*, *-lig*, *-skab*). Additionally, the results of implicit derivation and conversion are considered derivatives.

This investigation follows the following principles: (i) Cases like *studer(e)-ende* ‘studying’ which are originally verbs with the inflectional suffix *-ende* marking present participle, are treated as verbs if they follow an auxiliary (as in: *Hun kommer kravlende over gulvet* ‘She is crawling across the floor’), but as adjectives with the derivational suffix *-ende* changing the basis from verb to adjective if they are part

of a noun phrase (as in: *Har du brug for en hjælpende hånd?* ‘Do you need a helping hand?’) or function as a predicative (as in: *Du er irriterende* ‘you are annoying’); (ii) the suffix of static adverbs (e.g. *ind-e* ‘within’) is treated as a derivational suffix; (iii) the suffix of ordinal numbers (e.g. *syv-ende* ‘seventh’) is treated as a derivational suffix; (iv) adjectives functioning as adverbs (e.g. *hurtig-t* ‘fast’) are treated as adjectives.

In order to investigate when derivational processes start to become productive, we build on the mini-paradigm criterion of potential productivity which has been introduced for the acquisition of inflection (e.g. Kilani-Schoch & Dressler 2005; Dressler, Lettner & Korecky-Kröll 2010) and compounding (e.g. Korecky-Kröll, Sommer-Lolei & Dressler 2017). When adapted to Danish derivational morphology in CS, the mini-paradigm criterion of potential productivity is fulfilled when (a) a morphological process (affixation or conversion) is applied to at least three different bases, and when (b) each base occurs in at least three different types (uninflected or inflected words, compounds or other derivations) (see Introduction chapter and Sommer-Lolei et al. this volume).

4. Hypotheses

On the basis of our knowledge about the Danish derivational system and earlier studies on the early emergence and development of derivational morphology, we propose the following hypotheses:

1. The first derivatives to appear in CS are expected to be rote learned and treated as whole lexical items (see Clark 2014).
2. Later the child will start to analyse the derivatives and recognize the word constituents and the combinatorial processes in derivational word formation in order to link the roots and affixes to a specific meaning – hence starting to form derivatives productively.
3. Compounding is expected to appear before derivation due to higher morphosemantic transparency (see Clark & Hecht 1982; Clark 1993).
4. Innovative compounds (neologisms) are expected to appear before innovative derivatives due to higher morphosemantic transparency (see Clark & Hecht 1982; Clark 1993).
5. We expect frequency in CDS to affect acquisition, hence more frequent forms are expected to appear before less frequent ones. Derivatives with a high token frequency and derivational patterns with a high type frequency⁶ in the input are expected to be acquired before less frequent derivatives and derivational patterns (see Clark 1993, 2014; Bybee 1995; Ambridge et al. 2015).

6. In this chapter type frequency equals lemma frequency.

5. Data basis

The present study is based on empirical data from a corpus of longitudinal, naturalistic spontaneous CS and CDS from four Danish-speaking children and their families. The children are the twin girls Ingrid and Sara⁷ who are between the ages of 0;10–2;5, from the Odense Twin Corpus (Basbøll, Bleses, Cadierno, Jensen, Ladegaard, Madsen, Millar, Sinha & Thomsen 2002); and the girl Anne who is within the age range of 1;1–2;11 and the boy Jens between the ages of 1;0–3;11, both from the Plunkett Corpus (Plunkett 1985, 1986). All children come from middle-high SES families and are recorded approximately for about an hour once (or in the Plunkett Corpus two or three times) every month in a dining and/or playing situation. CDS includes only the speech of the children's mother and father (see Table 1 for an overview of the corpus). The data are transcribed in CHILDES (MacWhinney 2000a; b) in standard orthography and coded morphologically – and with standard phonetic transcriptions – in OLAM.⁸ To achieve as rich a data set as possible, we regard all words in the data set with at least one derivational affix as derivatives, i.e. also mixed forms (see Section 3.4), implicit derivation and conversion; although there are other possible analyses.

Table 1. CS and CDS data

	Age span	Recordings	Word tokens	
			CS	CDS
Ingrid	0;10–2;5	16	3,825	79,768
Sara	0;10–2;5	16	4,311	79,768
Anne	1;1–2;11	41	8,539	20,024
Jens	1;0–3;11	48	15,131	22,946

This corpus of CS and CDS, which spans over a number of years, can provide us with valuable knowledge about the kinds of linguistic features that make up the child's input and output at different stages of development.

7. The twin girls are recorded in the same sessions so the CDS data for them are the same throughout.

8. OLAM is a computational coding and analysis system for Danish (Madsen, Basbøll & Lambertsen 2002; Kjærbæk 2013).

6. Distribution of derivatives in CS and CDS

This section presents the distribution of derivatives in the corpus of CS and CDS (see Section 5). Table 2 gives an overview of the total number of words and derivatives in CS and CDS.

Table 2. Total number of words and derivatives (tokens and types) in the corpus of CS and CDS

	Tokens	Types
CS		
All words	31,806	1,137
Derivatives	1,312	256
<i>% derivatives among all words</i>	4.1	22.5
CDS		
All words	122,738	2,859
Derivatives	7,752	797
<i>% derivatives among all words</i>	6.3	27.9

In CS out of all words 4.1% tokens and 22.5% types are derivatives, whereas in CDS 6.3% tokens and 27.9% types are derivatives (Table 2), so the proportion of derivatives is smaller in CS compared to CDS (deviation of 2.2% tokens and 5.4% types).

Table 3 provides an overview of the distribution of prefixed and suffixed derivatives in CS and CDS.

Table 3. Number of prefixed and suffixed derivatives (tokens, types) in the corpus of CS and CDS

	Prefix		Suffix	
	Tokens	Types	Tokens	Types
CS	11	8	1,305	197
<i>% of all derivatives</i> [†]	0.8	3.1	99.5	77.0
CDS	198	58	7,646	775
<i>% of all derivatives</i> [†]	2.6	7.3	98.6	97.2

[†] Please note that a derived word can have both a prefix and a suffix.

Table 3 illustrates that suffixes are much more frequent than prefixes: Out of all derivatives in CS 0.8% tokens and 3.1% types are prefixed; whereas 99.5% tokens and 77.0% types are suffixed. In CDS, however, 2.6% tokens and 7.3% types are prefixed; whereas 98.6% tokens and 97.2% types are suffixed. Thus, prefixation is less frequent in CS than in CDS, particularly in types.

Table 4 shows the distribution of nouns, verbs and adjectives, including the proportion of derived words within the three word classes, in CS and CDS for each of the four children Ingrid, Sara, Anne and Jens.

Table 4. Distribution of nouns, verbs and adjectives incl. the proportion of derived words within each word class in the corpus of CS and CDS

	Ingrid		Sara	Anne		Jens	
	CS tok./ typ.	CDS tok./ typ.	CS tok./ typ.	CS tok./ typ.	CDS tok./ typ.	CS tok./ typ.	CDS tok./ typ.
Nouns	692/148	6,666/1,063	806/152	1,009/207	1,753/471	2,243/298	1,882/459
N DERIV	75/13	986/287	72/26	113/36	234/90	218/39	224/86
% N DERIV	10.8/8.8	14.8/27.0	8.9/17.1	11.2/17.4	13.3/15.8	9.7/13.1	11.9/18.7
Verbs	532/89	19,420/464	661/97	1,765/109	5,285/233	3,257/161	5,864/216
V DERIV	25/12	1,155/153	21/9	94/23	406/60	161/43	436/88
% V DERIV	4.7/13.5	5.9/33.0	3.2/9.3	5.3/21.1	7.7/25.8	4.9/26.7	7.4/40.7
Adjectives	238/38	4,521/274	246/45	283/50	1,116/137	583/75	1,003/128
ADJ DERIV	31/10	1,029/142	29/11	44/12	163/54	93/17	158/51
% ADJ DERIV	13.0/26.3	23.3/51.8	11.8/24.4	15.5/24.0	14.6/16.0	16.0/22.7	15.8/39.8

Verbs have the highest token frequency and nouns the highest type frequency in both CS and CDS; the only exceptions are CS from Ingrid and Sara where nouns have a higher token frequency than verbs. When we compare the proportion of derivatives between the word classes, we see that the token frequency for both CS and CDS for all children is highest for adjectives, then nouns and lowest for verbs. The same applies to type frequency for Ingrid and Sara, whereas for Anne it is highest for adjectives, then verbs and lowest for nouns. For Jens, type frequency is highest for verbs, then adjectives and lowest for nouns. The proportion of derived words are lower in CS than in CDS within all word classes; the only exception is Anne with 17.4% derivatives out of all nouns in CS compared to only 15.8% in CDS (Table 4).

The age of the four children when their first word, noun, verb and adjective occurred in the corpus of CS is shown in Table 5.

Table 5. First word, first noun, first verb and first adjective in the corpus of CS

	†First word	†First N	First v	First ADJ
Ingrid	0;10 [‡]	1;2	1;2	1;3
Sara	0;10 [‡]	1;2	1;3	1;4
Anne	1;1 [‡]	1;1 [‡]	1;1 [‡]	1;3
Jens	1;0 [‡]	1;0 [‡]	1;0 [‡]	1;1

† According to Kjærbaek (2013).

‡ Occurrence in the first session.

All four children produced their first word around their first birthday (Ingrid, Sara and Anne: *mmm* ‘yum-yum’ interjection (INTERJ); Jens: *det* ‘that’ pronoun (PRON)). Nouns are among the earliest words to appear (Ingrid *mælk* ‘milk’; Sara: *mam* child form of *mad* ‘food’; Anne: *mad* ‘food’; Jens: *bamse* ‘teddy bear’), then come verbs (Ingrid and Anne: *se* ‘look’; Sara: *ae* ‘pat’; Jens: *være* ‘be’) and adjectives (Ingrid: *dygtig* ‘talented’; Sara and Anne: *mere* ‘more’; Jens: *stor* ‘big’) (Table 5).

Table 6. First present tense form, first noun plural form, first past tense form, first noun definite form in the corpus of CS

	First v present tense	†First n plural	First v past tense	First n definite
Ingrid	1;3	1;4	1;4	2;0
Sara	1;4	1;4	1;9	1;10
Anne	1;2	1;6	1;8	1;8
Jens	1;2	1;10	1;5	1;5

† According to Kjærbæk (2013).

During the first months, the children mainly produce simple words, but gradually more complex words appear (such as inflected nouns and verbs, compounds and derivatives; see Table 6 and Table 7).

Table 7. First conventional compound (CPD), first innovative CPD, first conventional derivative and first innovative derivative in the corpus of CS

	†First conventional CPD	†First innovative CPD	First conventional derivative	First innovative derivative
Ingrid	1;7	2;4	1;3	–
Sara	1;7	2;4	1;7	–
Anne	1;7	2;2	1;7	–
Jens	1;10	3;8	1;5	–

† According to Kjærbæk and Basbøll (2017).

The first conventional compounds in CS appear in the ages of 1;7–1;10, whereas the first innovative compounds appear several months later in the ages of 2;2–3;8 (Kjærbæk & Basbøll 2017). The first conventional derivatives in CS appear in the ages of 1;3–1;7 (Ingrid: *dygt-ig* ‘clever’ ADJ; Sara: *kyl-ling* ‘chicken’ N (diminutive); Anne: *tegn-Ø* ‘draw!’ V (denominal conversion from *tegn* ‘sign’); Jens: *hul-Ø* ‘hole’ N (implicit derivation from *hul* ‘hollow’ ADJ (å:u)).⁹ No innovative derivatives appear in the corpus of CS.

9. Where *hul* N has a short vowel, *hul* ADJ has a long vowel, hence the difference in vowel quality as well (etymologically they are related, also with *hule* N ‘cave’, pronounced with long /u:/).

6.1 Noun derivation

Noun derivatives in CS and CDS are presented in Table 8 divided into prefixes and suffixes.

Table 8. Distribution of noun derivational patterns in CS and CDS (tokens/types)

	Ingrid		Sara	Anne		Jens	
	CS	CDS	CS	CS	CDS	CS	CDS
<i>tele-</i>	–	17/1	1/1	1/1	1/1	2/1	6/2
<i>u-</i>	–	–	–	–	–	1/1	1/1
<i>-ade</i>	–	3/1	–	–	–	–	–
<i>-ance</i>	–	2/2	1/1	–	–	–	–
<i>-ant</i>	1/1	4/2	–	–	–	–	2/1
<i>-at</i>	–	1/1	1/1	–	–	–	–
<i>-atur</i>	–	1/1	–	–	–	–	–
<i>-d</i>	–	5/1	–	–	–	–	–
<i>-de</i>	–	2/2	–	–	–	–	–
<i>-dje</i>	–	2/1	–	–	–	–	–
<i>-dom</i>	–	1/1	–	–	–	–	–
<i>-e</i>	2/1	25/10	–	10/3	22/9	1/1	12/6
<i>-el</i>	–	3/1	–	–	–	1/1	–
<i>-else</i>	–	18/11	–	–	2/2	–	8/4
<i>-ende</i>	–	2/2	–	–	–	–	–
<i>-ent</i>	–	1/1	–	–	–	–	–
<i>-er</i>	1/1	47/35	1/1	–	8/7	25/5	24/9
<i>-eri</i>	–	10/2	–	–	5/2	–	5/2
<i>-ers</i>	–	2/1	–	–	–	–	–
<i>-esse</i>	–	3/1	–	–	–	–	4/1
<i>-est</i>	–	1/1	–	–	–	–	–
<i>-gte</i>	–	1/1	–	–	–	–	–
<i>-hed</i>	–	7/5	–	–	2/2	–	1/1
<i>-ik</i>	–	7/3	1/1	1/1	3/1	–	1/1
<i>-inde</i>	–	4/2	–	–	–	–	–
<i>-ing/-ning</i>	–	35/29	1/1	2/1	10/6	1/1	6/4
<i>-ion/- (at)ion/- (t)ion</i>	–	5/4	–	–	2/2	2/1	–
<i>-itet</i>	–	7/2	–	–	–	–	–
<i>-l</i>	–	7/2	4/1	–	–	–	–
<i>-le</i>	–	1/1	–	1/1	2/1	–	–

Table 8. (continued)

	Ingrid		Sara	Anne		Jens	
	CS	CDS	CS	CS	CDS	CS	CDS
<i>-ling</i>	12/2	54/5	5/3	2/1	4/2	3/1	6/2
<i>-ment</i>	–	1/1	–	–	–	–	–
<i>-n</i>	–	7/3	–	–	1/1	–	–
<i>-nt</i>	–	1/1	–	–	–	–	–
<i>-o</i>	–	6/1	–	–	–	–	–
<i>-or</i>	–	2/1	–	–	–	28/1	–
<i>-ri</i>	–	7/2	–	–	1/1	–	1/1
<i>-s</i>	–	5/3	–	–	2/1	–	4/2
<i>-sel</i>	–	8/4	–	–	1/1	1/1	3/2
<i>-ser</i>	–	–	–	–	1/1	–	–
<i>-sk</i>	–	4/2	–	–	–	–	–
<i>-ske</i>	–	7/2	–	–	2/1	–	–
<i>-st</i>	–	3/1	1/1	1/1	3/1	–	2/2
<i>-t/-et</i>	–	19/5	1/1	–	3/2	1/1	3/3
<i>-tor</i>	–	2/2	–	–	–	–	6/1
<i>-um</i>	–	1/1	–	1/1	–	–	2/1
<i>-ur</i>	–	2/2	–	–	–	–	–
<i>Implicit derivation</i>	22/4	140/7	17/5	36/3	36/5	41/1	91/8
<i>Conversion</i>	37/4	494/117	38/8	50/19	132/44	95/19	96/36

Implicit derivation (e.g. *fjern+syn-Ø* ‘television’) and conversion (e.g. *sut-Ø* ‘dummy’) are the most frequent noun derivation patterns in CS. Only two noun prefixes occur in CS (*tele-*, *u-*) and 17 different derivational suffixes. Most of the affixes only have one occurrence; the derivational suffixes *-e* (e.g. *ung-e* ‘cub’), *-er* (e.g. *brus(e)-er* ‘shower’) and *-ling* (e.g. *kyl-ling* ‘chicken’) are the most frequently used. None of the children met the mini-paradigm criterion of potential productivity for any of the noun affixes; only partly by Anne for *-e* (at 2;4, used with three different bases) and Jens for *-er* (at 2;7, used with three different bases).

6.2 Verb derivation

Verb derivatives in CS and CDS are presented in Table 9 divided into prefixes and suffixes.

Table 9. Distribution of verb derivational patterns in CS and CDS (tokens/types)

	Ingrid		Sara	Anne		Jens	
	CS	CDS	CS	CS	CDS	CS	CDS
<i>be-</i>	–	69/20	2/2	1/1	19/9	2/2	17/11
<i>dis-</i>	–	3/2	–	–	–	–	1/1
<i>re-</i>	–	2/2	–	–	–	–	–
<i>und-</i>	–	7/1	–	–	1/1	–	–
<i>-er</i>	–	10/7	–	–	–	–	1/1
<i>-ficer</i>	–	1/1	–	–	–	–	–
<i>-l</i>	–	6/1	–	–	2/1	1/1	8/1
<i>-r</i>	–	6/4	–	–	–	–	3/1
<i>-s</i>	–	192/3	–	–	22/2	9/1	48/1
<i>-sk</i>	–	55/1	–	–	26/1	3/1	42/1
<i>-t</i>	–	4/1	–	–	9/1	–	1/1
<i>-ter</i>	–	1/1	–	–	–	–	–
<i>Implicit derivation</i>	3/2	297/19	3/3	5/2	81/10	30/4	84/6
<i>Conversion</i>	22/10	869/106	17/5	87/19	306/44	134/36	333/43

Conversion is by far the most frequent verb derivational pattern followed by implicit derivation in both CS and CDS. Only one prefix (*be-*) and three verb derivational suffixes (*-l*, *-s*, *-sk*) appear in CS. None of the children met the mini-paradigm criterion of potential productivity for any of the verb affixes.

6.3 Adjective derivation

Adjective derivatives in CS and CDS are presented in Table 10 divided into prefixes and suffixes.

The most frequent adjective derivational patterns are conversion (e.g. *lys-Ø* 'light') and addition of the derivational suffixes *-ig* (e.g. *dygt-ig* 'talented', *færd-ig* 'finished') and *-lig* (e.g. *dej-lig* 'lovely', *ro-lig* 'calm'). The only adjective prefixes which occur in CS is Anne producing the prefix *u-* (*u-syn-lig* 'invisible'), whereas ten derivational suffixes occur in CS out of which *-ig* and *-lig* are the most frequent. None of the children met the mini-paradigm criterion of potential productivity for any of the adjective affixes; only partly it was met by Ingrid, Sara and Jens for *-lig* (at 2;3, 2;3 and 2;6 respectively, used with three different bases).

Table 10. Distribution of adjective derivational patterns in CS and CDS (tokens/types)

	Ingrid		Sara	Anne		Jens	
	CS	CDS	CS	CS	CDS	CS	CDS
<i>aller-</i>	–	–	–	–	–	–	1/1
<i>ge-</i>	–	1/1	–	–	–	–	–
<i>multi-</i>	–	1/1	–	–	–	–	–
<i>u-</i>	–	31/9	–	1/1	5/2	–	2/1
<i>-e</i>	–	9/3	–	–	11/3	2/1	1/1
<i>-el</i>	–	1/1	–	–	2/1	–	–
<i>-en</i>	–	6/3	–	–	1/1	10/2	1/1
<i>-ende</i>	1/1	18/5	–	–	3/3	–	3/2
<i>-ig</i>	9/2	292/19	2/2	4/3	25/10	5/2	42/6
<i>-lig</i>	5/3	244/40	5/3	2/2	37/15	15/3	25/14
<i>-n</i>	–	10/2	–	–	–	–	–
<i>-p</i>	–	3/1	–	–	–	2/1	–
<i>-s</i>	–	3/2	–	–	1/1	2/1	1/1
<i>-sk</i>	–	15/11	1/1	–	5/4	–	5/3
<i>-st</i>	–	–	–	1/1	–	–	1/1
<i>-ste</i>	–	–	–	–	2/1	–	–
<i>-t/et</i>	–	66/20	–	1/1	13/4	3/2	13/8
<i>Conversion</i>	16/4	275/21	21/5	37/6	62/10	60/6	55/7

7. Discussion

Derivation has been a rather unexplored area in Danish-speaking children's first language acquisition, and this is, to our knowledge, the first systematic study of the topic. This endeavor is rather challenging given the large number of affix categories in Danish and the wide range of meanings that derivational affixes represent. Inflectional processes are almost exclusively semantically regular, whereas derivational processes frequently express a range of polysemous interpretations (Booij 2006). For example *-er* may denote an agent (*en tal-er* 'a speaker' N), an instrument (*proptraekk-er* 'cork-screw' N), experiencer (*lytt-er* 'listener' N), stimulus (*pleas-er* 'pleaser' N), denominational N (*dansk-er* 'Dane' N) and measure (*femm-er* 'fiver' N) (see Lieber 2004). Thus, not only is the child faced with the need to acquire a vast array of many-to-many form-function combinations in the same syntactic context, but a further prerequisite for the extraction of word-formation rules involves protracted exposure to a wealth of complex words in order that analogies can be drawn (Laws 2019).

The four children in the corpus produce their first word around their first birthday. The first words are typically simple uninflected words, but only a few months later the first inflected words start to emerge (noun plurals, age 1;4–1;10) and the first complex words, viz. conventional derivatives (1;3–1;7) and conventional compounds (1;7–1;10). This contradicts studies in other languages as well as Hypothesis 3 which predicts compounding to appear before derivation due to higher morphosemantic transparency (see Clark & Hecht 1982; Clark 1993). One explanation may be that the first derivatives to appear are acquired as isolated rote-learned words, i.e. they are still unanalysed and treated as whole lexical items (see Clark 2014), together with the fact that the first derivatives to appear (e.g. *dygt-ig* ‘talented’) have a high token frequency in CDS.

As expected, the proportion of derived words is generally lower in CS than in CDS within all word classes. A comparison of the percentages of derivatives between word classes shows that the token frequency for both CS and CDS are generally highest for adjectives, followed by nouns and then verbs. For type frequency, however, it varies among the four children (Table 4).

The most frequent derivational pattern in CS is conversion (e.g. *sut-Ø* ‘dummy’ N, *slut-Ø* ‘over’ ADJ, *tegn(e)* ‘draw’ V, cf. *sutt(e)* ‘suck’ V, *slutt(e)* ‘end’ V, *tegn* ‘sign’ N), and this is also the most frequent derivational pattern in CDS. Implicit derivation, i.e. zero-derivation with root change (such as ablaut or umlaut, which is fossilized in Danish, see Laaha et al. 2011; e.g. *syng(e)* ‘sing’ V → *sang-Ø* (y:a) ‘song’ N), is also rather frequent in both CS and CDS. This supports the hypothesis of effect of input frequency (Hypothesis 5). However, with regard to derivational affixes there is a rather unclear relation between input frequency and emergence of derivatives and derivational patterns.

Prefixation is less frequent in CS than in CDS, particularly in types. This could be explained by the recency effect which makes in early phases suffixes easier to notice/identify and thus to acquire than prefixes (see Slobin 1973; Peters 1997). Only two derivational prefixes occur in CS, namely *be-* as in *be-tyd(e)* ‘mean (V)’ and *u-* as in *u-synlig* ‘invisible’ (cf. *tyd(e)* ‘interpret’, *syn-lig* ‘visible’ from *syn* ‘sight’), whereas the children produce a larger number of derivational suffixes. Among the most frequent patterns of suffixation we find *-er* (e.g. *brus(e)-er* ‘shower’, cf. *brus(e)* ‘surge’) and *-lig* (e.g. *far-lig* ‘dangerous’, *dår-lig* ‘bad’, cf. *fare* ‘danger’, *dåre* ‘idiot’). Essentially, the patterns in CS reflect those present in CDS, although there is not a 1:1 relation between input frequency and acquisition of derivational morphology, and the frequency of derivatives is much higher in CDS compared to CS.

The first innovative compounds emerge at the ages of 2;2–3;8, whereas no innovative derivatives occur in the corpus. This is consistent with studies in other languages indicating that productive use of derivational processes appears later in language development than compounding, due to its greater complexity and

lower morphosemantic transparency (Berko 1958; Clark & Hecht 1982; Derwing & Baker 1986). This confirms Hypothesis 4 expecting innovative compounds to appear before innovative derivatives due to higher morphosemantic transparency (see Clark & Hecht 1982; Clark 1993).

None of the four children met the mini-paradigm criterion of potential productivity for any of the derivational affixes, which indicates that the first derivatives to appear in Danish child language are rote learned and treated as whole lexical items (see Clark 2014) confirming Hypothesis 1. However, the mini-paradigm criterion of potential productivity was partly met for the noun derivational suffix *-e* by Anne (2;4) and *-er* by Jens (2;7), as well as for the adjective derivational suffix *-lig* by Ingrid, Sara and Jens (2;3, 2;3 and 2;6 respectively), and this may be an early indication that the children start to analyse the derivatives and recognize the word constituents and the combinatorial process in derivational word formation in order to link the roots and affixes to a specific meaning – hence starting to form derivatives productively (Hypothesis 2).

8. Conclusion

Conventional derivatives appear in Danish-speaking children already from the age of 1;3–1;7 years. For all four children in the study this is either the same age or a few months earlier than the appearance of the first conventional compound. The derivatives and derivational patterns which appear most frequently in CS, are also among the most frequent in CDS, though there is not a 1:1 connection between input frequency and emergence of derivational affixes. We only find vague traces of an emerging knowledge of productive derivation, since no innovative derivatives appear in CS, and the mini-paradigm criterion of potential productivity is only partly met by each child for one or two suffixes; i.e. young Danish-speaking children seem to acquire derivatives as unanalysed lexical items and have not yet started to use derivation productively. Based on these observations we conclude that only a very small inventory of derivational patterns has emerged up to the age of 3;11. The acquisition of derivation could thus be assumed to increase in preschool age and even during school age up to adolescence.

Acknowledgement

We would like to thank Claus Lambertsen for his assistance with the OLAM-system. We are indebted to two anonymous reviewers and the editors for many useful remarks that have been of considerable benefit to the manuscript.

References

- Ambridge, B., Kidd, E., Rowland, C. F. & Theakston, A. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2): 239–253.
<https://doi.org/10.1017/S030500091400049X>
- Basbøll, H. 2005. *The Phonology of Danish*. Oxford: OUP.
- Basbøll, H. 2018. Danish imperative formation: A problem for the phonology/morphology interface. *Scandinavian Philology* 16(2): 219–244.
- Basbøll, H. 2019. Danish word-structure, the Non-Stød Model, and some comments about Norwegian tonal word accents. In *Fonologi, sociolingvistik og vitenskapsteori. Festschrift til Gjert Kristoffersen* (Phonology, sociolinguistics and theory of science. Festschrift to Gjert Kristoffersen), J. K. Hognestad, T. Kinn & T. Lohndal (eds), 47–64. Oslo: Novus.
- Basbøll, H., Bleses, D., Cadierno, T., Jensen, A., Ladegaard, H. J., Madsen, T. & Thomsen, P. 2002. The Odense language acquisition project. *Child Language Bulletin* 22: 11–12.
- Basbøll, H. Forthcoming. A general model of word-structure and its application to Danish, emphasizing the interaction between morphology and prosody. In *Inflection vs. Derivation*, W. U. Dressler (ed.). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Basbøll, H., Kjærbæk, L. & Lambertsen, C. 2011. The Danish noun plural landscape. *Acta Linguistica Hafniensia* 43(2): 81–105. <https://doi.org/10.1080/03740463.2011.689688>
- Bauer, L. 2011. IE, Germanic: Danish. In *The Oxford Handbook of Compounding*, 2nd edn., R. Lieber & P. Štekauer (eds), 400–416. Oxford: OUP.
- Berko, J. 1958. The child's learning of English morphology. *Word* 14(2–3): 150–177.
<https://doi.org/10.1080/00437956.1958.11659661>
- Bleses, D. 2000. Transparens og produktivitet i danske børns tilegnelse af verbers præteritum (Transparency and productivity in Danish children's acquisition of past tense verbs). *Ny-danske Studier og Almen Kommunikationsteori (NyS)* 26–27: 165–196.
<https://doi.org/10.7146/nys.v26i26-27.13414>
- Bleses, D., Vach, W., Wehberg, S., Faber, K. & Madsen, T. O. 2007. *Tidlig kommunikativ udvikling* (Early communicative development). Odense: Syddansk Universitetsforlag.
- Booij, G. 2006. Inflection and Derivation. In *Encyclopedia of Language & Linguistics*, 2nd edn., Vol. 5, K. Brown (ed), 654–661. Oxford: Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00115-2>
- Bybee, J. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455. <https://doi.org/10.1080/01690969508407111>
- Christensen, L. H. & Christensen, R. Z. 2019. *Dansk Grammatik* (Danish Grammar), 2nd edn. Odense: University of Southern Denmark.
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511554377>
- Clark, E. V. 2014. Acquisition of derivational morphology. In *The Oxford Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 424–439. Oxford: OUP.
- Clark, E. V. & Hecht, B. F. 1982. Learning to coin agent and instrument nouns. *Cognition* 12(1): 1–24. [https://doi.org/10.1016/0010-0277\(82\)90027-0](https://doi.org/10.1016/0010-0277(82)90027-0)
- Den Danske Ordbog* (The Danish Dictionary). Copenhagen: Det Danske Sprog- og Litteraturselskab & Gyldendal. <<https://ordnet.dk/ddo/>> (30 April 2021).
- Derwing, B. L. & Baker, W. J. 1986. Assessing morphological development. In *Language Acquisition: Studies in First Language Development*, 2nd edn, P. Fletcher & M. Garman (eds), 326–338. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511620683.019>

- Diderichsen, P. 1946. *Elementær Dansk Grammatik* (Elementary Danish Grammar). Copenhagen: Gyldendal.
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2010. First language acquisition of compounds with special emphasis on early German child language. In *Cross-Disciplinary Issues in Compounding* [Current Issues in Linguistic Theory 311], S. Scalise & I. Vogel (eds), 323–344. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.311.24dre>
- Dressler, W. U., Sommer-Lolei, S., Korecky-Kröll, K., Argus, R., Dabašinskienė, I., Kamandulytė-Merfeldiene, L., Ljäs, J. J., Kazakovskaya, V. V., Laalo, K. & Thomadaki, E. 2019. First-language acquisition of synthetic compounds in Estonian, Finnish, German Greek, Lithuanian, Russian and Saami. *Morphology* 29: 409–429. <https://doi.org/10.1007/s11525-019-09339-0>
- Gregersen, F. & Kristiansen, T. 2009. *Dansk sproglære* (Danish language learning). Copenhagen: Dansk lærerforeningens Forlag.
- Hansen, E. & Heltoft, L. 2011. *Grammatik over det danske sprog I* (Grammar of the Danish), Vol. 1–3. Copenhagen: Det Danske Sprog- og Litteraturselskab.
- Herslund, M. 2001. The Danish s-genitive: From affix to clitic. *Acta Linguistica Hafniensia* 33(1): 7–18. <https://doi.org/10.1080/03740463.2001.10412193>
- Herslund, M. 2002. *Danish* [Languages of the World 382]. Munich: Lincom.
- Jespersen, O. 1923. *Børnesprog* (Child Language). Copenhagen: Gyldendal.
- Katlev, J. 2000. *Etymologisk ordbog* (Etymological Dictionary). Copenhagen: Politikens Forlag.
- Kilani-Schoch, M. & Dressler, W. U. 2005. *Morphologie naturelle et flexion du verbe français* (Natural Morphology and French Verb inflection). Tübingen: Narr.
- Kjærbaek, L. 2013. Tilegnelse af bøjningsmorfologi: En undersøgelse af substantivernes pluralisbøjning hos normalt udviklede danske børn i alderen 0–10 år (The acquisition of inflectional morphology: A study of noun plural inflection in typically developing Danish children in the ages of 0–10 years). PhD dissertation, University of Southern Denmark.
- Kjærbaek, L. 2015. Danske børns tilegnelse af bøjningsmorfologi – udviklingen af substantivernes pluralisbøjning i alderen 0–10 år (Danish children's acquisition of inflectional morphology – the development of noun plural inflection in the ages of 0–10 years). *Nydanske Sprogstudier (NyS)* 48: 106–138. <https://doi.org/10.7146/nys.v48i48.21137>
- Kjærbaek, L. & Basbøll, H. 2015. Testing hypotheses on frequency effects – Noun plural inflection in Danish children. In *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015*, E. Babatsouli & D. Ingram (eds), 168–181. <<http://ismbs.eu/data/documents/Proceedings-ISMBBS-2015.pdf>> (30 April 2021).
- Kjærbaek, L. & Basbøll, H. 2016. Interaction between input frequency, transparency and productivity in acquisition of noun plural inflection in Danish. *Poznan Studies in Contemporary Linguistics* 52(4): 663–686. <https://doi.org/10.1515/psicl-2016-0026>
- Kjærbaek, L. & Basbøll, H. 2017. Compound nouns in Danish child language. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketrez & M. Kilani Schoch (eds), 39–62. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.03kja>
- Kjærbaek, L. & Basbøll, H. 2019. Competition in the acquisition of Danish noun plural markers. In *Nordic Child Language*, L. Kjærbaek & A. Esbensen (eds). *RASK – International Journal of Language and Communication* 49: 25–48.
- Kjærbaek, L., dePont Christensen, R. & Basbøll, H. 2014. Sound structure and input frequency on noun plural acquisition: Hypotheses tested on Danish children across different data types. *Nordic Journal of Linguistics* 37(1): 47–86. <https://doi.org/10.1017/S0332586514000067>

- Korecky-Kröll, K., Sommer-Loeie, S. & Dressler, W. U. 2017. Emergence and Early Development of German Compounds. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketrez & M. Kilani-Schoch (eds), 19–37. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.02kor>
- Laaha, S., Kjørnbæk, L., Basbøll, H. & Dressler, W. U. 2011. The impact of sound structure on morphology; An experimental study of children's acquisition of German and Danish noun plurals focusing on stem change. *Acta Linguistica Hafniensia* 43(2): 106–126. <https://doi.org/10.1080/03740463.2011.689689>
- Laws, J. 2019. Profiling complex word usage in the speech of English preschool children: frequency patterns and transparency characteristics. *First Language* 39(6): 593–617. <https://doi.org/10.1177/0142723719872669>
- Lieber, R. 2004. *Morphology and Lexical Semantics*. Cambridge: CUP.
- Lippeveld, M. & Oshima-Takane, Y. 2014. The effect of input on children's cross-categorical use of polysemous noun-verb pairs. *Language Acquisition* 22: 209–239. <https://doi.org/10.1080/10489223.2014.943902>
- Lippeveld, M. & Oshima-Takane, Y. 2015. Nouns to verbs to nouns: When do children acquire class-extension rules for deverbal nouns and denominal verbs? *Applied Psycholinguistics* 36: 559–588. <https://doi.org/10.1017/S0142716413000349>
- MacWhinney, B. 2000a. *The CHILDES Project: Tools for Analyzing Talk, Vol. 1: Transcription Format and Programs*. Mahwah NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. 2000b. *The CHILDES Project: Tools for Analyzing Talk, Vol. 2: The Database*. Mahwah NJ: Lawrence Erlbaum Associates.
- Madsen, T. O., Basbøll, H. & Lambertsen, C. 2002. OLAM – et semiautomatisk morfologisk og lydstrukturelt kodningssystem for dansk (OLAM – a semiautomatic morphological and phonological coding system for Danish). *Odense Working Papers in Language and Communication* 24: 43–56.
- Olsen, S. 2014. Delineating derivation and compounding. In *The Oxford Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 26–49. Oxford: OUP.
- Ordbog Over Det Danske Sprog* (Dictionary of the Danish Language). Copenhagen: Det Danske Sprog- og Litteraturselskab & Gyldendal. <<https://ordnet.dk/ods>> (30 April 2021).
- Paulin, M. N. & Beckman, M. S. 2004. *Danske børns tilegnelse af nominalmorfologi – Undersøgelse af substantivernes pluralisbøjning hos førskolebørn* (Danish children's acquisition of nominal morphology – Study of noun plural inflection in preschool children) Unpublished examination paper. University of Southern Denmark.
- Peters, A. M. 1997. Language typology, prosody, and the acquisition of grammatical morphemes. In *The Crosslinguistic Study of Language Acquisition*, Vol. 5, D. I. Slobin (ed.), 136–197. Mahwah NJ: Lawrence Erlbaum Associates.
- Plunkett, K. 1985. *Preliminary Approaches to Language Development*. Aarhus: Aarhus University Press.
- Plunkett, K. 1986. Learning strategies in two Danish children's language development. *Scandinavian Journal of Psychology* 27: 64–73. <https://doi.org/10.1111/j.1467-9450.1986.tb01188.x>
- Rajnik, E. 2011. *Wortbildung des Substantives im Dänischen – explizite und implizite Derivation* (Noun derivation in modern Danish – affixational and affixless derivation) [Seria Filologia Skandynawska 13]. Poznan: Adam Mickiewicz University Press.
- Slobin, D. I. 1973. Cognitive prerequisites for the development of grammar. In *Studies of Child Language Development*, C. A. Ferguson & D. I. Slobin (eds), 175–208. New York NY: Holt, Rinehart & Winston.

Early phases of development of German derivational morphology

Sabine Sommer-Lolei^{1,3}, Veronika Mattes², Katharina Korecky-Kröll³ and Wolfgang U. Dressler^{1,3}

¹Austrian Academy of Sciences / ²University of Graz / ³University of Vienna

This chapter presents for the first time an overall description of the development of German derivational morphology. The longitudinal spontaneous speech corpora of three children acquiring Standard Austrian German and their mothers is analysed up to 3;0, with outlooks up to 6;0. The analysis focusses on nominal, verbal and adjectival derivation patterns that play a role at this early age, as well as on age and order of emergence, productive use and development of various patterns and its possible correlation to the respective input. The mini-paradigm criterion is extended to derivational morphology for attesting potential productivity in child speech. Only a very small inventory of derivational patterns is acquired up to 3;0; acquisition of derivation accelerates remarkably during preschool age.

Keywords: mini-paradigm criterion, particle verbs, agent and instrument nouns, conversion, productive use

1. Introduction

German is a West Germanic language, closely cognate to Dutch and English, and thus is a compounding rather than a derivational language (cf. Korecky-Kröll, Sommer-Lolei & Dressler 2017). This chapter focusses on the emergence of derivational morphology (with the exception of diminutives dealt with in Savickienė & Dressler 2007 and Dressler & Korecky-Kröll 2015) in three children acquiring Standard Austrian German in Vienna up to the age of 3;0 including child-directed speech (CDS) with brief outlooks into the development up to the age of 6;0. The (derivational) morphology of Standard Austrian German, as widely spoken in Vienna, differs little from that of Standard German as spoken in Germany (cf. Klampfer 2003; vs. Bittner 2003; Korecky-Kröll & Dressler 2007, 2009; Dressler, Lettner & Korecky-Kröll 2010, 2012).

Our aim is to present for the first time an overall description of how derivational morphology as a whole arises and develops, with an additional focus on the impact of child-directed speech on age and order of acquisition of different types of morphological derivation. Our main research questions are:

1. Which derivational patterns emerge and are acquired first?
2. To what extent does this depend on lemma or token frequencies of children's input?
3. Does it also depend on productivity of patterns in the input?
4. How does complexity of derivational morphology rise?

2. German derivational morphology

Adult derivational morphology (in Section 2.1, 2.2 and 2.3), as described in grammars and specific reference works (such as Kühnhold & Wellmann 1973; Wellmann 1975; Kühnhold, Putzer & Wellmann 1978; Elsen 2011; Fleischer & Barz 2012) and as attested in electronic corpora, especially the Austrian Media Corpus¹ (AMC), is discussed with a focus on the patterns likely to be (at least partially) acquired by the age of six years.

German derivational morphology is rich in various types of affixation and makes use of some root modification, i.e. ablaut (apophony), either alone or accompanying suffixation (the latter being similar to umlaut/metaphony). It also uses conversion, but to a lesser degree than English. The bases and targets of derivational rules are mostly limited to nouns, verbs and adjectives, and, more often than not, suffixation changes the word class, in contrast to prefixation. Latinate derivational morphology is much less important than in English and plays no role in toddlers' and preschool children's acquisition. In general, German derivational morphology is highly polysemous/polyfunctional and is subject to manifold, and often subtle, constraints. For example, the nominal suffix *-er* has several meaning categories, such as persons, objects and actions, all of which have further subcategories (such as agents, professions, inhabitants or persons). The suffix is restricted to nominal and verbal bases but is not compatible with certain verb types, such as psych verbs or intransitives (Plank 1981: 124; Eisenberg 2013: 262; Fleischer & Barz 2012: 201–202). Eisenberg (2013: 32) specifies an inventory of approximately 100 word-formation affixes (of which only a quarter are productive), most of which modify nouns.

1. The Austrian Media Corpus (AMC) of the Austrian Academy of Sciences, established on data of the APA (Austrian Press Agency), covers all the Austrian Print Media of the last 30 years and consists of over 10 billion word tokens. <<https://www.oeaw.ac.at/acdh/tools/amc-austria-media-corpus/>> (5 September 2020).

Derivational prefixes and suffixes differ fundamentally in their phonological, prosodic and morphophonological characteristics. Prefixes are less salient for young children than suffixes due to their position because of the recency effect, but they are phonologically consistent, whereas suffixes are often involved in resyllabifications (see Eschenlohr 1999; Eisenberg 2013; Sommer-Lolei et al. to appear).

2.1 Noun formation

Derivational suffixes in German are heads determining word class, inflection class and gender. The most productive patterns of the more than 70 derivational affixes are deverbal agent and instrument nominalizations in *-er*:

- (1) *Fahr-er* 'driver' from *fahr-en*² 'drive'
- (2) *Staub+saug-er* 'vacuum cleaner, lit. dust sucker' from *saug-en* 'suck'

Denominal ones are much less productive:

- (3) *Gärtn-er* 'gardener' (with umlaut) from *Garten* 'garden'

Productive action and result noun formations can be found in *-ung*:

- (4) *Reinig-ung* 'cleaning' from *rein-ig-en* 'clean'

Other deverbal agent and instrument derivations, such as *-e* are frequent but weakly productive:

- (5) *Reib-e* 'grater' from *reib-en* 'grate'

And unproductive, if accompanied by ablaut:

- (6) *Gab-e* 'gift' from *geb-en* 'give'

Also, pure ablaut formation (often referred to as implicit derivation) is unproductive:

- (7) *Trank* or *Trunk* 'drink' from the verb *trink-en* 'drink'

The deverbal circumfixation *Ge-...-e* is productive:

- (8) *Ge-red-e* 'gossip' from *red-en* 'talk'

2. Unless specified otherwise, all verb examples throughout the chapter are given in the infinitive form (*-en* as the infinitive suffix).

Conversion comes in two patterns: fully productive syntactic nominalization of infinitives:

- (9) *das Les-en* ‘the reading’

And less productive morphological conversion of the verb stem (also called zero-derivation):

- (10) *Sitz* ‘seat’ from *sitz-en* ‘sit’
 (11) *Koch* ‘cook’ from *koch-en* ‘cook’

Denominal noun formation by prefixation is partially productive:

- (12) *Miss-erfolg* ‘failure, lit. mis-success’

2.2 Verb formation

The most productive pattern is particle verb formation from nominal, adjectival and verbal bases with 60 different particles (in our largest child corpus), e.g. *weg-räum-en* ‘put away’ – *weg-ge-räum-t* (away-PST.PTCP-put-PST.PTCP) ‘put away’ *Ich räum-e weg* (I put-1SG.PRS away) ‘I put away’) from *Raum* ‘space’, *aus-leer-en* ‘clear away, empty’ from the adjective *leer* ‘empty’, *an-mal-en* ‘paint sth.’ from *mal-en* ‘paint’. Particles always carry the only or main stress and switch into a postverbal, often sentence-final, position in finite forms, as in *Er mal-t XY an* ‘He paints XY’. The past participle prefix *ge-* and the infinitive particle *zu* ‘to’ are inserted between the particle and the verb root, as in *an-ge-mal-t* ‘painted’, *an-zu-mal-en* ‘to paint sth.’, thus verb particles (also called separable prefixes) are prosodically and positionally salient. Although they can contribute to abstract verb meanings (e.g. *aus-halt-en* ‘stand, support’ from *halt-en* ‘hold’), they all have a very concrete meaning when they recur as autonomous prepositions and/or adverbs (e.g. *auf* ‘up’, *mit* ‘with’, *raus* ‘out’). As particle verbs strongly combine syntactic and morphological features, their theoretical classifications differ considerably. Particles are treated as a category of prefixes or as a derivation pattern of their own, while other authors analyse particle verbs as compounds. They share properties with both prefixes and compound constituents and are homophonous with prepositions or adverbs, often having identical or similar meaning (more in Dehe 2015 and McIntyre 2015).

In contrast, the more than 20 inseparable verb prefixes are always unstressed. They are perceptually less salient than particles and have much more abstract meanings. The five fully morphologized (Eisenberg 2013: 244) prefixes in German are *be-*, *ent-*, *er-*, *ver-*, and *zer-*.

The most productive prefixes occurring in our data are *ver-* and *be-* (which are also the most frequent ones in German according to Duden 1998), as in *be-setz-en* ‘occupy’ from *setz-en* ‘set’, *ver-schlecht-er-n* ‘worsen’ from adjective *schlecht-er* ‘worse’. They are typically transitivising, but especially *ver-* is highly polysemous and covers a broad range of meanings, such as egressive, causative, or aberrant (details in Fleischer & Barz 2012; Eisenberg 2013; Dewell 2015). Besides *ver-* and *be-*, the prefix *ge-*, which is not productive in German and is often semantically opaque (e.g. *ge-fall-en* ‘like’ vs. *fall-en* ‘fall’) or even lacks a synchronic lexical base (e.g. *ge-winn-en* ‘win’), is also among the most frequent in our data.

Of the few verbal suffixes, the most productive ones are *-el* as in *tröpf-el-n* ‘drip’ from *tropf-en* ‘to drop’ and *-ier* as in *marsch-ier-en* ‘to march’ from *Marsch* ‘march’.

The verbal circumfixes are *be-...-ig* and *ver-...-ig*, e.g. *be-glaub-ig-en* ‘accredit sth.’, *ver-fest-ig-en* ‘compact, stiffen’.

Derivation via root modification is fossilized in a few more verb pairs than in English. Corresponding examples are *sitz-en* ‘sit’ and *setz-en* ‘set’, *fall-en* ‘fall’ and *fäll-en* ‘fell sth.’. Verb derivation from nouns or adjectives via conversion is much less productive than in English, e.g. *land-en* ‘land’ from *Land* ‘land’, *reif-en* ‘ripen’ from *reif* ‘ripe’. Prefixation is preferred over conversion, e.g. *ent-laus-en* ‘delouse’, *er-blass-en* ‘turn pale’.

Various devices can optionally or obligatorily be combined with each other, e.g. *auf-weck-en* ‘wake (sb.) up’ or *er-weck-en* ‘arouse (sth./sb.)’ from *wach* ‘awake’ (stem modification + particle or prefix), *um-ver-teil-en* ‘redistribute’ (particle + prefix), *um-lack-ier-en* ‘revarnish’ (particle + suffix).

2.3 Adjective formation

The most productive suffixes are denominal and deverbal *-ig* and deverbal *-bar*, along with deadjectival *-lich* and denominal and often pejorative *-isch*, as in *berg-ig* ‘mountainous’ from *Berg* ‘mountain’, *brenn-bar* ‘flammable’ from *brenn-en* ‘burn’, *gefähr-lich* ‘dangerous’ from *Gefahr* ‘danger’, *kind-isch* ‘childish’ from *Kind* ‘child’. The most productive prefix is negative *un-*, as in *un-brenn-bar* ‘incombustible’. Also, past and present participles frequently have an adjectival function, as in *schmerz-end* ‘hurting’, and *ge-fütter-t* as an adjective only in the sense of ‘lined’ instead of the past participle meaning of ‘fed’. Some adjectives can either function as adverbials or have the same form as adverbs (see Schäfer 2013).

Adverbs are nearly always only syntactically distinguishable from adjectives, thus, their role in derivation is very marginal (none in early child language).

3. Previous research on the acquisition of German derivational morphology

Although the development of word formation in German is an important feature of the language (e.g. Rainer 2010: 9; Schipke & Kauschke 2011: 69; Kauschke 2012: 72), it is still understudied. The only patterns well examined so far are nominal compounding (e.g. Dressler, Lettner & Korecky-Kröll 2010, 2012; Korecky-Kröll, Sommer-Lolei & Dressler 2017), diminutivization (e.g. Korecky-Kröll & Dressler 2007; Dressler, Lettner & Korecky-Kröll 2012) and noun derivation by the suffix *-er* (especially Meibauer 1999; Meibauer, Guttropf & Scherer 2004), i.e. dominant word-formation patterns in early language acquisition. Korecky-Kröll (2011) investigated noun derivation in two of the three child corpora reanalysed in this chapter (one up to age 3;0 and another up to 6;0) and found only evidence for productive use of the suffix *-er* as well as very few examples of neologisms of *-ung* and conversions. Noun prefixes were not used productively until age 6;0.

Even less has been reported about the acquisition of verbal derivation (e.g. Behrens 1998; Schulz, Wymann & Penner 2001), and virtually nothing about adjectival derivation. One reason is the minor role of these devices in early childhood. Rainer (2010) and Mattes (2018) offer detailed studies on the whole range of later acquired derivation patterns up to age 7;0.

We know from Schipke and Kauschke's (2011) investigation that German-acquiring children show a significant increase of word formation in their speech production between 1;9 (4.66% tokens/7.46% types) and 3;0 (8.76% tokens/17% types). The most impressive increase concerns compounding and affixation, whereas conversion and root modification remain infrequent.

Diary studies, such as Neugebauer-Kostenblut (1914); Stern and Stern (1928); Elsen (1999), and Rainer (2010), demonstrate that children detect the possibility of deriving words around their second birthday, as shown by the frequent production of derivational neologisms. However, they restrict themselves to a small subset of available patterns and elaborate their inventory only gradually during preschool and school age.

The first acquired and dominating nominal derivation pattern (besides diminutive suffixation) is the frequent and highly productive suffix *-er*, deriving – in first instance – instrument and agent nouns.³ Other early acquired nominal derivation patterns are the frequent, but weakly productive suffix *-e*,⁴ as well as morphological

3. Examples for early neologisms are (2;0) *Reib-er* 'rasp' instead of canonical *Feil-e* from Neugebauer-Kostenblut (1914), (2;0) *Schlaf-er* 'lit. sleep-er' for 'bed' instead of *Bett*, or (2;0) *Tank-er* 'someone who is refueling' from Elsen (1999).

4. (1;9) *Wisch-e* based on *wisch-en* 'to wipe' for 'mop' (instead of *Lappen*) or *Zähne+putz-e* 'lit. teeth+brush-e' for 'tooth powder' instead of canonical *Zahn+pulver* from Neugebauer-Kostenblut (1914).

conversion.⁵ Besides these first acquired derivation patterns, the suffix *-ung*, one of the most frequent nominalization suffixes of German, emerges usually during the 3rd year. It is sporadically used for coinages from early on by some children, however, it cannot be considered fully productive before age 5;0 or 6;0.⁶

The most important verb derivation pattern in early language acquisition is the highly frequent particle verb formation and, to a minor degree, conversion. Inseparable verbal prefixes are actually irrelevant up to 3;0 but are acquired gradually soon thereafter (Behrens 1998; Rainer 2010; Schipke & Kauschke 2011; Mattes 2019).

The only adjective derivational pattern that plays a noteworthy role up to age 3;0 is the suffix *-ig* (Rainer 2010; Mattes 2018: 313).

4. Acquisition data and methods

4.1 Data

We analyse the longitudinal spontaneous speech corpora of three typically developing children and their mothers (all of higher socio-economic status). All participants are monolingual native speakers of Standard Austrian German. They were recorded in their homes in Vienna 1 to 4 times per month in varying everyday situations (such as playing, eating, etc.). The data consist of 79.35 hours of recording and 285,905 transcribed word tokens in child speech (CS) and child-directed speech, see Table 1. All of the data were transcribed and coded using an adapted German version of CHILDES (cf. MacWhinney 2000).

Table 1. CS and CDS data

Corpus	Period	No. utterances		No. word tokens		No. rec.	No. hours of rec.
		CS	CDS	CS	CDS		
JAN (m)	1;3–3;0	15,602	33,476	36,458	159,636	91	45.75
KAT (f)	1;6–3;0	3,416	6,919	6,805	27,436	34	13.6
LEN (f)	1;7–3;0	6,066	11,883	13,551	42,019	40	20.0
Total				56,814	229,091	165	79.35

5. (2;1) *Tisch+deck* ‘table+set’ for the action of setting the table from Neugebauer-Kostenblut (1914).

6. (3;2) *Ver-bind-ung* ‘bandage’ instead of canonical *Ver-band* from Elsen (1999), or *Blink-ung* ‘blinking’ from Rainer (2010).

4.2 Methods

Since we focus on the analysis of the spontaneous speech of the children and their mothers, we excluded adult-directed speech, spontaneous imitations by the children (when the child repeated the utterance or word immediately after the mother) and all citation forms (i.e. book reading, singing, poems etc.).

All nouns, verbs and adjectives were annotated and categorized as either simplex, compound or derivational (including synthetic compounds), and in the latter case assigned to the respective derivation pattern. In cases of more complex forms with more than one derivation, we counted the last applied derivation pattern (e.g. *-nis* in *Ge-fäng-nis* ‘prison’ = *fang-en* ‘catch’ → *ge-fang-en* ‘caught’ → *Ge-fäng-nis* = suffixation including umlauting). CS and CDS were analysed separately, and the distributions of the derivation patterns were calculated.

The distinction between actual and pseudo-affixation, as in baseless *fertig* ‘done’ (LEN 1;11), affix polysemy and homophony, allomorphy and double affixation, was done following the morphological literature when available. For this purpose, we consulted printed and main electronic dictionaries (duden.de, canoonet.eu). In our analyses, very opaquely lexicalized forms, e.g. *Schling-el* ‘(little) rascal’ from *schling-en* ‘wind, devour’, were excluded. In our corpora, we measured lemma and token frequencies in CS and CDS, but not inflectional type frequencies, which is of minor relevance for word formation.

Mini-paradigm criterion (MPC)

For assessing when derivational processes start to become productive, we elaborated the mini-paradigm criterion (first introduced for the acquisition of inflection in Kilani-Schoch & Dressler 2005; Dressler, Lettner & Korecky-Kröll 2010 and adapted to compounding in Korecky-Kröll, Sommer-Lolei & Dressler 2017). Here, we adapt the criterion to derivational morphology; thus, the mini-paradigm criterion is fulfilled when a morphological process (affixation or conversion) is applied to (at least) three different bases, and when each base occurs in (at least) three different types (uninflected or inflected words, compounds or other derivations), as illustrated below:

(13) *-er* derivation

- a. Part 1: a suffix⁷ combined with at least three different bases: at age 1;8 Jan combines the suffix with six different bases:⁸ *Last-er* ‘truck’, *Schleck-er* ‘lollipop’, *Flitz-er* ‘speedster’, *An-häng-er* ‘trailer’, *Transport-er* ‘transporter’, *Ständ-er* ‘stand’ = partial fulfilment.

7. Note that we do not treat different semantic functions of one suffix separately but identify it on the basis of formal criteria (see Introduction of this volume).

8. Note that we do not differentiate between bases belonging to different word classes.

- b. Part 2: a base occurs in at least three different types: Jan uses the verbal base *fahr-* at 1;8: *fahr-en* (drive-INF) 'drive', *fahr+rad* 'drive+bike, bicycle', and one month later at 1;9 *weg-fahr-en* (away-drive-3PL) '(they) drive away' = partial fulfilment.
- c. Part 3: since *fahr-* is only combined another month later with the agentive suffix in *Fahr-er* 'driver' the mini-paradigm criterion is completely fulfilled only at 1;10.

5. Distribution of derivations in CS and CDS

Table 2 presents the total number of nouns, verbs and adjectives, contrasted with the number of derived nouns, verbs and adjectives and the resulting percentages of derived nouns, verbs and adjectives within their specific category.

Table 2. Frequencies of derived noun, verb and adjective lemmas and tokens in relation to noun, verb and adjective lemmas and tokens in input and output up to 3;0

Speaker	CDS (J)	JAN	CDS (K)	KAT	CDS (L)	LEN
N lemmas	1,959	809	730	317	778	308
N tokens	15,855	6,809	2,773	1,134	4,087	1,804
N DERIV lemmas	302	105	112	36	123	44
N DERIV tokens	1,407	481	369	261	603	263
% N DERIV lemmas	15.4	13.0	15.3	11.4	15.8	14.3
% N DERIV tokens	8.9	7.1	13.3	23.0	14.8	14.6
V lemmas	1,193	450	466	173	630	221
V tokens	26,770	6,190	5,255	850	7,176	2,003
V DERIV lemmas	921	299	316	96	437	124
V DERIV tokens	5,322	1,314	1,226	209	1,877	413
% V DERIV lemmas	77.2	66.4	67.8	55.5	69.4	55.9
% V DERIV tokens	19.9	21.2	23.3	24.6	26.2	20.6
ADJ lemmas	334	119	122	34	142	47
ADJ tokens	4,650	1,086	739	106	924	189
ADJ DERIV lemmas	118	26	19	2	39	2
ADJ DERIV tokens	618	124	61	2	82	2
% ADJ DERIV lem.	35.3	21.8	15.6	5.9	27.5	4.3
% ADJ DERIV tok.	13.3	11.4	8.3	1.9	8.9	1.1

When comparing CS and CDS frequencies of derived nouns (N DERIV), verbs (V DERIV) and adjectives (ADJ DERIV), as a percentage of noun, verb and adjective lemmas, we find that in all groups, children produce fewer derivations than their mothers.

Chi-square tests on simplex vs. derived nouns (lemmas and tokens) showed that all three children used significantly fewer derived nouns than their mothers (lemmas: Jan (JAN): $\chi^2 = 8.343$, $df = 1$, $p = .004$, Lena (LEN): $\chi^2 = 7.298$, $df = 1$, $p = .007$, Kathi (KAT): $\chi^2 = 8.990$, $df = 1$, $p = .003$; tokens: Jan: $\chi^2 = 8.269$, $df = 1$, $p = .004$; Lena: $\chi^2 = 28.255$, $df = 1$, $p < .001$; Kathi: $\chi^2 = 18.686$, $df = 1$, $p < .001$).

Similar effects were found for simplex vs. derived verb lemmas (Jan: $\chi^2 = 20.899$, $df = 1$, $p < .001$, Lena: $\chi^2 = 6.275$, $df = 1$, $p = .012$, Kathi: $\chi^2 = 8.208$, $df = 1$, $p = .004$). However, Jan used significantly more derived verb tokens than his mother ($\chi^2 = 4.859$, $df = 1$, $p = .027$), whereas the opposite effect was found for Lena ($\chi^2 = 24.148$, $df = 1$, $p < .001$) and Kathi showed a similar use as her mother ($\chi^2 = 1.240$, $df = 1$, $p = 0.266$).

Chi-square tests on simplex vs. derived adjectives revealed significantly fewer adjective lemmas only in Jan and Lena in comparison to their mothers (Jan: $\chi^2 = 8.530$, $df = 1$, $p = .003$, Lena: $\chi^2 = 9.142$, $df = 1$, $p = .002$), whereas in tokens, similar effects were only found for both girls (Lena: $\chi^2 = 12.694$, $df = 1$, $p < .001$, Kathi: $\chi^2 = 5.403$, $df = 1$, $p = .020$), in tokens only in Kathi ($p = .020$) and Lena ($p < .001$).

5.1 Noun derivation

The distribution (frequencies and percentages) of the patterns of derivations vs. simplex nouns and compounds are given for each child and their mothers in Table 3 (see Appendix); percentages for CS and CDS all together are given in Figures 1 and 2.

The distribution shows that diminutive suffixes take up the largest portion of nominal derivation in all mother-child-dyads. They play an essential role in the acquisition of derivation. However, we will exclude them from further analysis because of their special functional status. While diminutives have an evaluative function (i.e. on the pragmatic level), the other derivational patterns cause semantic and/or syntactic changes. For more on diminutives, see Korecky-Kröll & Dressler (2007) and Dressler & Korecky-Kröll (2015).

The most frequently used derivation patterns are suffixations of *-er*, *-e*, and conversion (see Table 3, Appendix; and Figures 1 and 2). The suffix *-ung* plays a minor role, but it emerges in all children before age 3;0 and reaches our criterion of potential productivity in Jan's and partially in Lena's speech (see Section 7.1, Table 6).⁹

9. Lena fulfils the mini-paradigm criterion only partially at 2;11 when combining the suffix *-ung* with three different bases, but none of the bases occurs in three different types.

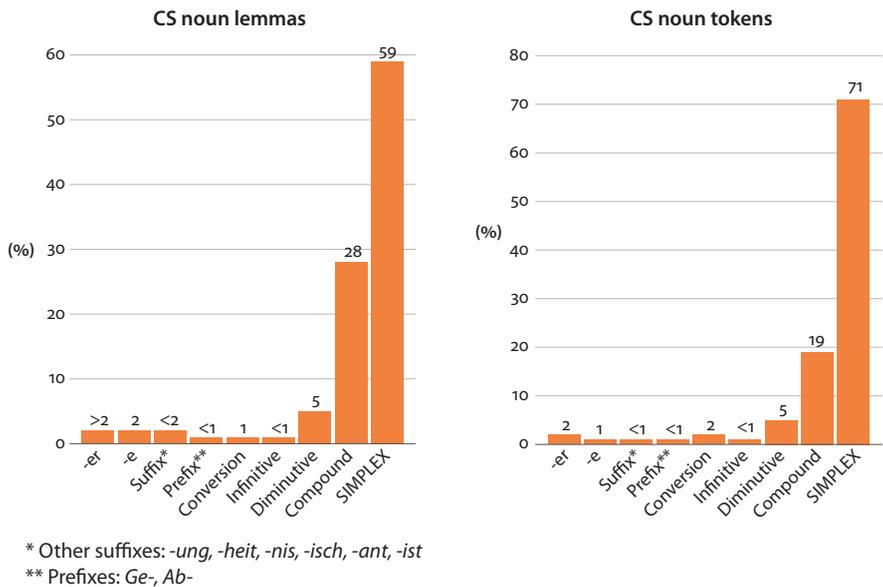


Figure 1. Distribution (percentages) of simplex nouns, compounds and derived nouns in lemmas and tokens in the three children's CS up to 3;0

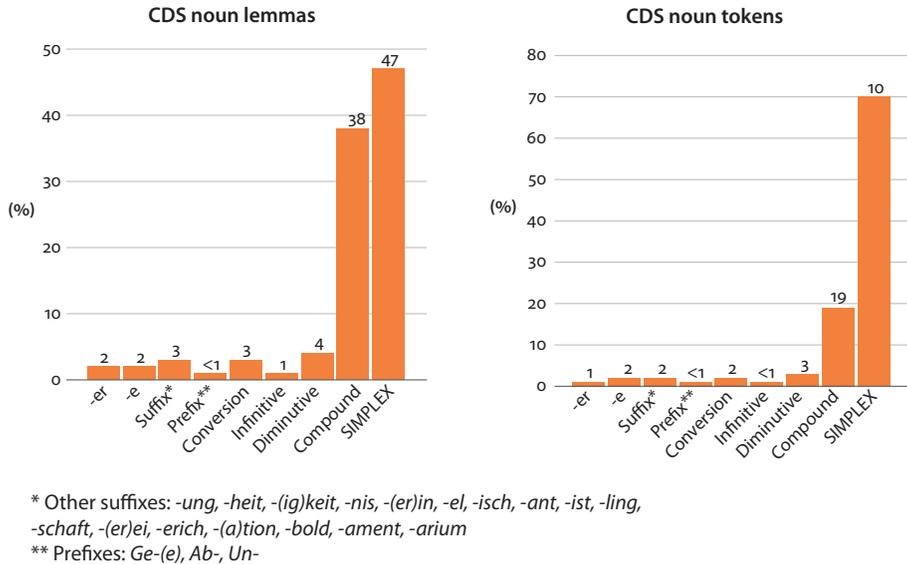


Figure 2. Distribution (percentages) of simplex nouns, compounds and derived nouns in lemmas and tokens in the three children's CDS up to 3;0

Chi-square tests on simplex vs. derived nouns revealed a similar use in lemmas in all three children in all observed suffixes (*-er*, *-e*, *-ung*), but significant differences in tokens in Jan (*-er/-e*: $p < .001$, *-ung*: $p = .005$) and in Lena (*-ung*: $p = .012$). For conversion, we find significant differences in lemmas in Jan ($p = .006$) and Lena ($p = .013$) and in tokens in Kathi ($p = .002$).

Children's use of suffixes that derive mainly abstract nouns, are limited to *-heit* and *-nis*, excluding *-(ig)keit* and *-schaft*. The Latinate suffixes *-ant* and *-ist* for deriving denominal agent nouns occur sporadically. All of the suffixes mentioned above (Figures 1 and 2, Suffix*) are extremely rare.

5.2 Verb derivation

The distributions (frequencies and percentages) of derived verbs and simplex verbs in lemmas and tokens (with the exception of auxiliary and modal verbs) in input and output are given in Table 4 (Appendix); percentages of CS and CDS are illustrated in Figures 3 and 4.

There is a clear preference for simplex verbs in tokens, but a preference for particle verbs in lemmas, which again shows the great productivity of particle verbs already present in early child language.

Conversion, as in *frühstück-en* 'to have breakfast', *trommel-n* 'to drum', plays a bigger role than prefixation in the beginning of acquiring derivational patterns, because distinction between morphological categories is less important in early

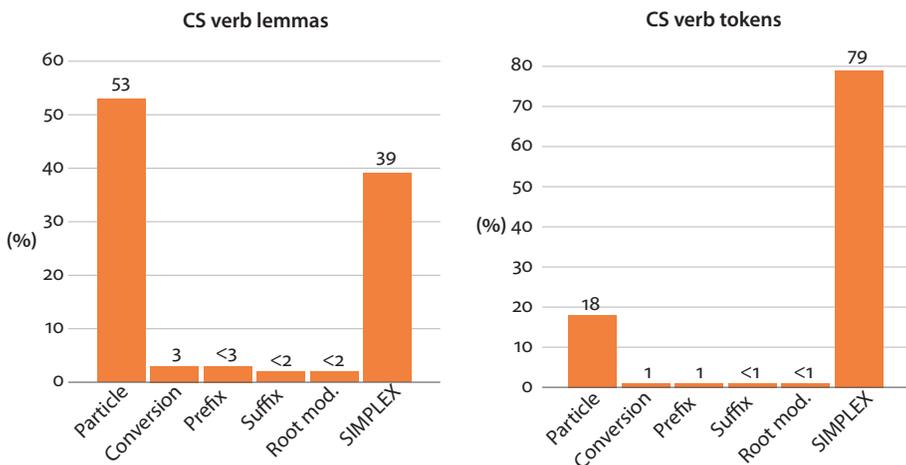


Figure 3. Distribution (percentages) of simplex and derived verb lemmas and tokens in the three children's CS up to 3;0

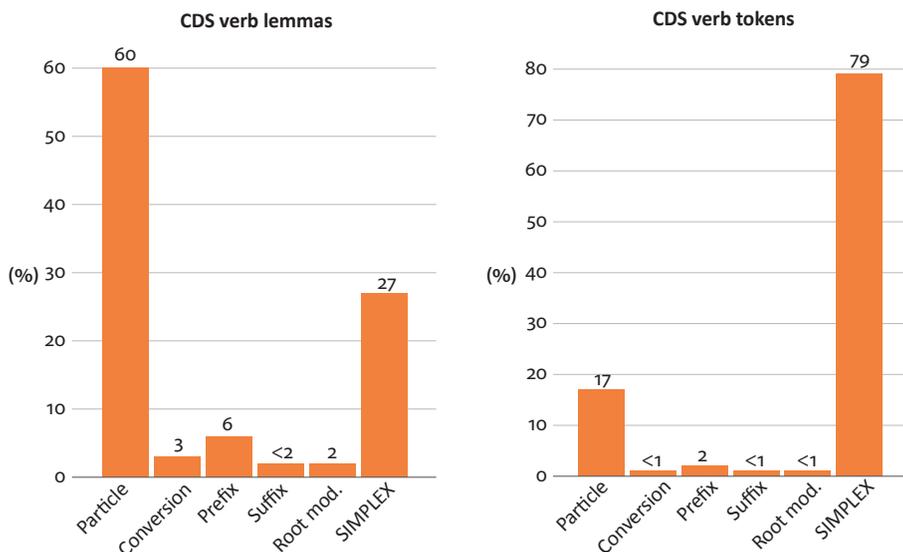


Figure 4. Distribution (percentages) of simplex and derived verb lemmas and tokens in the three children's CDS up to 3;0

child language than simplicity of formal relations, no formal change being simpler than affixation (Clark 1993: 199–203). Regarding the inventory of prefixes, we find more than double the number of prefixes in CDS than in CS (in all children).

It is obvious that particle verb formation dominates verb derivation, most impressively in lemmas (more than 50% of all verb lemmas, except for Kathi's CS), but also clearly in tokens (more than 15% in all speakers). This result is in accordance with the high productivity and frequency of verb particles in German (cf. also Behrens 1998). In German, almost every verb can be combined with various particles, and particle verbs can function as bases for further derivation (e.g. *An-häng-er* 'trailer' JAN 1;8).

In contrast, prefixes are much less frequent in the language, especially in our early CS data (less than 3% of all verb lemmas and 1.4% of tokens).

Furthermore, the frequencies of prefixed and particle verbs differ significantly between CS and CDS. Chi-square tests yielded significantly fewer prefixed and particle verb lemmas in all children in comparison to their mothers (prefixed verbs: Jan: $\chi^2 = 15.803$, $df = 1$, $p < .001$, Lena: $\chi^2 = 8.313$, $df = 1$, $p = .004$, Kathi: $\chi^2 = 4.283$, $df = 1$, $p = .039$; particle verbs: Jan: $\chi^2 = 17.057$, $df = 1$, $p < .001$, Lena: $\chi^2 = 7.746$, $df = 1$, $p = .005$, Kathi: $\chi^2 = 7.624$, $df = 1$, $p = .006$).

Similar effects were found for prefixed verb tokens in Jan ($\chi^2 = 6.678$, $df = 1$, $p = .010$) and prefixed as well as particle verb tokens in Lena (prefixed verbs:

$\chi^2 = 15.420$, $df = 1$, $p < .001$, particle verbs: $\chi^2 = 14.884$, $df = 1$, $p < .001$). In contrast, Jan used significantly more particle verb tokens than his mother ($\chi^2 = 8.994$, $df = 1$, $p = .003$), whereas Kathi showed a similar use of verb tokens as her mother in prefixed and particle verbs.

In the case of conversion, we find the exact opposite of prefixed and particle verbs, for all children in lemmas and tokens show a similar use to their mothers, with the exception of Kathi's token frequencies, which are significantly higher than her mother's ($p = .015$). The frequencies of conversion are variable among the three mother-child-dyads. Although Kathi does not have many verbs altogether (see Table 2), she uses a high number of derivations in tokens (24.6% of all verbs), of which there are 2.2% conversions (see Table 4, Appendix), due to the high token frequencies of *stempel-n* 'stamping' (2;10) and *tropfen* 'drip' (2;6), which account for more than half of the tokens.

Since suffixation plays only a marginal role in verb derivation, Jan's relatively high percentages of verb suffixes is surprising. In total, he produces only 10 different derived verb lemmas but 69 tokens: 4 deverbal lemmas derived by the diminutive suffix *-el*: e.g. *taps-el-n* 'plod' (1;10, plus 1 deadjectival *blöd-el-n* 'fool around' 2;11), which is expected to occur frequently in CDS and CS due to its expressive character, and 4 opaque derived verbs suffixed with *-ier*: e.g. *transport-ier-en* 'transport' (1;11), *prob-ier-en* 'try' (2;1). The last verb lemma has 18 tokens and together with the 26 tokens of *streich-el-n* 'to pet' (1;10) is responsible for the high token frequencies of *-el* and *-ier* (together 67 tokens out of 69). Thus, high-frequent verbs appear to have been taken over from CDS without morphological analysis.

We find a high variety of particle verb lemmas in both CS and CDS (see Figures 3 and 4), but a preference for simplex verbs in tokens up to 3;0. The most frequent particles in verbs (arranged according to decreasing frequency) are for Jan: *rein* 'in', *auf* 'up', *an* 'at', for Kathi: *an*, *auf* and *heraus* 'out', and for Lena: *aus* 'out', *an*, *auf*. Thus, the particles *an* and *auf* dominate, as in *an-schau-en* 'look at' (JAN 1;7, KAT 2;5, LEN 2;3) and *auf-mach-en* 'open (up)' (JAN 1;8, KAT 2;3, LEN 2;1).

As mentioned in Section 2.2, *be-* and *ver-* are the most frequent German prefixes. There are three prefixes that occur up to 3;0 in the speech of all three children and their CDS (in decreasing frequency): *ver-*, unproductive *ge-*, and *be-*, as in the highly opaque verbs *ver-steck-en* 'hide' (JAN 1;8), *ge-hör-en* 'belong to' (JAN 1;11), and *be-komm-en* 'get', the prefixed verbs with the highest token frequency. Nevertheless, we counted these verbs as complex, because both the prefix and the base occur in other combinations, i.e. structurally, morpheme boundaries are clear.

Additionally, only CDS frequently includes *zer-*, *er-*, *über-* (all three mothers), *ent-* in a few tokens in Lena's and Jan's mother and single tokens of *unter-*, *fern-* and Latinate *re-* in Lena's mother. For a detailed description of emergence and productivity see Section 7.2.

5.3 Adjective derivation

The distributions (frequencies and percentages) of derived and simplex adjectives in lemmas and tokens in CS and CDS are given in Table 5 (Appendix); percentages of the three children and their mothers are presented in Figures 5 and 6:

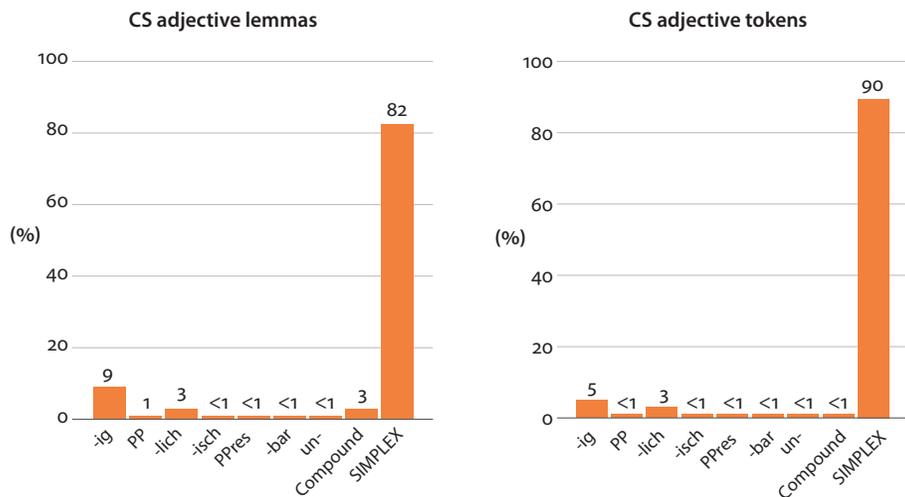


Figure 5. Distribution (percentages) of simplex and derived adjective lemmas and tokens in the three children's CS up to 3;0

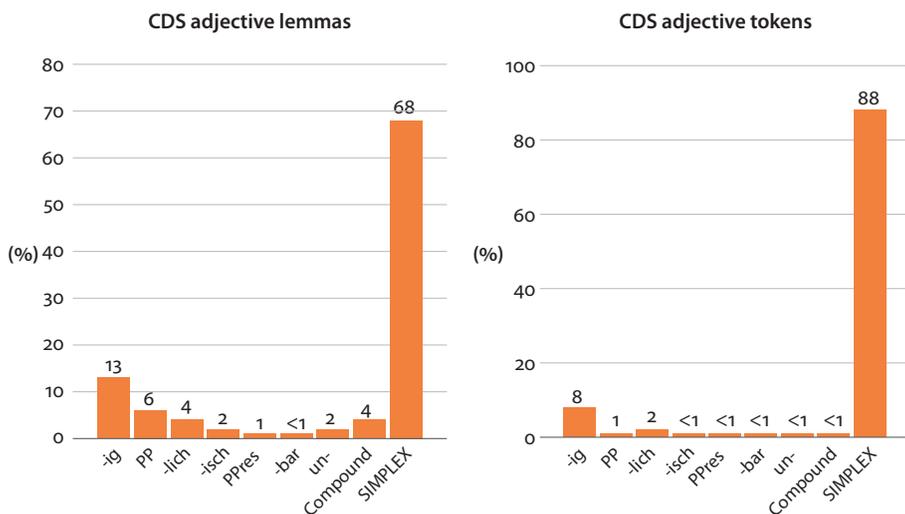


Figure 6. Distribution (percentages) of simplex and derived adjective lemmas and tokens in the three children's CDS up to 3;0

In the derivation of adjectives, the suffix *-ig* clearly dominates (e.g. in Jan at 1;9 denominal *durst-ig* (thirst-ADJ) ‘thirsty’, at 1;11 deverbal *rutsch-ig* (slide-ADJ) ‘slippery’ and only at 3;0 deadjectival *schwier-ig* (hard-ADJ) ‘difficult’); for Lena, it is even the only pattern used up to 3;0. The suffix *-lich* is used to a much more minor degree and only by Jan and Kathi. All other adjective derivation patterns appear only in Jan’s speech, e.g. *sicht-bar* (see-ADJ) ‘visible’ (2;0), *interess-ant* (interest-ADJ) ‘interesting’ (2;3), *quadrat-isch* (square-ADJ) ‘square/quadratic’ (2;9), *sprech-end* (talk-PRS.PTCP) ‘talking’ (2;10).

As the total numbers in Table 5 (Appendix) and the distributions in Figures 5 and 6 illustrate, adjectives and adjective formation play a marginal role in early child speech. Only less than 15% of lemmas and less than 10% of tokens in CS are derived adjectives, in CDS less than 28% of lemmas and less than 12% of tokens. Adjectives produced up to 3;0 are initially simple adjectives. For the emergence and productive use of derived adjectives see Section 7.3 and for their further development up to age 6;0 see Section 9.

6. Course of acquisition (nouns and verbs)

The course of acquisition differs for nouns and verbs; the first simplex nouns emerge earlier than simple verbs in Jan at 1;3 vs. 1;4, in Kathi at 1;6 vs. 2;0 and simultaneously only in Lena at 1;7. Concerning the occurrence of the first derivations, derived verbs tend to emerge earlier and develop faster than derived nouns: in Jan at 1;6 vs. 1;8, in Lena at 1;8 vs. 1;9 and only later in Kathi at 2;3 vs. 1;11 (see also Tables 6 and 7 in Section 7). The overall course of simplex vs. derived verbs and nouns is given in Figures 7 and 8.¹⁰

Figures 7 and 8 clearly demonstrate the dominance of simplex nouns in lemmas and tokens. This changes in token frequencies at mean age 2;5, when simplex verbs overtake simplex nouns. We find a constant increase of simplex and derived verbs which is parallel in lemmas in the course of acquisition, but also a priority of simplex verbs in tokens. There is a huge difference between verbs and nouns, because derived nouns do not show the same steady development as verbs and have lower frequencies in both lemmas and tokens throughout the course of development.

10. Because the number of recordings is highly variable from one month to another, we calculated mean values per month and child, to ensure comparability of the three children.

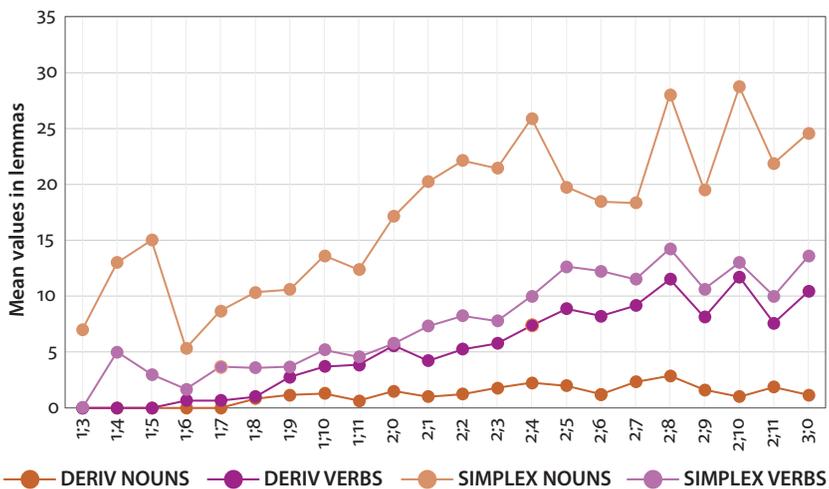


Figure 7. Course of acquisition in mean values of derived vs. simplex noun and verb lemmas per month in all three children (CS)

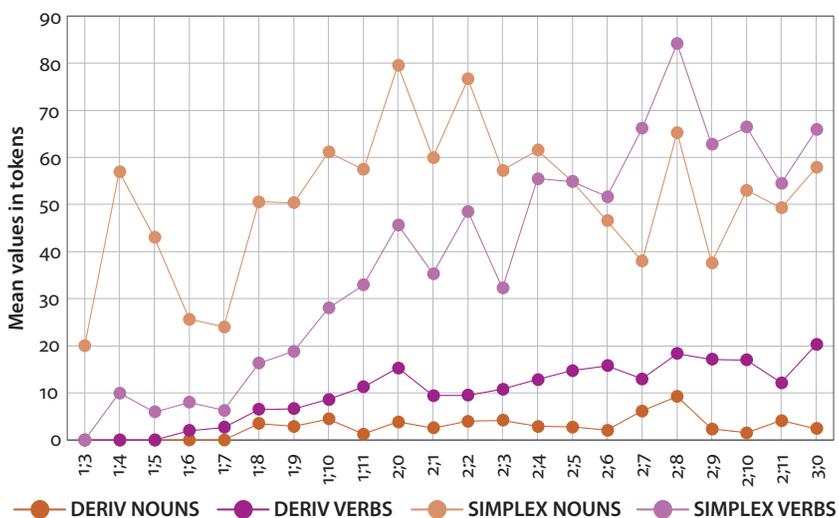


Figure 8. Course of acquisition in mean values of derived vs. simplex noun and verb tokens per month in all three children (CS)

The course of overall acquisition of different derived verb patterns is shown in Figure 9:

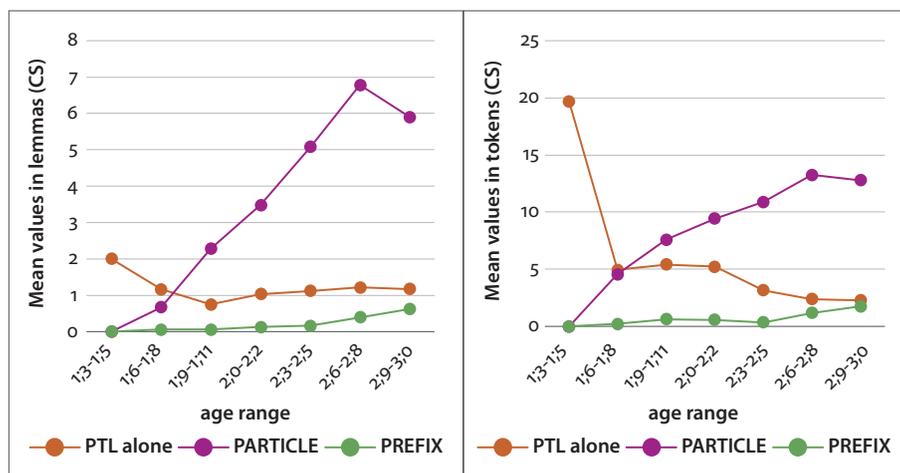


Figure 9. Mean values of particles alone (PTL), particle verbs and prefixed verbs in lemmas and tokens per quarter in all three children (CS)

Here, we observe a constant increase of particle verbs and a decrease of bare particles, but a much later slow increase of prefixed verbs towards the end of the investigated period. Thus, at 1;6 Jan first produces the particle *an* ‘on’ with the meaning of *anhaben* ‘lit. have on, wear’, later in the same month the particle verb *an-halt-en* (on-hold-INF) ‘hold on’, but also *halt-en* (hold-INF) ‘hold’ still with the meaning of *anhalten*, but without the particle, at 1;9 *fest-halt-en* (on-hold-INF) ‘hold on to’ and only at 2;10 *be-halt-en* (hold-INF) ‘keep’.

Due to the very low frequency of adjectives, we do not display their developmental course of acquisition.

7. Emergence and productive use

Frequency plays an important role for productivity (Bybee 2006), but “it is not the only factor in the learning process” (Cordes 2014: 117; see discussion in Gülzow & Gagarina 2007). It does not give us sufficient information for judging the status of productivity in the child’s language system. Therefore, in addition to the analysis of frequencies of derivation patterns in CS and CDS, we retraced for each child the course of development in terms of age of emergence (see Tables 6, 7 and Section 7.3) and we identified the month in which our mini-paradigm criterion had been met (see Section 4.2).

7.1 Derived nouns

Mean values of our CDS data up to 3;0 reveal the following percentages of lemmas that are based on word formation: 50.6% of all nouns (lemmas) (11.3% suffixed nouns including diminutive suffixes, 0.5% prefixed nouns, 2.7% conversions, 1% nominalized infinitives, 35.1% compounds). Thus, 15.5% are accounted for by derivation and conversion in nouns.

Emergences

Table 6. Emergences (Em.) and age of fulfilment of the mini-paradigm criterion (MPC) of the most relevant noun derivation patterns in CS up to 3;0

Child	JAN		KAT		LEN		Example
	Em.	MPC	Em.	MPC	Em.	MPC	
-er (agent nouns)	1;8	1;10	1;11	2;10	1;9	2;3	<i>Reit-er</i> '(horseback) rider'
-er (instr. nouns)	1;8	1;10	2;3	2;10	1;9	2;3	<i>Flitz-er</i> 'speedster' <i>Last-er</i> 'truck'
-e	1;9	2;0	2;6	2;6	2;5	(2;10)*	<i>Rutsch-e</i> 'slide'
Conversion	1;8	1;10	2;8	3;0	2;4	–	<i>Sitz</i> 'seat', <i>Fall</i> 'fall'
-ung (res./loc./instr. nouns)	2;0	2;3	2;9	–	2;7	(2;11)	<i>Rett-ung</i> 'ambulance' <i>Wohn-ung</i> 'flat'

* Age of partial fulfilment (if complete fulfilment is not attested) is given in parenthesis (see Section 4.2, Example (13a) and (13b)).

As listed in Table 6, one of the first suffixes to occur up to 3;0 is the deverbal or denominal suffix *-er* forming agent and instrument nouns, with instrument nouns having higher frequencies in lemmas and tokens in both CS and CDS in all children. Obviously, the necessity of naming instruments or toys the child is playing with or using, is more important than naming agents or other persons involved, although both categories emerge at the same time (except for Kathi, whose first instrument noun emerges four months after her first agent noun at age 2;3). Clark's (1993: 118) claim that agents come before instruments, especially in children's neologisms, is not supported by our (German) data. Conversion emerges also relatively early in Jan, and a couple of months later in the two girls. Early *-ung*-nouns are mainly either very opaque (e.g. *Ahn-ung* 'clue') or refer to concrete objects (e.g. *Rett-ung* 'ambulance', *Heiz-ung* 'heating').

Productivity

In regard to the mini-paradigm criterion illustrated in Section 4.2, Example (13), the three children fulfil the criterion for *-er* at 1;10 (Jan), at 2;3 (Lena) and at 2;10 (Kathi). For the suffix *-e*, only Jan (2;0) and Kathi (2;6) meet all the necessary

requirements during the observed period, and Lena only partially by combining the suffix with three different bases at 2;10. For *-ung*, the criterion is only fulfilled by Jan at 2;3 and partially by Lena at 2;11. Productivity of conversions greatly varies among the children: Jan fulfils the criterion at 1;10, Kathi at 3;0 and Lena not at all (see Table 6).

Inflectional forms of derivatives in nouns are extremely rare in CS (3 lemmas/4 tokens) or even non-existent (in Lena's corpus). Only the dative case marker occurs shortly before the end of their third year (at 2;8 and 2;11). In two of these three cases the derivative emerges in its base form (NOM.SG) before it gets inflected.

7.2 Derived verbs

Mean values of our CDS data up to 3;0 reveal that 71.4% of all verbs (lemmas) are based on word formation (58.5% particle verbs, 6.5% prefixed and circumfixed verbs, 4.7% conversions, including root vowel modification, and 1.7% suffixed verbs).

Emergences

Table 7. Emergences and age of fulfilment of the mini-paradigm criterion of the most relevant verb derivation patterns in CS up to 3;0

Child derivation	JAN		KAT		LEN		Example
	Em.	MPC	Em.	MPC	Em.	MPC	
Particle verbs	1;6 (<i>an</i>)	1;9	2;3 (<i>ein</i>)	2;6	1;8 (<i>auf</i>)	2;1	<i>auf-heben</i> 'pick up'
Conversion	1;6	2;1	2;3	3;0	2;2	–	<i>stempeln</i> 'stamp'
Prefixed verbs	1;8 (<i>ver-</i>)	2;6	2;6 (<i>ge-</i>)	(2;8)	2;6 (<i>ge-</i>)	–	<i>ver-stecken</i> 'hide'
	1;11 (<i>ge-</i>)	2;9	2;9 (<i>be-</i>)	–	2;7 (<i>ver-</i>)	(2;7)	<i>be-kommen</i> 'get'
	2;10 (<i>be-</i>)	3;0	2;10 (<i>ver-</i>)	–	2;8 (<i>be-</i>)	–	<i>ge-fallen</i> 'like'
Suffixed verbs	1;9 (<i>-el</i>)	(1;10)	2;8 (<i>-ier</i>)	–	–	–	<i>kling-eln</i> 'to ring'
	1;11 (<i>-ier</i>)	(2;0)	2;10 (<i>-el</i>)	–	–	–	<i>prob-ier-en</i> 'try'
	2;11 (<i>-ern</i>)	(2;11)	–	–	–	–	<i>steig-ern</i> 'bid'

Since up to 3;0 particle verb formation accounts for the majority of derived verbs in CS and CDS (5.2, Table 4, Appendix), this pattern also emerges first in all children (see Table 7). A comparison of the first emergences of patterns in CS vs. the children's input shows that there exists a general tendency in CDS to offer the child a large variety of patterns even at a very young age. The only larger gap can be found in prefixed verb lemmas (see also Figures 3 and 4). Prior to the first particle-base combinations, holophrastic particles emerge independently: in Jan at 1;3 *ab* (*drehen*)

‘(turn) off’, (*r*)*auf*‘up, reach the top of sth.’ vs. the first particle verb at 1;6 *an-halt-en* (on-hold-INF) ‘hold on’, in Kathi at 1;8 *weg(reiten)* ‘(ride) away’ vs. 2;3 *ein-kauf-en* (in-buy-INF) ‘shop’ and in Lena at 1;7 *auf(machen)* ‘(open) up’ vs. 1;8 *auf-schau* (up-look.IMP.2SG) ‘look up’.

Particle verbs also emerge before conversion (though within the same month in Jan and Kathi). Both patterns are followed by prefixed verbs and, only in Jan and Kathi, by suffixed verbs. Order of emergence goes hand in hand with the lemma and token frequencies of the respective categories as presented in Table 4 (Appendix), see also the development of frequencies in Figures 7–9.

Productivity

All children fulfil the mini-paradigm criterion for particle verbs early. The first particles to become productive for Jan are *an-*, *auf-* at 1;9, thus one month earlier than nominal *-er* (see Section 7.1), Lena had her first particle verb (*auf-*) at 2;1, two months prior to derived nouns, and Kathi at 2;6 simultaneously with her deverbal noun derivations in *-e* and four months prior to her *-er* derivations.

For verb conversions (e.g. *schaukel-n* ‘swing’ JAN 1;6) only Jan (2;1) and Kathi (3;0) meet the necessary requirements up to 3;0.

For prefixed verbs the criterion is only met by Jan at 2;6 for *ver-*, 2;9 for *ge-* and 3;0 for *be-*, always having the base in three different verbs before combining the prefix with three different bases (1;8 *ver-*, 2;0 *ge-*, 2;10 *be-*).

As to verb formation via root vowel modification (e.g. *fütter-n* ‘feed’ JAN 3;0), we find the criterion fulfilled in Jan at 2;0 and Kathi at 2;8. The reason why the criterion of this comparatively infrequent pattern is fulfilled so early is that it includes verbs which are low in transparency but high in frequency: *legen* ‘lay/put’ (vs. *liegen* ‘lie’) and *setzen* ‘sit/put’ (vs. *sitzen* ‘sit’).

As for suffixation, none of the children fulfils the mini-paradigm criterion completely up to 3;0. Lena has none at all and Kathi produces too few suffixed verbs. Jan shows some productive use of the suffixes *-el* (1;10) and *-ier* (2;0) in combination with three different bases (see Section 5.2).

Derived verbs, in the example of particle verbs, usually emerge first in the infinitive (e.g. *an-halt-en* (on-hold-INF) ‘hold on’) or in combination with just the stem (e.g. *an-halt* (on-hold.STEM) ‘hold on’). Early patterns of inflection in CS are mostly past participle and imperative. By and by, inflectional markers of person and number emerge. In the course of acquisition, the use of inflected particle verbs increases while infinitives decrease, as expected in syntactic development. The much greater use of inflection in derived verbs than in derived nouns (Section 7.1) shows again the greater importance of verb derivation over noun derivation in early German child language.

7.3 Derived adjectives

Mean values of our CDS data up to 3;0 reveal that 29.7% of all adjectives in lemmas are based on word formation (17.7% suffixed adjectives, 2.6% prefixed adjectives, 3.6% compounds, 4.9% converted past participles and 0.9% present participles). Thus, 26.1% account for derivation in adjectives.

Emergences and productivity

Adjective derivation is much too limited in our data up to the age of 3;0 for allowing to present a reasonable table. In Jan's corpus, *-ig* emerges at 1;8, *-lich* at 1;9, *-isch* at 2;9 and past participles at 2;7. Jan's (3;0) neologistic **bahn-ig* 'lit. train-y for belonging to the train/railway' fits to the dominance of the denominal suffix *-ig* in CS (see Section 5.3). In both girls, the first derived adjectives emerge at 2;9: *gefähr-lich* 'dangerous' from *Gefahr* 'danger' (with umlaut in Kathi) and morphosemantically opaque *lust-ig* 'funny' from *Lust* 'lust' in Lena, the (generally less frequent) bases do not appear in their CS. Therefore, the mini-paradigm criterion has only been met by Jan for the suffix *-ig* (at 1;11) and just partially for *-lich* (at 2;5, used with three different bases). A single prefixed adjective, *un-sicht-bar* 'invisible', occurs twice in Jan's corpus.

8. Frequencies vs. age and order of emergence

For German, as for most Indo-European languages, a general noun bias is assumed for CS (cf. Tomasello 2003: 45–49). This is clearly true in Jan's preponderance of nouns over verbs, but it decreases, as expected, from period to period. However, in morphological derivation there is a clear verb bias from the beginning. This finds a parallel in inflection but not in compounding (cf. Dressler, Ketrez & Kilani-Schoch 2017), where most languages show a clear noun bias, much more than in general.

We have seen that the relative frequencies of derived noun, verb and adjective lemmas and tokens in CS are closely related to frequencies in CDS, with few exceptions in noun tokens (Kathi) and adjectives in both lemmas and tokens (Kathi, Lena), see Table 2.

Regarding input frequency as an important factor for predicting the order of emergence of a specific derivational pattern in the output, we can conclude that when comparing first emergences in Jan's CS (the corpus with the richest data) to the respective input frequencies (Tables 3, 4, Appendix, and Tables 6, 7) our data show a diversified picture within word classes and word-formation patterns (see Table 8):

Table 8. Age and order of emergence of all simplex and derived nouns, verbs (except auxiliaries and modals) and adjectives in relation to the percentages of lemmas and tokens in CDS of the whole corpus

Child	Age of emergence	Derivation pattern	CDS % lemmas	CDS % tokens
JAN	1;3	simplex nouns	41.3	69.8
	1;3	simplex adjectives	57.8	85.9
	1;4	simplex verbs	22.8	80.1
	1;6	particle verbs	63.8	15.5
	1;6	conversion (verbs)	3.1	0.6
	1;6	diminutive suffixes (nouns)	3.8	1.7
	1;8	compounds (nouns)	43.3	21.3
	1;8	-er (nouns)	2.4	1.4
	1;8	conversion (nouns)	2.7	2.0
	1;8	prefixed verbs	6.5	2.2
	1;8	root vowel modification (verbs)	1.8	0.6
	1;8	-ig (adjectives)	14.7	8.8
	1;9	-e (nouns)	2.2	1.6
	1;9	suffixed verbs	1.7	0.9
	1;9	-lich (adjectives)	2.7	2.3
	2;0	-ung (nouns)	1.0	0.6

We find that high token frequencies of simplex nouns, verbs and adjectives in the input predict early emergence in all children. As Bybee (1995: 430) has already seen, for the emergence of grammar, type frequency (in the sense of what we specify as lemma frequency) is more important than token frequency. However, in terms of derivational patterns, our data show that input frequencies, regardless of whether lemmas or tokens, are not a reliable predictor, neither for age nor order of emergence. But frequency clearly must play a role in terms of a critical mass of lemmas in CDS (cf. Marchman & Bates 1994; van der Schuit et al. 2011; Serra 2014).

In verbs, high lemma frequencies in CDS correspond to early emergence of particle verbs only (over 50%), and do not hold for conversion, which emerges in Jan and Kathi simultaneously with particle verbs (with an input of barely over 3%), and second also in Lena (lemma frequency only 1.4%). Thus, simplicity is more important than input frequency (see Section 5.2). Given the rather high percentages of prefixed verbs in lemmas (around 6%) in CDS, it would be expected that they emerge earlier than conversion, root vowel modification and suffixation. In addition to simplicity of conversion, the low prosodic and semantic salience of verb prefixes may be the cause.

For derived nouns, we find no correlation between input frequencies and age/order of emergence. Regardless of CDS frequencies, *-er* derivations are the first

ones to emerge in all children, followed by either conversion (Jan and Lena) or *-e* (Kathi), *-ung* suffixation comes last in all children. Thus, as already mentioned above, it seems that emergence of a morphological category just needs a critical mass of it in CDS, but less importantly, high (lemma) frequency.

9. Rise of complexity

Simplex words always emerge before derivational patterns (see Section 6) within each category: nouns in 1;3 vs. 1;7 and verbs in 1;4 vs. 2;0. This supports our view that early emergence of a derivative is not simply due to chance or to frequency in CDS. Each child produces a higher percentage of simplex words than their mothers, as well as more simplex than derived verbs and nouns. As expected, we observed an overall increase of derivations (verbs and nouns) between age 1;9 and 3;0 of 12.4% in lemmas and 9.74% in tokens, from which verbs constitute the major part. This fits the findings of Schipke & Kauschke (2011), see Section 3.

In regard to still greater complexity, up to age 3;0 there are barely any derivations including more than one affix (but *Ge-fäng-nis* ‘prison’, *un-sicht-bar* ‘invisible’, *hin-ge-hör-en* ‘belong’). In the vast majority of such cases, it is a particle that derives a verbal base, being unlauted in this derivational process of stem modification. This parallels our findings on the rise of complexity in compounding (Dressler et al. 2017). A particularly complex development is the rise of synthetic compounds which combine compounding and derivation (Dressler et al. 2019; for other types of complexity in general see the introductory paper by Arkadiev & Gardani 2020).

10. Conclusions and outlook

Our investigation clearly shows that children detect some frequent derivation patterns (especially particle verb derivation, nominal suffixation and conversion) quite early (around 2;0) and that before their third birthday they successfully derive verbs, nouns and, to a considerably lesser extent, adjectives. Although their repertoire is limited to a small subset of the large inventory of derivation patterns of German, the early acquired patterns cover a useful range of functions for the communicative needs of toddlers.

In terms of the acquisition of word formation, we can conclude that derived verbs (in contrast to simplex verbs vs. simplex nouns) in general play a very important role from an early age onwards and that they emerge earlier and develop faster than derived nouns (except diminutives) and especially adjectives. In our corpus, only Kathi is a late developer of both simplex and then derived verbs.

All three children fulfil our mini-paradigm criterion for particle verbs before any other word-formation pattern. This points to the importance of this derivation pattern in early first language acquisition, in addition to input frequencies. Other verb derivation patterns only play a marginal role in this early stage.

Up to the age of 3;0, the only nominal suffixes that are productive are *-er* and *-e*, with *-er* being the first suffix to become productive in child speech. Only in one child does *-ung* also show some productivity, but to an insignificant extent. Other suffixes such as *-nis*, *-schaft* or *-heit*, which derive mainly abstract nouns, occur only later, i.e. in accordance with the general precedence of concrete before abstract concepts (Schwanenflugel 1991; Gleitman et al. 2005).

There is one suffix that seems to have a special status in children's language: deverbal *-e* (e.g. canonical *Rutsch-e* 'playground slide', innovative *Lutsch-e* 'sth. to suck' for candy, *Eintauch-e* 'a dip', etc. (Mattes 2018); it is virtually unproductive in adult speech (lexicalized examples *Wieg-e* 'cradle', *Würz-e* 'seasoning' etc.), but children use it quite frequently to form instrument and action/result nouns (cf. also Elsen 1999; Rainer 2010). This might be due to the relatively high input frequency (see Table 3, Appendix) and morphotactic transparency of *-e* suffixation.

From a functional point of view, the most frequently expressed relationship that is established between verbs and nouns by derivation (for all nominal patterns) is the agentive and instrumental relation. This suggests that the main function of noun derivation in earlier stages of language acquisition is relating an action to its tool or an agent.

The only relevant derivational patterns in adjectives up to age 3;0 are denominal and deverbal suffixations in *-ig* and *-lich*, with *-ig* emerging first (except in Kathi). It clearly remains the most frequent adjective pattern during preschool age and, confirmed by diary data, it is the only one that has a clear productive status from early on. Interestingly, the suffix *-bar*, which is the most productive deverbal adjectival suffix in German (cf. Tschirner 2010: 240), only appears in Jan's corpus in two lemmas. The late acquisition of the frequent and productive suffix *-bar* might have cognitive reasons: *-bar*-adjectives have a passive mode (e.g. *ess-bar* 'eat-able' is sth. that can be eaten, not someone who eats). This means that it does not refer to observable states but to possibilities. Also, in syntax, passive voice is acquired much later than active voice. The necessary conceptualizations require cognitive developments that are clearly beyond the age under consideration in this study.

None of the early uses of derivation patterns is fully mastered at age 3;0. The whole range of affixes (including all the Latinate ones), the various (and very often abstract or idiosyncratic) meanings and the subtle semantic constraints for base-affix combinations are detected only stepwise during preschool and school age.

We conclude with a brief outlook on the course of acquisition of derivation up to the seventh year of life (for details see Rainer 2010; Mattes 2018).

For nouns, the nominal suffix *-er* remains the most frequent and productive one during preschool age. At this age, at least for some children, the suffix *-ung* becomes an important means to derive action, result and instrument nouns from verbs, especially from age 5;0 onwards, whereas other patterns (such as *-heit* or *Ge-...-e*) for abstract or collective nouns still remain very infrequent up to age 6;0.

For verbs, we find a considerable increase of frequency of prefixation and its productivity (mainly of *ver-* and *be-*) between age 3;0 and 6;0 and as a consequence a decrease of the percentage of particle verbs and conversion.¹¹

Adjective derivation remains more or less limited to the early emerged suffixes *-ig* and *-lich*. The suffix *-bar* (e.g. *ess-bar* ‘edible’) remains infrequent in CS and is rarely used in neologisms up to age 6;0.

In general, by age 6;0 only the transparent and concrete semantic categories of derivations are produced and understood. The first abstract meanings and the respective patterns (e.g. nominal *-heit* or *-schaft*) appear around the seventh birthday. As we would have expected, borrowed affixes do not play a noteworthy role for children up to age 6;0, although they sporadically use rote-learned borrowed derivatives. If affixes have both concrete and abstract meanings, concrete meanings still prevail, with the partial exception of the increase of action nouns, which are more abstract than result and instrument nouns.

Of the many open problems to be solved, we only mention more precise semantic descriptions of the development of derivational morphology and we ourselves (in various co-author variations) want to study: the acquisition of prefixoids, first of intensifying expressive adjective “compounds”, such as *sau+dumm* ‘intensely stupid, lit. sow+stupid’; and of prototypical vs. non-prototypical derivations. Further investigations of the development of word formation of Jan’s CS and CDS up to 4;11 as well as of the transversal data of 29 monolingual children and their main caretakers of the INPUT project¹² in Sabine Sommer-Lolei’s Vienna doctoral thesis (Sommer-Lolei in prep.) to be submitted soon. Moreover, we are continuing to study (various aspects of) the later development of derivational morphology up to age 6;0 and beyond (as begun by Mattes 2018, 2019), when the acquisition of literacy influences language development more and more.

Smolka and colleagues have shown in several recent publications (e.g. Smolka, Libben & Dressler 2019) that adults process both morphosemantically transparent

11. For the development of semantic knowledge of verb prefixes, see the experimental study by Mattes (2019).

12. “Investigating Parental and Other Caretakers’ Utterances to Kindergarten Children” financed by the WWTF.

and opaque German prefixed and particle verbs in the same ways, i.e. according to morphotactic structure without regard for morphosemantic transparency, which means an important contrast to studies about English and French. She and Harald Baayen are finishing a paper titled, “When stems mean more than words: Developmental trajectories in the processing of complex verbs in German speaking 11–15-year-olds” where they demonstrate that these participants show some influence of the degree of morphosemantic transparency, whereas we have claimed a larger influence of it in early childhood (e.g. Dressler, Lettner & Korecky-Kröll 2010: 338–339). It would be important to study the gradual increase of the importance of morphotactics during this development. Moreover, Smolka and Eulitz (2018) have found some differences in the adult processing of prefixed and particle verbs. Further studies on the course of acquisition of these two complex verb types could clarify whether there are any age of acquisition effects due to the earlier acquisition of particle verbs that we have described.

Acknowledgements

Sabine Sommer-Lolei is a recipient of a DOC-team fellowship of the Austrian Academy of Sciences in Vienna.

We thank the Austrian Academy of Sciences for its support, especially the Austrian Centre of Digital Humanities and Cultural Heritage (ACDH-CH) and its director Alexandra N. Lenz.

We also thank the reviewers for their valuable comments and Learose Pinkham for her corrections of English.

References

- Arkadiev, P. & Gardani, F. 2020. Introduction: Complexities in morphology. In *The Complexities of Morphology*, P. Arkadiev & F. Gardani (eds), 1–19. Oxford: OUP.
<https://doi.org/10.1093/oso/9780198861287.003.0001>
- Behrens, H. 1998. How difficult are complex verbs? Evidence from German, Dutch and English. *Linguistics* 36(4): 679–712. <https://doi.org/10.1515/ling.1998.36.4.679>
- Bittner, D. 2003. The emergence of verb inflection in two German-speaking children. In *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective* [Studies on Language Acquisition 21], D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 53–88. Berlin: De Gruyter. <https://doi.org/10.1515/9783110899832.53>
- Bybee, J. L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455. <https://doi.org/10.1080/01690969508407111>
- Bybee, J. L. 2006. *Frequency of Use and the Organization of Language*. Oxford: OUP.
- Clark, E. V. 1993. *The Lexicon in Acquisition*. [Cambridge Studies in Linguistics 65]. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511554377>

- Cordes, A.-K. 2014. *The Role of Frequency in Children's Learning of Morphological Constructions*. Tübingen: Narr.
- Dehe, N. 2015. Particle verbs. In *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 1, P. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds), 611–626. Berlin: De Gruyter.
- Dewell, R. 2015. *The Semantics of Verb Prefixes* [Human Cognitive Processing 49]. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.49>
- Dressler, W. U., Ketz, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61>
- Dressler, W. U. & Korecky-Kröll, K. 2015. Evaluative morphology and language acquisition. In *Edinburgh Handbook of Evaluative Morphology*, N. Grandi & L. Körtvélyessy (eds), 134–141. Edinburgh: Edinburgh University Press.
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2010. First language acquisition of compounds with special emphasis on early German child language. In *Cross-Disciplinary Issues in Compounding* [Current Issues in Linguistic Theory 311], S. Scalise & I. Vogel (eds), 323–344. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.311.24dre>
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2012. Acquisition of German diminutive formation and compounding in a comparative perspective: Evidence for typology and the role of frequency. In *Current Issues in Morphological Theory. (Ir)regularity, Analogy and Frequency* [Current Issues in Linguistic Theory 322], F. Kiefer, M. Ladányi & P. Siptár (eds), 237–264. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.322.11dre>
- Dressler, W. U., Sommer-Lolei, S., Korecky-Kröll, K., Argus, R., Dabašinskienė, I., Kamandulytė-Merfeldienė, L., Ijäs, J. J., Kazakovskaya, V. V., Laalo, K. & Thomadaki, E. 2019. First-language acquisition of synthetic compounds in Estonian, Finnish, German, Greek, Lithuanian, Russian and Saami. *Morphology* 29(3): 409–429. <https://doi.org/10.1007/s11525-019-09339-0>
- Duden. 1998. *Grammatik der deutschen Gegenwartssprache* (Grammar of Contemporary German). 6., neu bearbeitete Aufl. Mannheim: Dudenverlag.
- Eisenberg, P. 2013. *Das Wort. Grundriss der deutschen Grammatik* (The word. An outline of German grammar), Bd. 1. 4. aktualisierte und überarbeitete Aufl. Stuttgart: Metzler.
- Elsen, H. 1999. *Ansätze zu einer funktionalistisch-kognitiven Grammatik: Konsequenzen aus Regularitäten des Erstspracherwerbs* (Approaches to a functional-cognitive grammar: Consequences from regularities in first language acquisition). Tübingen: Niemeyer. <https://doi.org/10.1515/9783110915976>
- Elsen, H. 2011. *Grundzüge der Morphologie des Deutschen* (Main features of German morphology). Berlin: De Gruyter. <https://doi.org/10.1515/9783110237900>
- Eschenlohr, S. 1999. *Vom Nomen zum Verb: Konversion, Präfigierung und Rückbildung im Deutschen* (From noun to verb: Conversion, prefixation and back formation in German) [Germanistische Linguistik Monographien 3]. Hildesheim: Olms.
- Fleischer, W. & Barz, I. 2012. *Wortbildung der deutschen Gegenwartssprache* (Word formation in Contemporary German), 4. Aufl. Berlin: De Gruyter. <https://doi.org/10.1515/9783110256659>
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A. & Trueswell, J. C. 2005. Hard words. *Language Learning and Development* 1(1): 23–64. https://doi.org/10.1207/s15473341l1d0101_4
- Gülzow, I. & Gagarina, N. (eds) 2007. *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110977905>

- Kauschke, C. 2012. *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze* (Child language acquisition of German: Courses of development, research methods, explanatory approaches). Berlin: De Gruyter. <https://doi.org/10.1515/9783110283891>
- Kilani-Schoch, M. & Dressler, W. U. 2005. *Morphologie naturelle et flexion du verbe français* (Natural Morphology and French Verb Inflection). Tübingen: Narr.
- Klumpfer, S. 2003. Emergence of verb paradigms in one Austrian child. In *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective* [Studies on Language Acquisition 21], D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 297–322. Berlin: De Gruyter. <https://doi.org/10.1515/9783110899832.297>
- Korecky-Kröll, K. 2011. Der Erwerb der Nominalmorphologie bei zwei Wiener Kindern: Eine Untersuchung im Rahmen der Natürlichkeitstheorie (The acquisition of noun morphology in two Viennese children: an investigation within the framework of Naturalness theory). PhD dissertation, Universität Wien.
- Korecky-Kröll, K. & Dressler, W. U. 2007. Diminutives and Hypocoristics in Austrian German (AG). In *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43], I. Savickienė & W. U. Dressler (eds), 207–230. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.09kor>
- Korecky-Kröll, K. & Dressler, W. U. 2009. The acquisition of number and case in Austrian German nouns. In *Development of Nominal Inflection in First Language Acquisition: A Cross-linguistic Perspective*, U. Stephany & M. Voeikova (eds), 265–302. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110217117.265>
- Korecky-Kröll, K., Sommer-Lolei, S. & Dressler, W. U. 2017. Emergence and Early Development of German Compounds. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketz & M. Kilani-Schoch (eds), 19–37. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.02kor>
- Kühnhold, I., Putzer, O. & Wellmann, H. 1978. *Deutsche Wortbildung: Das Adjektiv* (German word formation. The adjective). Düsseldorf: Schwann-Bagel. <https://doi.org/10.1515/9783110877151>
- Kühnhold, I. & Wellmann, H. 1973. *Deutsche Wortbildung: Das Verb* (German word formation. The verb). Düsseldorf: Schwann-Bagel.
- MacWhinney, B. 2000. *CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Marchman, V. A. & Bates, E. 1994. Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language* 21: 339–366. <https://doi.org/10.1017/S0305000900009302>
- Mattes, V. 2018. Derivationsmorphologie und Wortarten im Erstspracherwerb des Deutschen. Mit einem Fokus auf dem Schuleintrittsalter (Derivational morphology and lexical categories in first language acquisition of German: focussing on the transition from preschool to school age). Habilitation thesis, Universität Graz.
- Mattes, V. 2019. What do children know about German verb prefixes? A study on the development of verb derivation from preschool age to school age. *The Mental Lexicon* 14(2): 274–297. <https://doi.org/10.1075/ml.00007.mat>
- McIntyre, A. 2015. Particle-verb formation. In *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 1, P. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds), 434–449. Berlin: De Gruyter. <https://doi.org/10.1515/9783110246254-025>

- Meibauer, J. 1999. Über Nomen-Verb-Beziehungen im frühen Wortbildungserwerb (On noun-verb-relations in early word formation acquisition). In *Das Lexikon im Spracherwerb* (The lexicon in language acquisition), J. Meibauer & M. Rothweiler (eds), 184–207. Tübingen: Francke.
- Meibauer, J., Guttropf, A. & Scherer, C. 2004. Dynamic aspects of German -er-nominals: A probe into the interrelation of language change and language acquisition. *Linguistics* 42(1): 155–193. <https://doi.org/10.1515/ling.2004.002>
- Neugebauer-Kostenblut, H. 1914. Sprachliche Eigenbildungen meines Sohnes (My son's linguistic inventions). *Zeitschrift für Kinderforschung* (Journal for child research) 19: 174–181, 242–246, 362–370.
- Plank, F. 1981. *Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie* (Morphological (Ir-)Regularities. Aspects of word structure theory). Tübingen: Narr.
- Rainer, F. 2010. *Carmens Erwerb der deutschen Wortbildung* (Carmen's acquisition of German word formation). Vienna: Verlag der Österreichischen Akademie der Wissenschaften. <https://doi.org/10.1553/oxo0237d8b>
- Savickienė, I. & Dressler, W. U. 2007. *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Schäfer, M. 2013. *Positions and interpretations: German adverbial adjectives at the syntax-semantic interface*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110278286>
- Schipke, C. & Kauschke, C. 2011. Early word formation in German language acquisition: A study on word formation growth during the second and third year. *First Language* 31(1): 67–82. <https://doi.org/10.1177/0142723709359240>
- Schulz, P., Wymann, K. & Penner, Z. 2001. The early acquisition of verb meaning in German by normally developing and language impaired children. *Brain and Language* 77: 407–418. <https://doi.org/10.1006/brln.2000.2411>
- Schwanenflugel, P. J. 1991. Why are abstract concepts hard to understand? In *The Psychology of Word Meanings*, P. J. Schwanenflugel (ed.), 223–250. Hillsdale NJ: Lawrence Erlbaum Associates.
- Serra, M. 2014. Interfaces between language components: Bootstrapping in sensitive periods after reaching a critical mass of usage. *Anuario de Psicología* 44: 85–93.
- Smolka, E. & Eulitz, C. 2018. Psycholinguistic measures for German verb pairs: Semantic transparency, semantic relatedness, verb family size, and age of reading acquisition. *Behavior Research Methods* 50(4): 1540–1562. <https://doi.org/10.3758/s13428-018-1052-5>
- Smolka, E., Libben, G. & Dressler, W. U. 2019. When morphological structure overrides meaning: Evidence from German prefix and particle verbs. *Language, Cognition, and Neuroscience* 34(5): 599–614. <https://doi.org/10.1080/23273798.2018.1552006>
- Sommer-Lolei, S. In preparation. Der frühe Erwerb der deutschen Wortbildung (Early acquisition of word formation in German). PhD dissertation, Universität Wien.
- Sommer-Lolei, S., Korecky-Kröll, K., Christiner, M. & Dressler, W. U. To appear. Acquisition and processing of (mor)phonotactic consonant clusters in German. In *Experimental, Acquisitional and Corpus Linguistic Approaches to the Study of Morphonotactics*, W. U. Dressler, B. Calderone, S. Sommer-Lolei & K. Korecky-Kröll (eds). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

- Stern, C. & Stern, W. 1928. *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung* (Child language: a psychological and language-theoretic study). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Tomasello, M. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: CUP.
- Tschirner, E. 2010. Wortschatz (Lexicon). In *Deutsch als Fremd- und Zweitsprache 1* (German as foreign and second language 1), H.-J. Krumm, C. Fandrych, B. Hufeisen & C. Riemer (eds), 236–245. Berlin: Walter de Gruyter.
- van der Schuit, M., Segers, E., van Balkom, H. & Verhoeven, L. 2011. How cognitive factors affect language development in children with intellectual disabilities. *Research in Developmental Disabilities* 32: 1884–1894. <https://doi.org/10.1016/j.ridd.2011.03.015>
- Wellmann, H. 1975. *Deutsche Wortbildung. Das Substantiv* (German word formation. The noun). Düsseldorf: Schwann-Bagel.

Appendix

Table 3. Distribution (frequencies and percentages) of derivational patterns in noun lemmas and tokens, simplex nouns and compounds among all noun lemmas and tokens in input and output up to 3;0

Derivation	CDS (J) lem/tok % lem/tok	JAN lem/tok % lem/tok	CDS (K) lem/tok % lem/tok	KAT lem/tok % lem/tok	CDS (L) lem/tok % lem/tok	LEN lem/tok % lem/tok
-er	47/224 2.4/1.4	24/152 3.0/2.2	14/38 1.9/1.4	4/7 1.3/0.6	17/56 2.0/1.3	6/22 1.9/1.2
-e	43/260 2.2/1.6	18/63 2.2/0.9	9/25 1.2/0.9	3/7 0.9/0.6	18/39 2.1/0.9	8/14 2.5/0.7
Conversion	52/310 2.7/2.0	11/124 1.4/1.8	21/58 2.9/2.1	4/6 1.3/0.5	20/44 2.3/1.0	2/13 0.6/0.7
-ung	20/98 1.0/0.6	9/22 1.1/0.3	5/9 0.7/0.3	1/2 0.3/0.2	8/44 0.9/1.0	3/9 0.9/0.5
diminutive suffixes	75/274 3.8/1.7	26/71 3.2/1.0	40/178 5.5/6.4	20/231 6.3/20.4	37/287 4.8/7.0	22/169 7.1/9.4
Compounds	848/3,374 43.3/21.3	283/1,562 35.0/22.9	226/428 31.0/15.4	77/138 24.3/12.2	242/610 31.1/14.9	49/107 15.3/5.6
Simplex	809/11,074 41.3/69.8	421/4,766 52.0/70.0	392/1,976 53.7/71.3	204/735 64.4/64.8	413/2,876 53.1/70.3	217/1,440 70.5/79.8

Table 4. Distribution (frequencies and percentages) of the most relevant derivational patterns in verb lemmas and tokens and simplex verbs among all verb lemmas and tokens (with exception of auxiliary and modal verbs) in input and output up to 3;0

Derivation	CDS (J)	JAN	CDS (K)	KAT	CDS (L)	LEN
	lem/tok	lem/tok	lem/tok	lem/tok	lem/tok	lem/tok
	% lem/tok	% lem/tok	% lem/tok	% lem/tok	% lem/tok	% lem/tok
Particles	761/4,162	253/1,060	252/938	76/162	363/1,590	115/375
	63.8/15.5	56.2/17.1	54.1/17.8	43.9/19.1	57.6/22.2	51.8/18.7
Prefixes	78/578	13/99	27/116	5/16	38/144	5/16
	6.5/2.2	2.9/1.6	5.8/2.2	2.9/1.9	6.0/2.0	2.3/0.8
Conversion	37/170	14/48	14/49	8/19	9/33	2/20
	3.1/0.6	3.1/0.8	3.0/0.9	4.6/2.2	1.4/0.5	0.9/1.0
Root mod.	22/167	8/37	13/79	4/9	13/54	2/2
	1.8/0.6	1.8/0.6	2.8/1.5	2.3/1.1	2.1/0.8	0.9/0.1
Suffixes	20/240	10/69	9/43	3/3	10/49	0/0
	1.7/0.9	2.2/1.1	1.9/0.8	1.7/0.4	1.6/0.7	0.0/0.0
Simplex	272/21,448	151/4,876	153/4,042	78/639	192/5,215	96/1,575
	22.8/80.1	33.6/78.8	32.8/76.9	45.1/75.2	30.5/72.7	43.2/78.6

Table 5. Distribution (frequencies and percentages) of the most relevant derivational patterns in adjective lemmas and tokens and simplex adjectives among all adjective lemmas and tokens in input and output up to 3;0

Affixes	CDS (J)	JAN	CDS (K)	KAT	CDS (L)	LEN
	lem/tok	lem/tok	lem/tok	lem/tok	lem/tok	lem/tok
	% lem/tok	% lem/tok	% lem/tok	% lem/tok	% lem/tok	% lem/tok
<i>-ig</i>	49/408	15/66	11/47	1/1	13/46	2/2
	14.7/8.8	12.6/6.1	9.0/6.4	2.9/0.9	9.2/5.0	4.3/1.1
<i>-lich</i>	9/107	4/44	5/7	1/1	7/9	0/0
	2.7/2.3	3.4/4.1	4.1/0.9	2.9/0.9	4.9/1.0	0.0/0.0
Past	23/45	2/6	2/6	0/0	9/9	0/0
Participle	6.9/1.0	1.7/0.6	1.6/0.8	0.0/0.0	6.3/1.0	0.0/0.0
Prefixes	14/23	1/2	0/0	0/0	5/9	0/0
	4.2/0.5	0.8/0.2	0.0/0.0	0.0/0.0	3.5/1.0	0.0/0.0
Other*	46/74	9/11	4/4	1/1	7/11	0/0
	13.8/1.6	7.4/1.1	3.3/0.5	2.9/0.9	4.9/1.1	0.0/0.0
Simplex	193/3,993	88/957	100/675	31/103	101/840	45/187
	57.8/85.9	73.9/88.1	82.0/91.3	91.2/97.2	71.1/90.9	95.7/98.9

* This category includes all other suffixes, present participles (PPres) and compounds.

Derivational morphology in Croatian child language

Gordana Hržica
University of Zagreb

Few studies have analysed derivation in languages other than English, including Croatian. This study shows that by the age of three years, Croatian children already use a high number of derived words, especially prefixed verbs. Both the number of acquired affixes and proportion of derived words increase with age. Verbs are mainly prefixed and nouns suffixed. The most prevalent semantic categories in nouns are objects and instruments, and derivation in verbs marks differences in aspect and/or Aktionsart. Later child speech shows productive usage of derivation, including derivational pairs and mini-paradigms, neologisms and neosemantism. The frequency of derived words in child-directed speech influences their frequency in child language. Compared to children, adults employ a broader repertoire of derivational patterns.

Keywords: child language, derivation, Croatian

1. Introduction

Words can have a simple or complex morphemic structure. Complex words arise through compounding (when two lexemes are combined) or derivation (when an affix is added to a lexeme). How children acquire derived and other complex words in different languages remains to be explored. As Haman (2003) explains, the acquisition of simple words means learning the relation between the word and the reality, while the acquisition of complex words means learning the relation between linguistic signs, because the meaning of a complex word is defined by its relation to a base word. She suggests that developmental psycholinguistics has neglected derivation and compounding, perhaps because the field has been dominated for so long by English, which has less productive derivational morphology than, for example, Slavic languages. This is at least true for English child-directed speech (CDS) characterized by little use of rich Latinate derivational morphology.

Both corpus-based and experimental studies have shown that children can productively use derivational patterns and compounding from an early age (e.g. Becker 2006; Berko 1958; Berman 2000; Bowerman 1982; Clark 1993, 2000, 2009), but that word formation continues to develop through late primary school age (Aitchison 1994; Anglin 1993; Berman 2003; Ravid 2004). Ravid (2004) suggests that children's ability to form words may relate to their literacy skills, since they may exhibit certain patterns of word formation only or predominantly in written language (see also Jarmulowicz 2002).

How children acquire and develop derivational morphology varies with the language. Monolingual children speaking Polish prefer derivation over compounding, while the opposite is true of monolingual children speaking English (Haman, Zevenbergen, Andrus & Chmielewska 2009). The typological characteristics of a language influence a child's acquisition of compounding (Dressler, Ketrez & Kilani-Schoch 2017).

This chapter aims to describe the early derivational patterns in Croatian, a language rich in morphology and more prone to derivation than to compounding. The goal is to assess the productive use of morphological devices to form new words.

1.1 Derivation in Croatian

Many word-formation processes have been described in the literature of Croatian (Barić et al. 2005; Kuna 2006; Silić & Pranjković 2005; overview: Tafra & Košutar 2009). The two most frequent, derivation consisting mainly of prefixation and suffixation, and compounding, are based on concatenation (Marković 2018).

There are several detailed descriptions of Croatian derivational morphology (Babić 2002; Barić et al. 2005; Silić & Pranjković 2005) that offer a comprehensive overview of affixes and their use. The comprehensive overview of Croatian word formation by Babić (2002) remains the authoritative reference on the *productivity* of affixes, which is defined by the number of words bearing a given affix from among a sample of words retrieved from a corpus, which was dictionaries and newspapers in this case. Less is known about the *frequency* of affixes in Croatian. Relevant research has focused on verbs (Kolaković & Jelaska 2009; Šojat, Srebačić & Tadić 2012), including a comprehensive overview of meanings of verb prefixes (Šojat loc. cit.). Studies of derivation in Croatian have often focused on the semantics of an individual or have been restricted to a few affixes (e.g. Budja 2008), such as in the research to assemble the Croatian WordNet (Bekavac, Šojat & Tadić 2004; Raffaelli et al. 2008).

1.1.1 Derivation of nouns

Suffixation is the most frequent type of word formation for nouns in Croatian: Babić (2002) lists 526 suffixes. The most frequent suffixes are those that name male or female persons, with female nouns typically derived from masculine nouns; or inanimate objects, such as tools and devices. Less frequent suffixes name animals, plants, or places and occur as collective nouns, gerunds or abstract nouns. Diminutives, hypocoristics¹ and augmentatives are also suffixed. Several suffixes of different productivity and frequency are used within each of these word categories (for a detailed overview, see Babić 2002). A suffix alters the meaning of a lexeme and can change its word class. For example, the suffix *-ač* can derive nouns from verbs (e.g. *pjevač* ‘singer’ from *pjevati* ‘to sing’) and from nouns (*trubač* ‘trumpeteer’ from *truba* ‘trumpet’). Each suffix can have several meanings. For example, the suffix *-ač* can derive names for persons (e.g. *pjevač* ‘singer’ from *pjevati* ‘to sing’) or inanimate objects (e.g. *otvarač* ‘opener’ from *otvoriti* ‘to open’).

Prefixation is less frequent in nouns. There are 14 productive prefixes for deriving nouns. Prefixes are mainly of prepositional origin, such as *među-* ‘between’/‘inter-’, *nad-* ‘above’, *po-* ‘below’, *pred-* ‘in front’, *protu-* ‘anti-’ and *su-* ‘co-’. Prefixes that are not prepositional denote negation (*ne-*) or ancient time (*pra-*), or they reveal the noun as untrue (*nadri-*, *nazovi-*).

1.1.2 Derivation of verbs

Derivation in Croatian verbs cannot be observed without tackling the subject of verbal aspect, which in Croatian is expressed by both prefixation and suffixation. Other types of derivation of Croatian verbs exist, such as conversion or back formation. However, they are not as prominent as prefixation or suffixation (Šojat, Srebačić & Štefanec 2013).

As in other Slavic languages, each verb is marked morphologically for perfective or imperfective aspect using a prefix or suffix. This affects the use of the verb: only imperfective verbs can be used in present tense. Both types of affixes can be applied to create pure aspectual pairs, but only suffixation is used to generate iterative forms of perfective verbs, through a process called secondary imperfectivisation (Šojat et al. 2012). There are five suffixes for imperfectivisation, and one suffix, *-nu(ti)*, that changes imperfective verbs to perfective to express punctual actions (Kocijan, Šojat & Poljak 2018). Suffixation of Croatian verbs nearly always alters the original verb’s aspect; in rare cases, it conserves the aspect but alters the original meaning, producing iterative, diminutive or pejorative verbs. New verbs can also be derived from nouns and adjectives by stem-forming suffixes.

1. A hypocoristic or pet name is a name used to show affection for a person or object.

Like other Slavic languages, Croatian features a rich system of prefixes, of which Barić et al. (2005) list 18, that make verbs perfective and may also alter verbs in additional ways. Prefixes can be added to most verbs, and each verb can usually take more than one prefix. If the prefix does not alter the original verb's meaning, the original and prefixed verbs are considered to form a primary aspectual pair (Šojat et al. 2012). Often, prefixes do change the original meaning by specifying information about how the action is performed (Aktionsart), thereby altering the verb's semantics (Martinot, Anđel & Kumar 2003).

1.2 Derivation in Croatian language acquisition

Relatively few studies have examined the acquisition of derivation in Croatian. Palmović (2007) as well as Kuvač and Palmović (2001) analysed the role of diminutives in child speech (CS), explaining how children use diminutives to avoid inflectional complexity. Pavličević-Franić and Sikirić (2005) showed how children use affixes to generate neologisms. Children coin new words by following standard Croatian morphological rules and taking affix semantics into consideration. Studies of verb prefixation have aimed to explain the acquisition of aspect and actionality (Hržica 2011; Kovačević, Kuvač & Ceganec 2005; Martinot et al. 2003). These studies have shown that typically developing children rarely err in verbal aspect; errors tend to occur among younger children and involve less frequent prefixes.

1.3 Aims of the study

This study aims to describe comprehensively the productivity and frequency of derivation in Croatian CS and CDS. The study focuses on nouns and verbs. Based on previous work on the acquisition of morphology, in which frequency (e.g. Ambridge et al. 2015), productivity (Clark 2009; Dressler, Dziubalska-Kořaczyk & Katićić 1996) and semantics (Clark & Hecht 1982) were determined as relevant factors, hypotheses were formed.

Hypotheses

- H1. As in ADS, CS and CDS reflect the duality of word-formation models in nouns and verbs, which often show typological differences: most derived nouns are produced by suffixation; most derived verbs, by prefixation.
- H2. As usual in acquisition, type and token frequency of derived words and patterns in CS reflects frequency in CDS: more frequent words and patterns in CDS are also more frequent in CS.
- H3. As found in our previous joint project publications, CDS and CS reflect affix productivity: productive affixes are more frequent and appear first in the course of language development.

- H4. As in general rise of complexity in acquisition, CS shows gradual development of derivation as an active morphological process, observable as an increase in the number of affixes and lemmas as well as in productive use of derived words and neologisms.

2. Methodology

The present study used data from the Croatian Corpus of Child Language, which consists of longitudinally collected speech samples of three monolingual children acquiring Croatian from the onset of speech to approximately three years of age and is available in the CHILDES database (Kovačević 2002; MacWhinney 2000). The speech samples were fully transcribed and morphologically coded.

Antonija was recorded approximately three times per month between the ages of 1;3 and 2;8, corresponding to a total of 10 hours of audio recordings of CS and CDS. These recordings were made at home during spontaneous interactions with family members. Marina was also recorded at home two to three times per month between the ages of 1;5 and 2;11, corresponding to a total of 21 hours of CS and CDS. Vjeran was recorded approximately three times a month, sometimes at home and sometimes at the playground, between the ages of 0;10 to 3;2, giving a total of 35 hours of CS and CDS. Further information about the corpora is presented in Table 1.

Table 1. Information about the Croatian corpus of child language

Child	Antonija	Marina	Vjeran
Age	1;3–2;8	1;5–2;11	0;10–3;2
Hours of recording	10	21	35
Number of lemmas			
in CS	1,322	1,745	3,632
in CDS	2,865	2,538	10,738
Number of tokens			
in CS	15,019	25,302	49,481
in CDS	35,899	37,984	133,533

The present study drew on transcripts from the onset of speech until 2;8 in the case of Antonija and Vjeran, or from the beginning of the recording period until 2;8 in the case of Marina. The transcripts were analysed using the *FREQ* frequency analysis programme in the Computerized Language Analysis suite *CLAN* (MacWhinney 2000). Frequency lists of nouns and verbs were compiled for all three children and the adults captured on the recordings. The adults who provided CDS on the recordings were the mother in the case of Antonija and Marina, or the mother, a nanny and an investigator in the case of Vjeran.

Two spreadsheets in Microsoft® Excel were created. The first included all lemmas, which were categorised according to whether they were nouns or verbs and whether they occurred in CS or CDS. Each lemma was marked for the type of derivation, affix and semantic features. Standard filters in Microsoft® Excel were used to analyse productivity and frequency of affixes as well as the meanings of all lemmas in CS and CDS.

The second spreadsheet consisted of complete lists of nouns and verbs in CS and CDS per month of chronological age. Each row contained the following information: child, age, lemma, type, number of appearances (tokens), and whether the lemma was derived or not. For each row containing a derived word, the type of derivation and affix were determined. Standard filters in Microsoft® Excel were used to extract the following data:

1. lemmas and tokens (nouns), altogether and only derived ones, in CS and CDS per month and child
2. lemmas and tokens (verbs), altogether and only derived ones, in CS and CDS per month and child
3. frequency of affixes in CS per month and child

Both spreadsheets were used to find derivational pairs and mini-paradigms. Possible mini-paradigms were detected by reading the lists of lemmas, then they were checked against the data for each child.

3. Derivational morphology in CS and CDS

All derived lemmas in the corpus were analysed together in order to obtain an overview of prefixation and suffixation in CS and CDS. Diminutives and hypocoristics were not analysed, since they have been examined in the data from the same three children in previous work (Kuvač & Palmović 2001; Palmović 2007).

3.1 Suffixation

3.1.1 *Nouns*

Children and their caretakers used a limited number of Croatian noun suffixes (Table 1). Out of 159 suffixes listed in Barić et al. (2005), only 19 appeared in CS. All suffixes that appeared in CS also appeared in CDS, which contained another 10. In CS, only 82 of the 584 nouns (14%) were derived with those suffixes, much lower than the rate of 32% (253/801) in CDS.

Children used lemmas with certain suffixes more frequently, with the suffixes *-a* and *-ica* the most frequent, while some suffixes appeared in only one or two

lemmas. Only five suffixes appeared in five or more lemmas in CS. Four of these five suffixes were used in both CS and CDS to name the instrument of an action. Certain suffixes were used to convey multiple meanings; the most striking example was *-a*, which was used to express five meanings (Table 2).

CDS contained a higher proportion of derived lemmas. The most frequent suffixes were the same as in CS (*-a* and *-ica*), as well as the more abstract suffix *-(a/e)nje* which was never recorded in CS. Altogether 17 suffixes appeared in at least five lemmas in CDS. The three most frequent suffixes were used to derive names for instruments or female persons and to form deverbal nouns. CS also contained nouns that named male persons, female animals, objects, places, collective nouns, abstract nouns and adjectival nouns (Table 2).

Children differed from adults in the number and proportion of suffixed lemmas, the total number of suffixes and the total number of meanings of those suffixes. Adults used a more diverse repertoire of suffixes to form deverbal nouns; in fact, one of the most frequent suffixes served exclusively to form for deverbal nouns.

Suffixes in CS and CDS were predominantly productive. Less productive or non-productive suffixes were found in only one token (e.g. *kraljev-na* 'princess' in CDS) or one lemma (e.g. *dvor-ana* 'hall').

Table 2. All noun suffixes appearing in CS and CDS

Suffix	No. lemmas CS/CDS	Meanings (CS)	Additional meanings (CDS)	Example/derived from
<i>-a</i>	15/23	– person (f., m. and f.) – animal (f.) – object – instrument – deverbal noun	– abstract nouns	<i>ograd-a</i> (fence-N.F.) 'fence' ← <i>ograd-iti</i> (fence.V-INF) 'to fence'
<i>-ac</i>	4/14	– person (m.) – instrument	– object – animal (m.)	<i>glum-ac</i> 'actor' ← <i>gluma</i> 'acting'
<i>-ač</i>	2/6	– person (m.) – instrument	=CS	<i>pjev-ač</i> 'singer' ← <i>pjevati</i> 'to sing'
<i>-ača</i>	1/2	– instrument	– place	<i>kuh-ača</i> 'cooking spoon' ← <i>kuhati</i> 'to cook'
<i>-ak</i>	5/16	– animal (m.) – place	– animal (m.) – person (m.) – object – deverbal noun	<i>krtičnj-ak</i> 'entrance to mole burrow' ← <i>krtica</i> 'mole'
<i>-ar</i>	2/8	– person (m.)	– instrument	<i>zub-ar</i> 'dentist' ← <i>zub</i> 'tooth'
<i>-aš</i>	1/2		– person (m.)	<i>nogomet-aš</i> 'soccer player' ← <i>nogomet</i> 'soccer'

(continued)

Table 2. (continued)

Suffix	No. lemmas CS/CDS	Meanings (CS)	Additional meanings (CDS)	Example/derived from
<i>-(l)ica</i>	20/43	- person (f., m. and f.) - instrument - object	- animal (f.)	<i>boj-ica</i> 'colouring pencil' ← <i>bojiti</i> 'to colour'
<i>-ić</i>	3/5	- person (m.) - place	=CS	<i>brat-ić</i> 'cousin' ← <i>brat</i> 'brother'
<i>-ina</i>	0/5		- abstract noun - adjectival noun	<i>brz-ina</i> 'speed' ← <i>brz</i> 'fast'
<i>-je/e</i>	1/7	- collective nouns	- place	<i>cvijeć-(j)e</i> 'flowers' ← <i>cvijet</i> 'flower'
<i>-(l)ište</i>	1/8	- place	=CS	<i>igra-lište</i> 'playground' ← <i>igra</i> 'play'
<i>-ka</i>	6/11	- person (f.) - instrument - object	=CS	<i>frizer-ka</i> 'female hairdresser' ← <i>frizer</i> 'hairdresser'
<i>-lo</i>	3/5	- instrument - object	- person (m.) - abstract noun	<i>šilji-lo</i> 'pencil sharpener' ← <i>šiljiti</i> 'to sharpen'
<i>-na</i>	0/2		- person (f.)	<i>kraljev-na</i> 'princess' ← <i>kraljev</i> 'king's'
<i>-(o) n(ic)a</i>	2/10	- place	- object	<i>kupa-onica</i> 'bathroom' ← <i>kupati</i> 'to take a bath'
<i>-nik</i>	2/10	- person (m.) - object	=CS	<i>pomoć-nik</i> 'helper' ← <i>pomoći</i> 'to help'
<i>-nje</i>	0/4		- deverbial noun	<i>igra-nje</i> 'playing' ← <i>igrati</i> 'to play'
<i>-o</i>	1/4	- instrument - object	=CDS	<i>zvon-o</i> 'bell' ← <i>zvon</i> 'sound'
<i>-nje</i>	4/43	- deverbial noun	- deverbial noun (activity or action)	<i>kuha-nje</i> 'cooking' ← <i>kuhati</i> 'to cook'
<i>-ost</i>	0/4		- abstract noun	<i>mudr-ost</i> 'wisdom' ← <i>mudar</i> 'wise'
<i>-stvo</i>	0/4		- abstract noun	<i>zadovolj-stvo</i> 'satisfaction' ← <i>zadovoljan</i> 'satisfied'
<i>-telj</i>	0/2		- person (m.)	<i>učit-elj</i> 'teacher' ← <i>učiti</i> 'to learn'
<i>-o</i>	5/5	- instrument - object	- deverbial noun	<i>namaz</i> 'spread' ← <i>namazati</i> 'to spread'
other	4/10	- food - day/season	various	<i>pro-lječ-(j)e</i> 'spring' ← <i>ljeto</i> 'summer'
Total	82/253			

Most suffixes and lemmas in CS were used to name objects and instruments. Like children, adults used nouns more often to name objects and instruments, but unlike children, they also frequently used nouns to name male and female persons and to form deverbal nouns denoting activity or action. Abstract nouns appeared only in CDS (Table 3).

Table 3. Semantics of derived nouns in CS and CDS

	CS	CDS
Person (male)/actor	5/2	29/14
Suffixes	<i>-ac, -ar, -ić</i>	+ <i>-ač, -ak, -ant, -aš, -ek, -er, -nik, -telj</i>
Person (female)	7	26
Suffixes	<i>-ica, -ka</i>	+ <i>-na</i>
Person (male and female)	2	3
Suffixes	<i>-a, -ica</i>	+ <i>-lo</i>
Animal (male)	2	3
Suffixes	<i>-ak</i>	+ <i>-ac</i>
Animal (female)	2	3
Suffixes	<i>-a</i>	+ <i>-ica</i>
Instrument nouns	17	24
Suffixes	<i>-a, -ac, -ač, -ača, -er, -(n)ica, -ka, -lo, -o</i>	=CS
Objects	29	39
Suffixes	<i>-a, -ica, -ik, -ka, -lo, -o</i>	+ <i>-ac, -ača, -ak, -ar, -av, -iš, -(n)ica, -nik, -taj</i>
Place	9	29
Suffixes	<i>-ak, -ić, -(l)ište, -(o)n(ic)a</i>	+ <i>-ac, -ača, -ana, -e, -nik, -nja</i>
Collective nouns	2	3
Suffixes	<i>-e/je</i>	<i>-e/je</i>
Deverbal nouns	3	59
Suffixes	<i>-a, -nje</i>	+ <i>-ak, -ba, -nja, -o</i>
Abstract nouns	–	16
		<i>-a, -ina, -lo, -(s)tvo, -ost</i>
Others	4	19
Suffixes	<i>-ada, -e, -ić, -uca</i>	+ <i>-ka, -nik, -ak</i>

Nouns in CS and CDS were derived most often from verbs (52% of nouns in CS, 60% in CDS), less often from other nouns and least often from adjectives (5% in CS and CDS). These derivations more often involved a change in word class.

3.1.2 Verbs

Children and adults produced relatively low numbers of verb suffixes (Table 4). Most suffixes were used to derive verbs from other verbs; only 3 lemmas in CS and 13 in CDS involved suffixed verbs derived from other parts of speech. To derive one verb from another, with or without change of aspect, children and adults together used 10 suffixes, including diminutive ones. Only five appeared in CS, and three of these were found in only one lemma. Beyond what children produced, adults produced one suffix for imperfectivisation and four diminutive suffixes, two of which had iterative meanings. The diminutive suffixes were limited to only one or two lemmas. Of the 419 verbs in CS, 27 (6%) were derived using suffixes, compared to 110 of 1024 verbs (11%) in CDS.

The most frequent suffix in CS and CDS was *-nu(ti)*, used to form verbs expressing punctual actions (Table 4). This suffix was followed by the suffix for imperfectivisation *-a(ti)*. These suffixes made up similar proportions of all suffixes in CS and CDS. Adults occasionally produced imperfective verbs with the suffixes *-ava(ti)* and *-iva(ti)*.

Table 4. All verb suffixes appearing in CS and CDS

Suffix	No. in lemmas CS/CDS	Role	Example	Derived from
<i>-a(ti)</i>	11/38	imperfectivisation	<i>bac-a-ti</i> 'to throw'	<i>baciti</i> 'to throw'
<i>-ava(ti)</i>	1/9	imperfectivisation	<i>otkopč-ava-ti</i> 'to unbutton'	<i>otkopčati</i> 'to unbutton'
<i>-iva(ti)</i>	0/5	imperfectivisation	<i>pokaz-iva-ti</i> 'to show'	<i>pokazati</i> 'to show'
<i>-nu(ti)</i>	13/48	perfectivisation	<i>puh-nu-ti</i> 'to blow'	<i>puhati</i> 'to blow'
<i>-ka(ti)</i>	0/4	diminutivisation	<i>lup-ka-ti</i> 'to pound repeatedly and gently'	<i>lupati</i> 'to pound'
<i>-uca(ti)</i>	1/1	diminutivisation	<i>kašlj-uca-ti</i> 'to cough a little bit'	<i>kašljati</i> 'to cough'
<i>-ulji(ti)</i>	1/1	diminutivisation	<i>smij-ulji-ti</i> 'to snicker'	<i>smijati</i> 'to laugh'
<i>-uši(ti)</i>	0/1	diminutivisation	<i>pjev-uši-ti</i> 'to sing softly'	<i>pjevati</i> 'to sing'
<i>-uta(ti)</i>	0/2	diminutivisation	<i>skak-uta-ti</i> 'to jump repeatedly in small jumps'	<i>skakati</i> 'to jump'
<i>-ucka(ti)</i>	0/1	diminutivisation	<i>greb-ucka-ti</i> 'to scratch gently'	<i>grepsti</i> 'to scratch'
Total	27/110			

The suffixes used by children and adults for perfectivisation and imperfectivisation were predominantly productive. The notable exceptions were the perfective suffix *-nu(ti)* and the imperfective suffix *-a(ti)*; the latter appeared in most imperfectivised verbs used by children and adults.

3.2 Prefixation

3.2.1 Nouns

One prefixed noun (*prez-ime* ‘family name’) appeared in CS and CDS, while the prefixes *ne-* and *po-* appeared in three lemmas in CDS.

3.2.2 Verbs

There were a large number of verb prefixes, mostly of prepositional origin. Each prefix had several meanings. For example, the prefix *u-* from the preposition *u* ‘in’ was used to mark location (‘into’) as well as a property change (e.g. *u-bit* ‘to kill’). Children and adults used the prefix *ne-* from the particle *ne-* ‘no’ differently from the way they used other prefixes. They used it with only one meaning: negation, such as in *ne-mati* ‘not to have’ from the verb *imati* ‘to have’. In addition, the prefix was never used to change verb aspect.

Based on the semantic analysis of prepositions by Šojat et al. (2012), we classified prefixed verbs in CS and CDS (Table 5). The list was extended by adding supplementary meanings for three prefixes: *po-*, *pri-* and *sa-* (Babić 2002). Based on their meaning, prefixes were classified as pure aspectual (primary aspectual pairs), location, time, quantity or manner (Šojat et al. 2012).

Children and adults produced relatively small numbers of verb prefixes. Of the 16 prefixes, 14 appeared in CS, two of which were found in only one lemma. Adults produced two additional prefixes, *nad-* and *pred-*. Of the 419 verb lemmas, 159 (38%) were derived using these prefixes in CS, compared to 539 of 1024 (53%) in CDS.

In the two groups, verb prefixation was the most prevalent type of derivation, but children used this derivation less often and in a smaller proportion of lemmas than adults. In CS and CDS, the most frequent prefix was *po-*. Prefixes were used to confer multiple meanings on verbs. In CS, most prefixes were associated with three meanings, and some prefixes with more, such as *po-* (six meanings). Eight of the 14 prefixes used by children formed aspectual pairs (Table 5). In CDS, the most frequent five prefixes (*po-*, *za-*, *na-*, *pro-*, *pre-*) were used to derive verbs with at least six different meanings. A given prefix used by children and adults showed similar distributions in the two groups.

Table 5. All verb prefixes appearing in CS and CDS

Prefix	No. lemmas CS/CDS	Meanings (CS)	Additional meanings (CDS)	Example/derived from
<i>u-</i>	11/51	– aspectual – location: into – manner: change of property	– quantity: intensity	<i>u-gurati</i> ‘to push into’ ← <i>gurati</i> ‘to push’
<i>za-</i>	18/57	– location: top-down, around – time: inchoativity – manner: change of property	– aspectual – location: to/toward, behind	<i>za-motati</i> ‘to wrap around’ ← <i>motati</i> ‘to wrap’
<i>iz-</i>	10/46	– aspectual – location: bottom-up, from – time: completion	– time: distributivity – quantity: sufficiency, excessiveness	<i>is-kopati</i> ‘to dig from’ ← <i>kopati</i> ‘to dig’
<i>o(b)-</i>	12/28	– aspectual – location: around	= CS	<i>o-brisati</i> ‘to wipe from all sides’ ← <i>brisati</i> ‘to wipe’
<i>do-</i>	8/24	– location: to/toward – time: completion	– aspectual – time: finiteness	<i>do-ći</i> ‘to arrive to’ ← <i>ići</i> ‘to go’
<i>po-</i>	42/99	– aspectual – location: top-down – time: inchoativity, distributivity, completion – quantity: intensity	= CS	<i>po-piti</i> ‘to drink up’ ← <i>piti</i> ‘to drink’
<i>na-</i>	16/48	– aspectual – time: distributivity – quantity: sufficiency	– location: top-down, to/towards – time: inchoativity – quantity: intensity excessiveness, addition	<i>na-pisati</i> ‘to finish writing’ ← <i>pisati</i> ‘to write’
<i>pre-</i>	5/25	– aspectual – location: over – time: completion	– location: re-location – time: completion – manner: change of property – quantity: intensity	<i>pre-trčati</i> ‘to run over’ ← <i>trčati</i> ‘to run’
<i>pro-</i>	9/28	– aspectual – location: through – time: completion	– location: proximity – time: inchoativity, preceding – manner: intensity	<i>pro-čitati</i> ‘to read through’ ← <i>čitati</i> ‘to read’
<i>s-</i>	16/57	– aspectual – location: top-down – manner: connection	– time: completion	<i>sa-šiti</i> ‘to sew together’ ← <i>šiti</i> ‘to sew’
<i>pri-</i>	4/17	– location: proximity	– aspectual – location: to/toward – manner: connection – quantity: intensity	<i>pri-tiskati</i> ‘to push close’ ← <i>tiskati</i> ‘to push’

Table 5. (continued)

Prefix	No. lemmas CS/CDS	Meanings (CS)	Additional meanings (CDS)	Example/derived from
<i>od-</i>	6/38	– location: apart	– aspectual – location: from – time: completion	<i>ot-kvačiti</i> ‘to unhook’ ← <i>kvačiti</i> ‘to hook’
<i>raz-</i>	1/16	– location: apart	– aspectual – quantity: intensity	<i>raz-bacati</i> ‘to throw around’ ← <i>bacati</i> ‘to throw’
<i>ne-</i>	1/1	– negation	– negation	<i>ne-mati</i> ‘not to have’ ← <i>imati</i> ‘to have’
<i>nad-</i>	0/2		– location: over	<i>nad-gledati</i> ‘watch from above/oversee’ ← <i>gledati</i> ‘to watch’
<i>pred-</i>	0/2		– time: preceding	<i>pred-vidjeti</i> ‘to foresee’ ← <i>vidjeti</i> ‘to see’
Total	159/539			

These suffixes formed aspectual pairs, as well as verbs with meanings of location, time and manner (Table 6). Most of the prefixed verbs in CS were primary aspectual pairs, followed by verbs of location or time. Similarly, more prefixed verbs in CDS expressed location and time than quantity or manner, but the proportion of prefixed verbs serving only as aspectual pairs was lower than in CS.

Table 6. Semantics of derived (prefixed) verbs

	CS	CDS
Aspectual	64	151
Suffixes	<i>iz-, na-, o-, po-, pre-, pro-, s-, u-</i>	+ <i>do-, od-, pri-, raz-, za-</i>
Location	44	162
Suffixes	<i>do-, iz-, o-, od-, po-, pre-, pri-, pro-, raz-, s-, u-, za-</i>	+ <i>na-, nad-</i>
Time	30	125
Suffixes	<i>do-, iz-, na-, po-, pre-, pro-, s-, za-</i>	+ <i>od-, pred-</i>
Quantity	6	40
Suffixes	<i>na-, po-</i>	+ <i>iz-, pre-, pri-, pro-, raz-, u-</i>
Manner	15	61
Suffixes	<i>s-, u-, za-</i>	+ <i>pre-, pri-</i>

Nearly all prefixes in CS and CDS were productive, as are most prefixes in the Croatian language (Babić 2002). The sole exception was the prefix *ne-* ‘no’, whose productivity was low, and which was used in a small number of frequent lemmas.

4. Frequency and acquisition of derived words in CS

In this section, the data from the three children were separated, in order to study possible individual differences. To see the prevalence of derived words, types of derivation, and the timeframe of their appearance, both lemmas and tokens in each individual corpus were analysed.

4.1 Frequencies of derived nouns and verbs in CS and CDS

The percentage of derived nouns (lemmas) in CS was high, reaching more than 55% (Table 7). However, most of the derived nouns in CS and CDS were diminutives. After diminutives were discarded from the analysis, the percentage of derived nouns fell to 10% for CS and 14% for CDS. In contrast, the percentage of derived verbs (lemmas) was high in CS (50%) and CDS (56%).

For nouns and verbs, the percentage of derived lemmas was higher than the percentage of derived tokens for children and adults (Table 7). In other words, both groups have available a broader inventory of derived nouns and verbs than they actually use. Nearly half of all verb lemmas in CS and CDS were derived, compared to only up to 22% of verb tokens. The percentages of lemmas and tokens were clearly higher among adults than children.

Table 7. Frequencies of lemmas and tokens of derived nouns and verbs

	CDS Antonija	CS Antonija	CDS Marina	CS Marina	CDS Vjeran	CS Vjeran
N lemmas	620	326	610	407	1479	439
N tokens	3722	1284	4335	2469	13081	1942
N DERIV lemmas	267	141	244	162	545	158
(excl. DIM)	(72)	(29)	(68)	(42)	(210)	(37)
N DERIV tokens	1468	712	1821	1312	3704	730
(excl. DIM)	(229)*	(126)*	(281)*	(155)*	(973)*	(100)*
% N DERIV lemmas	43%	43%	40%	40%	37%	36%
(excl. DIM)	(12%)	(9%)	(11%)	(10%)	(14%)	(8%)
% N DERIV tokens	39%	55%	42%	53%	28%	38%
(excl. DIM)	(6%)*	(9%)*	(6%)*	(6%)*	(7%)*	(5%)*
V lemmas	436	200	454	242	1088	250
V tokens	6099	2142	6745	3581	21563	2029
V DERIV lemmas	212	77	214	95	612	108
V DERIV tokens	1134	304	1069	445	4825	332
% V DERIV lemmas	49%	38%	47%	39%	56%	43%
% V DERIV tokens	19%	14%	16%	12%	22%	16%

* Numbers outside the brackets take into account all derived nouns, while numbers inside the brackets exclude diminutives

4.2 Frequencies of different derivational types in CS and CDS

Most derived nouns in CS were produced through suffixation; only up to 5% were prefixed. Conversely, most derived verbs were prefixed verbs, with only up to 12% suffixed (Table 8). Similar results were observed in CDS.

Table 8. Frequencies of types of noun and verb derivation in lemmas and tokens

	CDS Antonija	CS Antonija	CDS Marina	CS Marina	CDS Vjeran	CS Vjeran
N DERIV lemmas	72	29	68	42	210	37
N DERIV tokens	229	126	281	155	973	100
N SUFF lemmas	65	24	60	38	183	33
N SUFF tokens	216	114	258	134	917	95
% N SUFF lemmas	90%	83%	88%	91%	87%	90%
% N SUFF tokens	94%	91%	92%	86%	94%	95%
N PREF lemmas	0	0	0	0	4	0
N PREF tokens	0	0	0	0	20	0
% N PREF lemmas	0%	0%	0%	0%	2%	0%
% N PREF tokens	0%	0%	0%	0%	2%	0%
N PREF SUFF lemmas	4	2	3	1	20	2
N PREF SUFF tokens	7	8	12	14	35	2
% N PREF SUFF lemmas	6%	7%	5%	2%	10%	5%
% N PREF SUFF tokens	3%	6%	4%	9%	3%	2%
N other lemmas	3	3	5	3	3	2
N other tokens	6	4	11	7	6	3
% N other lemmas	4%	10%	7%	7%	1%	5%
% N other tokens	3%	3%	4%	5%	1%	3%
V DERIV lemmas	212	77	214	95	612	108
V DERIV tokens	1134	304	1069	445	4825	332
V SUFF lemmas	32	11	40	16	112	11
V SUFF tokens	88	41	137	97	475	31
% V SUFF lemmas	15%	14%	19%	17%	18%	10%
% V SUFF tokens	8%	13%	13%	22%	10%	10%
V PREF lemmas	180	66	174	79	500	97
V PREF tokens	1046	163	932	348	4350	301
% V PREF lemmas	85%	86%	81%	83%	82%	90%
% V PREF tokens	92%	87%	87%	78%	90%	90%

4.3 Acquisition: Appearance and development of derived words

Derived words appeared early in CS, at 1;5 for Antonija and Marina and at 1;6 for Vjeran. Derived verbs appeared simultaneously with, or shortly after, derived nouns. The first derived form to appear in all three children was the verb *ne-mati* 'not to have'. Around the same time, Marina produced several derived nouns, such as the neologism *na-rezalj-ka* 'instrument for cutting'. Months after producing their first derived verbs, Antonija and Vjeran produced their first derived nouns: *glum-ac* 'actor' in the case of Vjeran, *na-oča-le* 'glasses' in the case of Antonija.

Over time, the numbers of derived words and their frequency in lemmas increased. The proportion of derived lemmas increased from 10% at the beginning to more than 25% toward the end (Table 9). Spearman correlation analysis showed a strong positive correlation of month of chronological age with the total number of derived words ($r_s = .89, p < .05$) and the percentage of words that were derived ($r_s = .97, p < .05$).

Table 9. Percentages of derived words (lemmas) in CS, by month in chronological order

	Antonija		Marina		Vjeran	
	DERIV lemmas	% DERIV lemmas	DERIV lemmas	% DERIV lemmas	DERIV lemmas	% DERIV lemmas
1;5	8/1	13%	58/6	10%	6/0	0%
1;6	17/2	12%	114/14	12%	29/2	7%
1;7	61/6	10%	128/16	13%	91/10	11%
1;8	61/6	10%	145/20	14%	124/12	10%
1;9	100/14	14%	157/21	13%	157/18	11%
1;10	101/14	14%	169/22	13%	274/24	9%
1;11	141/23	16%	204/28	14%	320/40	13%
2;0	175/29	17%	228/32	14%	340/44	13%
2;1	209/39	19%	254/36	14%	381/55	14%
2;2	254/52	20%	275/39	14%	399/60	15%
2;3	298/73	24%	287/45	16%	435/68	16%
2;4	325/80	25%	287/45	16%	476/81	17%
2;5	358/92	26%	376/75	20%	523/89	18%
2;6	375/96	26%	447/112	25%	546/101	20%
2;7	403/102	25%	488/123	25%	569/123	22%
2;8	414/106	26%	529/137	26%	583/145	25%

Also increasing over time was the number of suffixes used to derive words (Figure 1). Children initially produced words with few affixes, but this number grew steadily until 2;5. Spearman correlation analysis showed strong positive correlation of month of chronological age with number of affixes ($r_s = .95, p < .05$) as well as the number of derived lemmas and affixes ($r_s = .84, p < .05$).

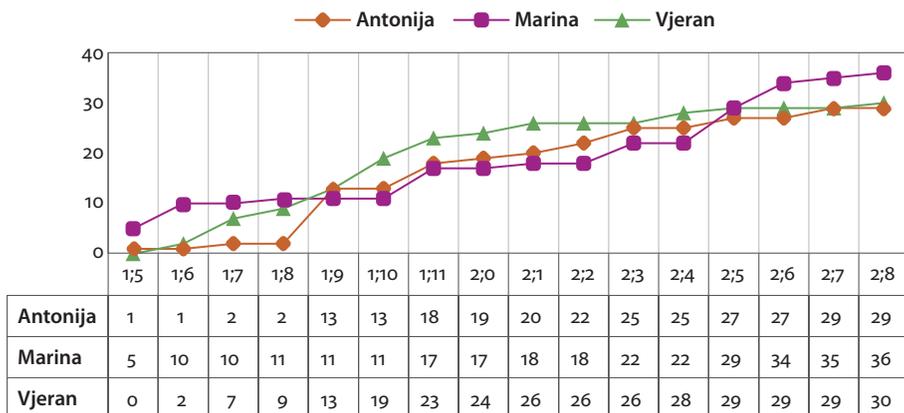


Figure 1. Number of affixes in CS

4.4 Acquisition: Derivational pairs and mini-paradigms

In order to analyse when the children began to detect patterns of derivational morphology, we identified derivational pairs and mini-paradigms from the CS of each child. Derivational pairs consist of two words sharing the same root, while mini-paradigms consist of at least three words that share the same root and have a recurring affix. Derivational pairs produced by all three children consisted of two nouns, two verbs or a noun and a verb (1). Mini-paradigms consisted nearly entirely of verbs, nouns or their combination, with adjectives rarely involved (2).

- (1) a. Derivational pair (verbs)
 - i. Marina: *tući* ‘to hit someone’ (2;2) – *po-tući* ‘to pick a fight’ (2;5)
 - ii. Antonija: *o-tvoriti* ‘to open’ (1;9) – *za-tvoriti* ‘to close’ (2;1)
 - iii. Vjeran: *iz-laziti* ‘to come out’ (2;6) – *pro-laziti* ‘to pass by’ (2;6)
- b. Derivational pair (nouns)
 - i. Marina: *mač-ka* (cat-F) ‘cat’ (2;1) – *mač-ak* (cat-M) ‘cat’ (2;6)
 - ii. Antonija: *dim* ‘smoke’ (2;6) – *dimnj-ak* ‘chimney’ (2;4)
 - iii. Vjeran: *ruka* ‘hand’ (1;9) – *rukav-ica* ‘glove’ (1;9)
- c. Derivational pair (nouns and verbs)
 - i. Marina: *ljuljati* ‘to swing’ (1;5) – *ljuljač-ka* ‘swing’ (2;0)
 - ii. Vjeran: *broj* ‘number’ (2;0) – *broj-a-ti* ‘to count’ (2;2)
- (2) a. Derivational mini-paradigms (verbs)
 - i. Marina: *praviti* ‘to make something’ (2;1) – *na-praviti* (1;7) ‘to finish something’ – *po-praviti* ‘to fix/mend something’ (2;5)
(recurring prefixes: *na-*, as in *na-crtati* ‘to finish drawing’ or *na-točiti* ‘to complete pouring’; *po-*, as in *po-brisati* ‘to wipe off’ or *po-gledati* ‘to finish watching’)

- ii. Antonija: *zvati* 'to call' (2;0) – *po-zvati* 'to call upon' (2;6) – *na-zvati* 'to call someone on the telephone' (1;11)
(recurring prefixes: *na-*, as in *na-crtati* 'to finish drawing' or *na-pisati* 'to finish writing'; *po-*, as in *po-jesti* 'to finish eating' or *po-gledati* 'to finish watching')
 - iii. Vjeran: *kazati* 'to point' (1;9) – *po-kazati* 'to point at' (1;10) – *pri-kazati* 'to display' (2;0)
(recurring prefixes: *po-*, as in *po-goditi* 'to hit (target)' or *po-gaziti* 'to stamp over'; *pri-*, as in *pri-tiskati* 'to push closed' or *pri-premiti* 'to prepare')
- b. Derivational mini-paradigms (nouns)
- i. Antonija: *ruka* 'hand' (2;3) – *rukav-ica* 'glove' (2;5) – *na-rukv-ica* 'bracelet' (2;7)
(recurring suffix: *-ica*, as in *kvač-ica* 'laundry clip' or *boj-ica* 'colouring pen')
 - ii. Antonija: *krt* 'mole' (1;11) – *krt-ica* 'mole' (1;11) (fem.) – *krtičnj-ak* 'mole burrow' (1;11)
(recurring suffixes: *-ica*, as in *prijatelj-ica* 'female friend' or *doktor-ica* 'female doctor'; *-ak*, as in *dimnj-ak* 'chimney')
 - iii. Vjeran: *miš* 'mouse' (1;9) – *miš-ica* 'mouse (fem.)' (2;3) – *miš-ić* 'little mouse' (pup) (2;4)
(recurring suffixes: *-ica*, as in *prijatelj-ica* 'female friend' or *doktor-ica* 'female doctor'; *-ić*, as in *ps-ić* 'puppy' or *brat-ić* 'cousin')
- c. Derivational mini-paradigms (nouns and verbs)
- i. Marina: *kuhati* 'to cook' (1;6) – *s-kuhati* 'to finish cooking' (1;11) – *kuha-nje* 'cooking' (2;6) – *kuh-ača* 'wooden spoon' (2;6)
(recurring affixes: *s-*, as in *s-trgati* 'to tear apart' or *s-močiti* 'to finish wetting'; *-nje*, as in *pliva-nje* 'swimming' or *staja-nje* 'standing')
 - ii. Antonija: *bojiti* 'to colour' (1;11) – *boj-ica* 'colouring pen' (2;3) – *po-boj-a-ti* 'to colour something fully' (2;3)
(recurring affixes: *-ica*, as in *kvač-ica* 'laundry clip' or *gum-ica* 'rubber eraser'; *po-*, as in *po-jesti* 'to finish eating' or *po-gledati* 'to finish watching')
 - iii. Vjeran: *igrati* 'to play' (1;7) – *po-igrati* 'to play' (iterative) (2;6) – *igrač-ka* 'toy' (2;6) – *igra* 'game' (2;8)
(recurring affixes: *po-*, as in *po-jesti* 'to finish eating' or *po-gledati* 'to finish watching'; *-ka*, as in *ljuljač-ka* 'swing' or *trenir-ka* 'track suit')

Children produced many derivational pairs but fewer mini-paradigms, as expected. More pairs and mini-paradigms involved verbs than nouns. All three children

productively used suffixes and prefixes in different words. Prefixes and suffixes were sometimes used to alter aspect and occasionally to alter both aspect and semantics (Tables 10 and 11).

Table 10. Numbers and types of derivational pairs

Type of derivational pair	CS	CS	CS
	Antonija	Marina	Vjeran
Nouns	14	24	13
Verbs	29 (18)*	40 (20)*	30 (18)*
Combination of verbs and nouns	18	11	15
Combination of adjectives and nouns	6	6	8
Combination of adjectives and verbs	5	3	6
Total	72 (61)	84 (64)	72 (60)

* Counts included primary aspectual pairs (outside brackets) or excluded them (inside brackets).

Table 11. Numbers and types of derivational mini-paradigms

Type of derivational mini-paradigm	CS	CS	CS
	Antonija	Marina	Vjeran
Nouns	3	2	5
Verbs	7	11	12
Combination of verbs and nouns	4	5	11
Combination of adjectives and nouns	–	–	4
Combination of adjectives and verbs	1	–	2
Combination of adjectives, nouns and verbs	2	1	1
Total	17	19	35

4.5 Acquisition: Neologisms and neosemanticism

All three children produced some derived neologisms or neosemanticism,² always in accordance with the rules of Croatian morphology (Table 12). While some words were lexical innovations that do not exist in the standard lexicon, sometimes children simply added a non-conventional prefix to an existing word, preferring more frequent prefixes that expanded the original semantics.

2. The assignment of a new meaning to an existing word.

Table 12. Neologisms and neosemantism in CS

Child	Neologism/ neosemantism	Meaning	Observed in adult language	Possible reason
Antonija	<i>po-davati</i> (instead of <i>do-davati</i>)	'to pass' (the ball)	yes (different meaning)	prefix <i>po-</i> is more frequent than <i>do-</i>
Antonija	<i>reket-a-ti</i>	'to play a game with a racket'	no	lexical innovation
Marina	<i>na-rezalj-ka</i>	'a thing to cut with'	no	lexical innovation
Marina	<i>na-paliti</i> (instead of <i>u-paliti</i>)	'to turn on' (the TV)	yes (different meaning)	prefix <i>na-</i> is more frequent than <i>u-</i>
Vjeran	<i>po-plašiti</i> (instead of <i>u-plašiti</i>)	'to scare someone'	yes (different meaning)	prefix <i>po-</i> is more frequent than <i>u-</i>

5. Discussion

5.1 Derivational patterns in CS and CDS

Acquisition of inflectional morphology and of language more generally can be analysed in terms of frequency (e.g. Ambridge et al. 2015), productivity (Clark 2009; Dressler et al. 1996) and semantics (Clark & Hecht 1982). Therefore, we analysed these factors in the present study.

5.1.1 *Derivational types*

In CS and CDS, derived nouns were less frequent than derived verbs, after excluding diminutives and hypocoristics. Consistent with Hypothesis 1, derived nouns were produced predominantly by suffixation, but derived verbs predominantly by prefixation, reflecting the rich system of verbal prefixation in Slavic languages.

The distribution of different derivational types in CS and CDS of the present study is largely consistent with what has been described in Croatian adult language. Our results show that most suffixed verbs in CS and CDS were derived using one suffix for imperfectivisation or one for perfectivisation. The perfective suffix may add the meaning of punctuality and/or diminutivisation, as described for Croatian in Katunar (2013). However, two deviations were observed between the present sample and samples in other studies. In Croatian, perfective verbs are typically derived by prefixation, and imperfectives by suffixation (e.g. Kocijan et al. 2018). The morphological database of Croatian verbs CroDeriv (Šojat et al. 2012, 2013) indicates greater numbers of lemmas containing verbs with the imperfective suffixes *-ava(ti)* (more than 1000) or *-iva(ti)* (more than 1500) than the perfective suffix *-nu(ti)* (more than 800). However, the opposite distribution was found in

the present analysis of CS and CDS. This discrepancy deserves attention in future work: is the present result limited to CS and CDS, or is it a characteristic of spoken language in general and why?

The second deviation was that prefixation of nouns, which was not observed in the present sample of children and adults, does occur in Croatian, albeit less often.

5.1.2 *Productivity*

Productivity may be an important determinant of the acquisition of Croatian inflectional morphology (Dressler et al. 1996; Hržica 2012), so it may affect the emergence of derived words and their patterns in CS. Most derived nouns in CS were suffixed, and all those suffixes were productive. However, one of the most frequent suffixes in CS was *-a*, which is less productive (Barić et al. 2005). In addition, less productive suffixes such as *-ac* (for male individuals) outnumbered highly productive suffixes such as *-ak* in some semantic categories. The two most prevalent suffixes in the suffixation of verbs, *-ati* and *-nuti*, are non-productive (Babić 2002).

It is not easy to assess the role of productivity in prefixed verbs, since most verbal prefixes in Croatian are productive (Babić 2002; Barić et al. 2005). The only prefix that appeared in CS was *ne-*, which is not productive and is used to negate verbs (*ne-mati* 'not to have'). This prefix appeared in the CS and CDS of all three children, but in only one verb.

These results, together with the observation that the derived words that appeared first in CS contained less productive or non-productive affixes, suggest that productivity plays a different role in the acquisition of derivation than in inflectional morphology in the Croatian language. Overall, the present results do not support Hypothesis 2, which predicted that productive affixes would be more frequent in CDS and would appear first during language development. However, it should be taken into consideration that other types of analyses, such as observing affixes appearing in mini-paradigms, might yield different conclusions about the role of productivity in the acquisition of derivation (the most frequent affixes in mini-paradigms are productive: *-ica*, *-ak* and *-ka* for nouns and *-po* and *-na* for verbs).

5.1.3 *Frequency*

Analysis of derivational morphemes in CDS and CS showed that the children in the present study produced suffixed nouns in a similar manner as adults: the five suffixes most frequent in CS were also most frequent in CDS. The only notable difference was that the suffixes used to derive deverbal nouns denoting activity or action were much more frequent in CDS. Similarly, the relative frequencies of verb prefixes were similar between CS and CDS. Children in the present study rarely erred in the usage of prefixes and produced neologisms, replacing less frequent

prefixes with more frequent *po-* and *na-*. This is consistent with the “cause error thesis” of Ambridge et al. (2015), according to which children err by selecting a more frequent but incorrect prefix instead of the correct but less frequent prefix in a given context.

All the prefixes used by children in the present study are among the most frequent ones on two lists of prefixed verb lemmas in Croatian (Kolaković & Jelaska 2009; Šojat et al. 2012). However, the children and adults in the present study used most often *po-*, which occupies third or fourth position on those lists. In Croatian, the prefix *po-* is often used to form primary aspectual pairs and it is the most frequent prefix in tokens (Matovac 2018). In some Slavic languages, *po-* primarily has a function of perfectivisation and has lost other meanings (Dickey 2011, 2012), but these other meanings have been better preserved in Croatian (Matovac 2018). Our results are consistent with the token frequency effect (Ambridge et al. 2015), which predicts that children are more likely to acquire words that they hear more often.

The present study found that most frequent affixes in CS are also the most frequent affixes in CDS. These results support the type frequency effect (Ambridge et al. 2015), which predicts that children are more likely to use morphemes that they hear in multiple words in CDS than morphemes that occur in only one word.

The present study confirms Hypothesis 3 and the pervasive effects of frequency in child language acquisition previously observed in experimental studies (e.g. Endress & Hauser 2011) and longitudinal corpus analysis (Behrens 2003) of derivation patterns. The results here suggest that the frequency effects described by Ambridge et al. (2015) based on single words, inflectional morphology and syntactic constructions apply also to the acquisition of derivation.

5.1.4 *Semantics*

Children in the present study used derived nouns to name objects and instruments most often and agents only rarely. This contrasts with previous studies that suggested that agent nouns are acquired earlier (Clark & Hecht 1982). These differences may reflect the larger number of polysemic suffixes in Croatian than in English, and the fact that the present study analysed children in their natural environment rather than in an artificial experimental setting as in the previous work. In the present study, adults derived nouns naming persons more often than children did. This difference may mean that younger children prefer to talk about instruments and objects than agents and therefore acquire the relevant nouns earlier.

The children in the present study correctly used a range of prefixes to produce prefixed verbs with diverse meanings. Their low rate of errors is consistent with previous research that found aspectual differences emerging first during children’s spontaneous production of prefixed verbs in Croatian (Hržica 2011). Indeed, children predominantly produced prefixed verbs that marked only aspect, whereas

adults most often produced verbs that marked not only aspect but also location or temporality. Future studies should examine whether prefixation in CS serves mainly to express primary aspectual differences, while other meanings of prefixes are acquired later, perhaps in phases similar to those during the acquisition of prepositions (for an overview, see Rivière, Lécuyer & Hickmann 2009).

5.2 Derivation in development – evidence of productivity

Early experimental studies of derivation in child language (Berko 1958) showed that children acquire derivational patterns relatively late, up to 10 or 12 years of age. In contrast to these findings, early studies of spontaneous speech of children speaking English showed that agent nouns constitute a prominent category, even in three-year-olds (Clark 1981). In parallel, children produce many novel forms, including compounds. This raises the question of the extent to which children productively derive words or simply apply memorised words. Previous research provided evidence that children productively derive diminutives in Croatian (Kuvač & Palmović 2001; Palmović 2007). The present study examined other aspects of derivational morphology to detect possible indicators of productivity.

5.2.1 *Derived words in CS*

All three children in this study produced high percentages of derived words. This was especially true for verb lemmas, which accounted for around 40% of all verbs. Derived nouns accounted for only 10% of all nouns after excluding diminutives and hypocoristics. The corresponding percentages in CDS were only slightly higher. Both the number and percentage of derived words increased with the chronological age of the children.

5.2.2 *Affixes and lemmas in CS*

The number of affixes grew gradually and steadily with the age of the children in this study. In addition, the number of affixes correlated with the number of derived lemmas. Thus, with time, the children used more derived forms and more derivational patterns (see Figure 1).

The children and adults in our study used nearly the same repertoire of derivational affixes for verbs. However, children used a much narrower range of affixes for nouns. This suggests that specifically during early language development, children may apply a range of prefixes to a range of verbs. Indeed, half of the prefixes used by children in the present study appeared in at least 10 lemmas. The appearance of the same prefix in such large groups of verbs implies productive usage of derivational patterns.

5.2.3 *Spontaneous usage of derived forms*

Mini-paradigms have also been interpreted as indicators of productivity in children's use of different inflectional forms in various languages (e.g. Dressler, Kilani-Schoch & Klampfer 2003; Gagarina 2003; Kovačević, Palmović & Hržica 2009). Along these lines, the present study analysed derivational mini-paradigms as an indicator of children's active use of derivational morphology.

Earlier in language development children in this study used words within a single derivational family to form large numbers of derivational pairs, and later they also produced some derivational mini-paradigms. Although many mini-paradigms contained nouns, the mini-paradigms with the highest number of members (five) included either five verbs or four verbs and one noun. This is in accordance with the prevalence of verbs within the overall number of derived words and of verb affixes in CS in the present study. The children produced large groups of verbs with the same root but different prefixes that slightly altered the meaning. The lower number of nouns in mini-paradigms reflects both the lower percentage of derived nouns and the lower number of suffixes used by the children than the adults in the present study.

5.2.4 *Neologisms*

When using affixes, the children in the present study sometimes produced new forms that were not observed in adult speech, consistent with Hypothesis 4. Such new words may be cases when children rely on their knowledge of morphology to create words (e.g. Becker 2006; Clark 2009). Such neologisms are likely rare as in the present study, though they may occur more often than can be detected in studies in which recordings are up to three times per month (Rowland, Fletcher & Freudenthal 2008).

6. Conclusion

The qualitative and quantitative analysis of changes in the early lexicon of three Croatian children revealed numerous derivational patterns and provided evidence that they use derivational morphology productively. The frequency of derived words in CDS had a strong impact on the acquisition of derivational morphology. Nevertheless, the derivational productivity in adult speech did not necessarily influence the development of derivational patterns in children. Comparing the Croatian children's derivational behaviour to that of the adults shows that they have a long way to go until they master the complex derivational morphology of the language.

Funding

This work has been supported in part by the Croatian Science Foundation under the project Multilevel approach to spoken discourse in language development (UIP-2017-05-6603) and by the project Computer assistant for text input for persons with language impairment (RAPUT, RC.2.2.08 – 0050), co-funded by European Structural and Investment Funds.

References

- Aitchison, J. 1994. *Words in the Mind. An Introduction to the Mental Lexicon*. Oxford: Blackwell.
- Ambridge, B., Kidd, E., Rowland, C. & Theakston, A. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2): 239–273.
<https://doi.org/10.1017/S030500091400049X>
- Anglin, J. M. 1993. Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development* 58(10): 1–16. <https://doi.org/10.2307/1166112>
- Babić, S. 2002. *Tvorba riječi u hrvatskom književnom jeziku* (Derivation in literary Croatian language). Zagreb: HAZU i Globus.
- Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V. & Znika, M. 2005. *Hrvatska gramatika* (Croatian grammar). Zagreb: Školska knjiga.
- Becker, J. 2006. A child's invention of new words. In *Child Language: A Resource Book for Students*, J. S. Peccie (ed.), 134–147. London: Routledge.
- Behrens, H. 2003. Verbal prefixation in German child and adult language. *Acta Linguistica Hungarica* 50(1–2): 37–55. <https://doi.org/10.1556/ALing.50.2003.1-2.4>
- Bekavac, B., Šojat, K. & Tadić, M. 2004. Zašto nam treba hrvatski WordNet? (Why do we need Croatian WordNet?). In *Semantika prirodnog jezika i metajezik semantike* (Semantics of natural language and semantic metalanguage), *Proceedings of HDPL Annual Conference*, J. Granić (ed.), 733–743. Zagreb-Split: HDPL.
- Berko, J. 1958. The child's learning of English morphology. *Word* 14(2–3): 150–177.
<https://doi.org/10.1080/00437956.1958.11659661>
- Berman, R. A. 2000. Children's innovative verbs versus nouns: Structured elicitations and spontaneous coinages. In *Methods for Studying Language Production*, L. Menn & N. Bernstein-Ratner (eds), 69–92, Hillsdale NJ: Lawrence Erlbaum Associates.
- Berman, R. A. 2003. Children's lexical innovations: Developmental perspectives on Hebrew verb structure. In *Language Processing and Acquisition in Language of Semitic, Root-based Morphology* [Language Acquisition and Language Disorders 28], J. Shimron (ed.), 243–292. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.28.13ber>
- Bowerman, M. 1982. Reorganization processes in lexical and syntactic development. In *Language Acquisition: The State of Art*, E. Wanner & L. Gleitman (eds), 319–346. Cambridge: CUP.
- Budja, J. 2008. Značenja glagolskoga prefiksa za- u hrvatskom (Meanings of the verbal prefix za- in Croatian). PhD dissertation. University of Zagreb.
- Clark, E. V. 1981. Lexical innovations: How children learn to create new words. In *The Child's Construction of Language*, W. Deutsch (ed.), 299–328. London: Academic Press.
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511554377>

- Clark, E. V. 2000. Coining new words: Old and new word forms for new meanings. In *Methods for Studying Language Production*, L. Menn & N. Bernstein-Ratner (eds), 53–67. Mahwah NJ: Lawrence Erlbaum Associates.
- Clark, E. V. 2009. *First Language Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511806698>
- Clark, E. V. & Hecht, B. F. 1982. Learning to coin agent and instrument nouns. *Cognition* 12: 1–24.
[https://doi.org/10.1016/0010-0277\(82\)90027-0](https://doi.org/10.1016/0010-0277(82)90027-0)
- Dickey, S. M. 2011. The varying role of *po-* in the grammaticalization of Slavic aspectual systems: Sequences of events, delimitatives, and German language contact. *Journal of Slavic Linguistics* 19(2): 175–230. <https://doi.org/10.1353/jsl.2011.0009>
- Dickey, S. M. 2012. Orphan prefixes and the grammaticalization of aspect in South Slavic. *Jezikoslovlje* 13(1): 71–105.
- Dressler, W. U., Dziubalska-Kołodziejczyk, K. D. & Katičić, A. 1996. A contrastive analysis of verbal inflection classes in Polish and Croatian. *Suvremena Lingvistika* 41–42(1–2): 127–138.
- Dressler, W. U., Ketz, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins.
<https://doi.org/10.1075/lald.61>
- Dressler, W. U., Kilani-Schoch, M. & Klampfer, S. 2003. How does a child detect morphology? Evidence from production. In *Trends in Linguistics*, W. Bisang, H. H. Hock & W. Winter (eds), 391–425. Berlin: Mouton de Gruyter.
- Endress, A. D. & Hauser, M. D. 2011. The influence of type and token frequency on the acquisition of affixation patterns: Implications for language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(1): 77–95.
- Gagarina, N. 2003. The early verb development and demarcation of stages in three Russian-speaking children. In *Development of Verb Inflection in First Language Acquisition. A Cross-Linguistic Perspective*, D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 131–169. Berlin: De Gruyter. <https://doi.org/10.1515/9783110899832.131>
- Haman, E. 2003. Early productivity in derivation. A case study of diminutives in the acquisition of Polish. *Psychology of Language and Communication* 7(1): 37–56.
- Haman, E., Zevenbergen, A., Andrus, M. & Chmielewska, M. 2009. Coining compounds and derivations – A crosslinguistic elicitation study of word-formation abilities of preschool children and adults in Polish and English. *Polish Psychological Bulletin* 40(4): 176–192.
<https://doi.org/10.2478/s10059-009-0013-3>
- Hržica, G. 2011. Glagolske kategorije aspekta, vremena i akcionalnosti u usvajanju hrvatskog jezika (The verbal categories of aspect, tense and actionality in acquiring Croatian as a first language). PhD dissertation, University of Zagreb.
- Hržica, G. 2012. Daj mi to napisaj: Preopćavanja glagolske osnove u usvajanju hrvatskog jezika (Daj mi to napisaj: Overgeneralisations of the verb stem in the acquisition of Croatian language). *Suvremena Lingvistika* 74: 189–208.
- Jarmulowicz, L. D. 2002. English derivational suffix frequency and children's stress judgments. *Brain and Language* 81: 192–204. <https://doi.org/10.1006/brln.2001.2517>
- Katunar, D. 2013. Diminutives in action: A cognitive account of diminutive verbs and their suffixes in Croatian. *Suvremena Lingvistika* 39(75): 1–23.
- Kocijan, K., Šojat, K. & Poljak, D. 2018. Designing a Croatian aspectual derivatives dictionary: Preliminary stages. In *First Workshop on Linguistic Resources for Natural Language Processing*, A. Barreiro, K. Kocijan, P. Machonis & M. Silberztein (eds), 28–37. Stroudsburg PA: Association for Computational Linguistics.

- Kolaković, Z. & Jelaska, Z. 2009. Proširenost glagolskih predmetaka (The spread of verbal pre-fixes). *Lahor* 1(7): 59–76.
- Kovačević, M. 2002. CHILDES Database Kovacevic Corpus. <https://doi.org/10.21415/T5F55X>
- Kovačević, M., Kuvač, J. & Capanec, M. 2005. Je li glagolska prefiksacija morfološka ili leksička kategorija (Is verbal prefixation a morphological or lexical category)? Presented at the Third Conference of Speech and Language Pathologists of Croatia, Dubrovnik, Croatia.
- Kovačević, M., Palmović, M. & Hržica, G. 2009. The acquisition of case, number and gender in Croatian. In *Development of Nominal Inflection in First Language Acquisition*, U. Stephany & M. D. Voeikova (eds), 153–178. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110217117.153>
- Kuna, B. 2006. Nazivlje u tvorbi riječi (Terminology in word formation). *Filologija* 46–47: 165–182.
- Kuvač, J. & Palmović, M. 2001. Računalna obrada dječjeg jezika na primjeru usvajanja umanjena (Computer analysis of child language on the example of the acquisition of diminutives). *Suvremena Lingvistika* 1–2(51–52): 101–112.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Marković, I. 2018. *Uvod u jezičnu morfologiju* (An introduction to linguistic morphology). Zagreb: Disput.
- Martinot, C., Anđel, M. & Kumar, S. 2003. Acquisition of aspect and Aktionsart by children in Croatian and French. In *Acquisition of Aspect* [ZAS Papers in Linguistics 29], D. Bittner & N. Gagarina (eds), 133–148. Berlin: ZAS.
- Matovac, D. 2018. Glagoli kretanja i prefix *po-* u hrvatskom (Verbs of motion and the prefix *po-* in Croatian). *Croatica* 42(62): 205–219. <https://doi.org/10.17234/Croatica.42.13>
- Palmović, M. 2007. The acquisition of diminutives in Croatian. In *The Acquisition of Diminutives: A Cross-Linguistic Perspective* [Language Acquisition and Language Disorders 43], I. Savickienė, W. U. Dressler (eds), 73–88. Amsterdam: John Benjamins.
<https://doi.org/10.1075/lald.43.04pal>
- Pavličević-Franić, D. & Sikirić, M. 2005. Inovacijski semantemi u procesu usvajanja jezičnoga znaka (Innovatory semantic units in the process of linguistic sign acquisition). *Jezik* 52(3): 92–102.
- Raffaelli, I., Bekavac, B., Agić, Ž. & Tadić, M. 2008. Building Croatian WordNet. In *Proceedings of the Fourth Global WordNet Conference*, A. Tanács, D. Csendes, V. Vincze, C. Fellbaum & P. Vossen (eds), 349–359. Szeged: GWC.
- Ravid, D. 2004. Later lexical development in Hebrew: Derivational morphology revisited. In *Language Development across Childhood and Adolescence: Psycholinguistic and Crosslinguistic Perspectives* [Trends in Language Acquisition Research 3], R. A. Berman (ed.), 53–82. Amsterdam: John Benjamins. <https://doi.org/10.1075/tilar.3.07rav>
- Rivière, J., Lécuyer, R. & Hickmann, M. 2009. Early locomotion and the development of spatial language: Evidence from young children with motor impairments. *European Journal of Developmental Psychology* 6(5): 548–566. <https://doi.org/10.1080/17405620701345712>
- Rowland, C. F., Fletcher, S. L. & Freudenthal, D. 2008. How big is enough? Assessing the reliability of data from naturalistic samples. In *Corpora in Language Acquisition Research: History, Methods, Perspectives* [Trends in Language Acquisition Research 6], H. Behrens (ed.), 1–24. Amsterdam: John Benjamins. <https://doi.org/10.1075/tilar.6.04row>
- Silić, J. & Pranjković, I. 2005. *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta* (Croatian Grammar for Secondary Schools and Universities). Zagreb: Školska knjiga.

- Šojat, K., Srebačić, M. & Štefanec, V. 2013. CroDeriV and the morphological analysis of Croatian verb. *Suvremena Lingvistika* 39(75): 75–96.
- Šojat, K., Srebačić, M. & Tadić, M. 2012. Derivational and semantic relations of Croatian verbs. *Journal of Language Modelling* 10(1): 111–142. <https://doi.org/10.15398/jlm.voi1.34>
- Tafra, B. & Košutar, P. 2009. Rječotvorni modeli u hrvatskom jeziku (Word-formation models in Croatian). *Suvremena Lingvistika* 35(67): 87–107.

Acquisition of derivational morphology in Russian

Victoria V. Kazakovskaya and Maria D. Voeikova
Russian Academy of Sciences, Saint Petersburg

The chapter examines the onset of derivational morphology in the speech production of two Russian boys (aged between 1;5–2;8 and 1;8–3;0). This process involves early nouns, verbs and adjectives built by different derivation methods. Suffixation is typical for nouns and adjectives, whereas new verbs are mostly built by prefixation. Derived verbs and nouns belong to the same grammatical class as the originals and adjectives are formed from nouns. Productive use of derivation was registered several months after the first derivatives occur. The differences between adult-child dyads relate to the set of semantic categories of derived words and the proportion of derivatives belonging to different grammatical classes. Further development of derivation (4;2 to 5;10) in one of the boys was tracked.

Keywords: derivation, nouns, verbs, adjectives, Russian first language acquisition

1. Introduction

Russian, an East Slavic language, can be characterised as a morphology-rich language. With regard to its word-formation system, there are derivation methods using both affixation and non-affixation.

1.1 Affixation vs. non-affixation morphology

Morphological patterns are expressed by productive and frequent affixation, as well as, less frequently, by compounding (*grjaz+e+lečenij(e)*¹ ‘mud therapy’ ← *grjaz*’

1. Hereafter the sign ‘(...)’ will be used for marking the endings of feminine and neuter nouns and adjectives as well as the infinitive suffix *-t’* (*-ti*) in verbs to distinguish them from derivational suffixes. Most masculine nominative singular endings are zero.

'mud' +INTERF+ *lečenij(e)* 'treatment', *sero-golub(oj)* 'grey and blue' ← *ser(yj)* 'grey' +INTERF+ *golub(oj)* 'light blue').²

Nominal affixation consists of suffixation (*syn-ok* 'son-DIM' ← *syn* 'son'), prefixation (*pra-ded* 'great-grandfather' ← *ded* 'grandfather'), or a combination of prefixation and suffixation (*so-besed-nik* 'interlocutor' ← *besed(a)* 'conversation'). Lastly there is a combination of compounding and suffixation resulting in a synthetic compound (*mjas+o+rub-k(a)* 'meat grinder' ← *mjas(o)* 'meat' +INTERF+*rubi(t')* 'chop.INF' + *-k(a)*). The motivating stems for noun derivatives are other nouns, e.g. *syn-ok* 'son-DIM' (see above), verbs (*podar-ok* 'present' ← *podar-i(t')* 'give'), adjectives (*glup-ost'* 'foolishness' ← *glup(yj)* 'foolish'), adverbs (*potom-ok* 'descendant' ← *potom* 'then'), numerals (*sot-nj(a)* 'a hundred' ← *st(o)* 'hundred'), and prepositions in isolated cases (*pred-ok* 'ancestor' ← *pered* 'before'). The most productive patterns from them are 'N+affix', 'ADJ+affix' and 'V+suffix'. The most frequent model is 'N+suffix'.

Adjectives are derived from nouns (*golod-n(yj)* 'hungry' ← *golod* 'hunger'), verbs (*bol'-n(oj)* 'ill' ← *bole(t')* 'be ill'), other adjectives (*tix-on'k(ij)* 'quiet-DIM' ← *tix(ij)* 'quiet') and adverbs (*ran-n(ij)* 'early' ← *ran-o* 'early'). Only parametric adjectives like *bolš(oj)* 'big' may be non-derived. Relational adjectives corresponding to compound constituents in Germanic languages, such as *sel'-sk(ij)* 'rural', or possessive adjectives such as *mam-in* 'mommy's' are all derived and mostly built by suffixation. Prefixation is less diverse and frequent: *ne-bolš(oj)* 'not big' ← *bolš(oj)* 'big'. A potentially productive pattern is *-n-* suffixation, attested in adjectives derived from modern loan words, such as *komfort-n(yj)* 'comfortable' ← *komfort* 'comfort'.

Patterns for verb formation comprise (1) prefixation (delimitative *po-igra(t')* 'play for a while' ← *igr-a(t')* 'play'), (2) suffixation, including zero-suffixation, or conversion³ (*sol-i(t')* 'to salt' ← *sol'* 'salt'), (3) postfixation⁴ of the reflexive semiclitic *-sja* (*my(t')*-*sja* 'wash oneself' ← *my(t')* 'wash'), (4) a combination of them: prefixation with postfixation as in saturative verbs (*na-igra(t')*-*sja* 'play enough' ← *igr-a(t')* 'play'). Prefixes are often used for different aktionsarts modifying the lexical meaning, whereas suffixes change the aspectual meaning, e.g. by forming secondary

2. The tendency for nominal compounding has intensified in modern Russian, which is also reflected in a child-centered situation (Kazakovskaya 2017), whereas verb compounds are not productive.

3. This kind of verb formation is treated here as conversion (Manova & Dressler 2005: 73–79), although in Russian grammatical tradition thematic vowels are considered to be true derivational suffixes (Švedova 1980).

4. The notion of postfix, initially introduced by Baudouin de Courtenay (1904/1963: 101) for suffixes, later on, was used by Russian grammarians for semiclitic affixes following the inflectional ending (e.g. Jakobson 1957/1971: 145; Švedova 1980: 124).

imperfectives (*poigr-yva(t)* ‘use to play’ ← *poigr-a(t)* ‘play for a while’). The semi-clitic (postfix) *-sja* (shortened from *sebja* ‘oneself’) has several meanings derived from the reflexive. Prefixation is the most frequent pattern for forming verbs, although secondary imperfective suffixation is the most regular and productive. Verbs are mostly motivated by other verbs but may also be derived from nouns and adjectives (*krasn-e(t)* ‘become red’ ← *krasn(yj)* ‘red’). Prefixation, especially aktionsarts, and some suffixes, are potentially productive in applying to new loan words, such as *kser-it* ‘copy by xerox’ ← *kseroks* ‘xerox’.

The main non-affixing methods are by syntactic conversion and fusion. Syntactic conversion is defined as “a derivational process linking lexemes of the same form but belonging to different word classes” (Bauer & Valera 2005: 8), or a simple shift of word class (Clark 2009: 255) without changing inflection, e.g. from adjectives to nouns: *voenn(yj)* ‘military.N’ (about a person) ← *voenn(yj)* ‘military.ADJ’. Fusion (*mat’-i-mačex(a)* ‘coltsfoot.N’ ← *mat* ‘mother’ + *i* ‘and’ + *mačex(a)* ‘stepmother’) is much less frequent and often accompanied by suffixation, such as *vremjaprovodženij(e)* ‘pastime’ ← *vremj(a)* ‘time’ + *provod-i(t)* ‘spend’ + *-enij*. Both non-affixing methods are unproductive and quite rare.

Thus, Russian word classes can only be changed by suffixation (including a zero one) or syntactic conversion, whereas prefixes and the postfix always maintain the grammatical class but change the meaning. Prefixation is more productive and typical for verbs than for nouns and adjectives. The verb prefixes are very numerous, and many are polysemous with an unclear semantic connection between meanings: e.g. *po-igra(t)* ‘play for a while’ ← *igra(t)* ‘play’, but *po-j(ti)* ‘start going’ ← *id(ti)* ‘go’. Adjective prefixes create a closed class with a limited set of meanings, e.g. they often form antonyms: *ne-sčastliv(yj)* ‘unhappy’ ← *sčastliv(yj)* ‘happy’.

1.2 State of the art

Earlier Russian studies on the acquisition of derivational morphology were based mainly on diary observations, beginning with the famous pioneering parental diary of A. N. Gvozdev (1961), as well as observations of pre-schoolers’ speech and the results of some experiments (Leontjev 1965; Saxarnyj 1979; Kubrjakova 1981; Negnevickaja & Šaxnarovič 1981; Tambovceva 1983; Lepskaja 1997; Ušakova 2004; Jurjeva 2006; Eliseeva 2015, among others). The main aspects of early word formation including compounding were described in detail, regarding child-created words (Ceitlin 1989, 2009; Xarčenko & Ozerova 1999). Cross-linguistically, Russian was investigated in a study of diminutives (Savickienė & Dressler 2007) and of synthetic compounding in L1 (Dressler et al. 2019) and also in comparison with typologically different languages (Argus, Kazakovskaya & Laalo 2013; Argus &

Kazakovskaya 2013, 2018). The present study is one of the first investigations of derivatives in Russian child speech (CS) based on longitudinal data of spontaneous speech in interaction with caregivers.

2. Data

Our data consist of two longitudinal corpora of recorded spontaneous speech (74,359 tokens), which were transcribed and coded morphologically, according to CHILDES conventions (MacWhinney 2000).⁵ The children are typically developing Russian-speaking monolingual boys from two Saint Petersburg families with middle-class SES (see Table 1).

Table 1. Data CS: Grammatical classes and their percentage to all tokens

	Age	Hours of rec.	Total tokens	Noun tokens	Verb tokens	Adjective tokens
Filipp (F)	1;5–2;8	28	16,486	3,803 (23%)	1,920 (12%)	930 (6%)
Kirill (K)	1;8–3;0	6	5,769	1,518 (26%)	1,237 (21%)	279 (5%)
Total		34	22,255	5,321	3,157	1,209

Additionally, recordings available of Kirill from a later period (4;2–5;10) were analysed to trace his further development.

The percentage of noun and adjective tokens among all tokens in both CS corpora is comparable. However, the proportion of verbs in Kirill's speech is almost twice as big as in Filipp's. This difference may be explained by the properties of the child-directed speech (CDS), the different acquisition strategies of children and also by the effect of the smaller size of Kirill's corpus.

2.1 Nominal derivatives in CS

The proportion of nominal derivatives amongst all corresponding tokens in Filipp's speech is higher (31%) than in Kirill's (12%) (see Table 2). However, the percentage of new⁶ derivatives relative to all nouns is quite similar (12% and 8%).

5. We thank T. Pranova and K. Bayda for collecting these corpora and making them available.

6. New derivatives are those that were not registered in the previous recordings.

Table 2. Nominal derivatives in CS (lemmas/tokens)

	Nouns	Nominal derivatives	% of nominal derivatives among nouns	New nominal derivatives	% of new derivatives to a. all nouns b. all nominal derivatives
F	874/3,803	575/1,168	66%/31%	280/456	a. 32%/12% b. 49%/39%
K	580/1,518	89/176	15%/12%	69/126	a. 12%/8% b. 78%/72%

The total number of nominal derivatives in both corpora increases gradually until the end of the observation. However, in Filipp's speech this process seems to be more evident and the proportion of derivative tokens is larger (see *a.*), whereas in Kirill's data the percentage of new derivatives among all nouns derived is higher, both in lemmas and tokens (see *b.*).⁷ Figures 1a and 1b show the distribution of new derivatives within the period of observation, in rough numbers for each corpus.

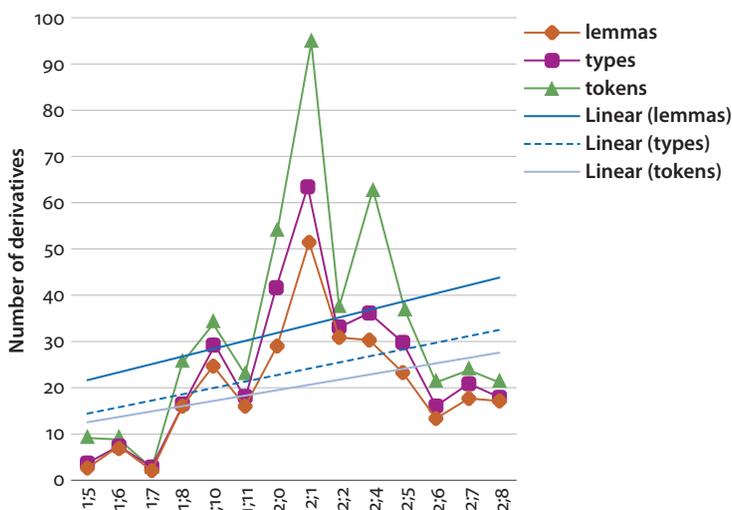


Figure 1a. Distribution of new nominal derivatives in Filipp's speech

Filipp begins to use nominal derivatives quite early, at 1;5 (see Figure 1a) with three denominal lexemes: *miš-k(a)* 'bear' (1/1), *dyr-k(a)* 'hole' (1/7), *mam-anj(a)* 'mother' (1/1).⁸ They are not frequent and occur only in the nominative singular.

7. Additionally, Kirill has more compounds. Hereafter, relevant results on compounding in the speech of both subjects will be cited after Kazakovskaya (2017).

8. Relevant simplicia were documented earlier or simultaneously.

At 2;1 a peak in derivatives occurs (51/95). The age of 1;10 might be considered to be a pre-peak (24/34). Trends of lemmas, types and tokens offer some evidence in favor of fairly uniform and parallel development.

Kirill starts to use nominal derivatives later, at 2;0 (see Figure 1b). His initial derivatives are deadjectival *prjan-ik* ‘gingerbread’ (1/1) and deverbal *beg-un* ‘runner’ (1/1). At the age of 2;6 and 2;9, two peaks of derivatives occur.

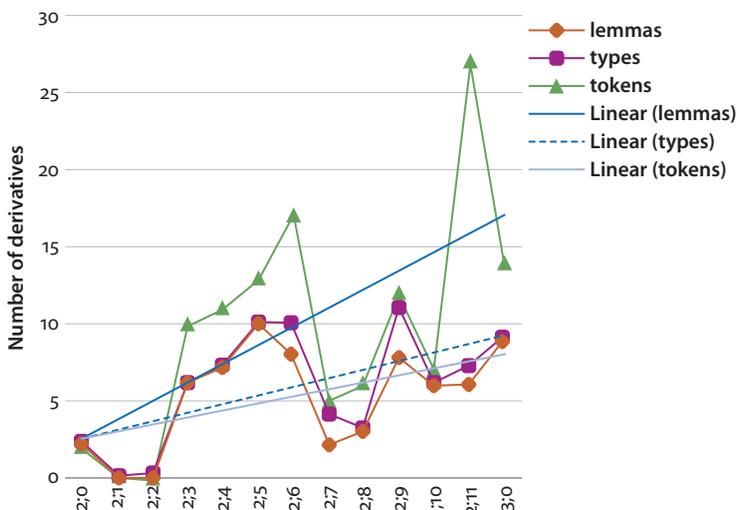


Figure 1b. Distribution of new nominal derivatives in Kirill's speech

In contrast to Filipp's derivative development, the token frequency of derivatives in Kirill's speech grows faster than both their diversity (lemmas) and different inflectional forms emerge. In both CS the development of nominal inflection starts after derivation.

2.2 Adjective derivatives

Adjectives are scarce in both CS corpora. Their number in Filipp's speech is four times smaller than that of nouns and half of that of verbs. In Kirill's speech adjectives occur even more seldom (compare Table 1).⁹ Table 3 shows the total number of adjective derivatives in CS and the proportion of new ones.

9. In both CS corpora adjectives make up about 5–6% of all tokens, which is a normal average value for children under 3;0 as stated for 11 different languages in Tribushinina, Voeikova & Noccetti (2015).

Table 3. Adjective derivatives in CS (lemmas/tokens)

	ADJ	ADJ derivatives	% of ADJ derivatives among ADJ	New ADJ derivatives	% of new derivatives to a. all ADJ b. all ADJ derivatives
F	143/930	96/170	67%/18%	64/117	a. 45%/69% b. 67%/17%
K	164/279	67/128	41%/46%	45/82	a. 27%/29% b. 67%/64%

Both boys show “adjective spurts” (F. 1;11, K. 2;4) when the rough number of adjectives per recording exceeds 30.¹⁰ These periods are also extremely rich in the percentage of adjective derivatives that, however, do not exhibit any permanent increase in rough numbers, unlike their inflectional diversity (Tribushinina et al. 2015: 13; Voeikova 2015: 212–213). Thus, adjective derivative patterns are not acquired as consistently as their inflectional forms.

Adjective frequency depends on the topic of conversation (Tribushinina et al. 2015: 177, 212–213, 325). Mothers initiate the discussion of several properties of objects, probably having some expectations about the child’s “zone of the proximal development”, as postulated by Vygotskij (1935: 42). In such discussions children use about 40 different adjectives, including numerous color names, such as *višn-ev(yj)* ‘cherry-coloured’ ← *višnj(a)* ‘cherry’ (K. 2;5), *ser-en’k(ij)* ‘grey-DIM’ ← *ser(yj)* ‘grey’ (F. 1;11). Their number is three times greater than in the previous recordings. The boys tend to use antonyms and gradually learn to form them: *ne-vkusn(yj)* ‘unappetizing’ ← *vkus-n(yj)* ‘tasty’ (F. 2;1, K. 2;11). Suffixed derivatives predominate absolutely: *krug-l(yj)* ‘round-ADJ’ ← *krug* ‘round.N’ (F. 2;2). The number of adjectives derived from nouns in CS reaches 50% of all adjective lemmas.

2.3 Verb derivatives in CS

The percentage of verb derivatives among all verbs is higher than that of nominal derivatives: 34% vs. 31% (F.) or even 40% vs. 12% (K.) (see Table 4). The same percentages in both CS corpora are quite similar, ranging between 42% and 54% for lemmas and 34% and 40% for tokens.

10. An adjective spurt is defined as a “rapid transition from several frozen and rote-learned adjective forms to dozens of tokens per hour” (Tribushinina et al. 2015: 225) attested 2–3 months after the first occurrence of frozen adjective forms.

Table 4. Verb derivatives in CS (lemmas/tokens)

	Verbs	Verb derivatives	% of verb derivatives among verbs	New verb derivatives	% of new derivatives to a. all verbs b. all verb derivatives
F	910/1,920	382/660	42%/34%	237/467	a. 26%/24% b. 62%/71%
K	412/1,237	222/498	54%/40%	175/346	a. 42.5%/28% b. 79%/69.5%

The similarly high percentage of verb derivatives in CS is an indicator that derivation patterns in verbs are more productive than in nouns. At the end of the observation period the proportion of verb derivatives reaches about 50% of all verbs in both CS corpora, whereas for nouns a similar percentage is only achieved by Filipp. The children show different developmental strategies in the acquisition of verb derivation: as regards Filipp, derivatives already exceed half of all verb tokens at 1;8, and later, at 1;11, 2;2, 2;4 and 2;6.¹¹ Figures 2a and 2b represent the correlation of lemmas and tokens of verb derivatives for the children.

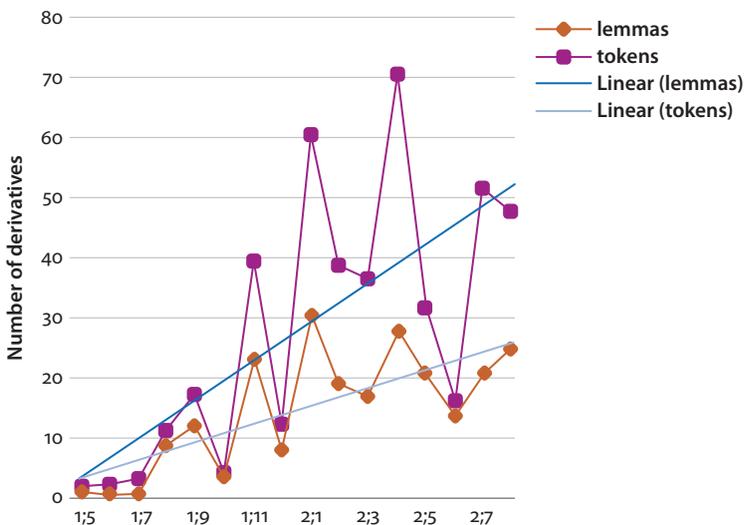


Figure 2a. Distribution of new verb derivatives in Filipp's speech

The uneven distribution of derived verbs in CS shows that their acquisition is not a gradual process but a joint technique of coining new words and repeating derived

11. This results from his repetitive strategy of language acquisition: the mean share of repetitions by Filipp after his mother, exceeds 27% (Voeikova 2015: 53–57), whereas for Kirill it is about 13%.

verbs from CDS. In CS we find the use of polysemic prefixes: compare *u-beža(t')* ‘PFV-run away’ ← *beža(t')* ‘run’, *u-kusi(t')* ‘PFV-bite once’ ← *kusa(t')* ‘bite’ (F. 1;9). At the same time, Filipp also starts to use derivation models with one and the same meaning, e.g. inchoative *po-exa(t')* ‘PFV-start moving by car’ and *po-j(ti)* ‘PFV-start going’ (1;8). This effect might be due to the great frequency of certain prefixes, especially *po-* in colloquial Russian (Janda et al. 2013: 93–97).

Kirill starts to use derived verbs much later: there are almost no derivatives before 2;4, followed by four months of very scarce use and a constant increase from 2;8 to 3;0 (see Figure 2b).

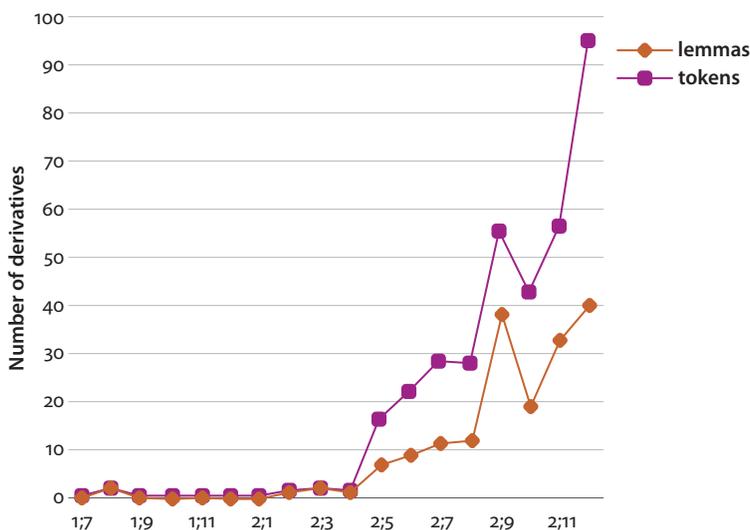


Figure 2b. Distribution of new verb derivatives in Kirill's speech

The number of derived verb forms (tokens) in Kirill's CS significantly increases starting from 2;5, but only after 2;9 does their number exceed that of non-derived verbs in lemmas. However, this difference is very small.

3. Early derivational patterns in CS

3.1 Derivational nominal patterns

In the early phases nominal derivation is limited to forming new nouns from nouns, i.e. within one grammatical class of words (including diminutives), which corresponds to Gvozdev's (1961) results. The earliest derivatives are used without inflectional oppositions up to 1;10 in Filipp's speech and up to 2;5 in Kirill's. This period

lasts for five months for both children, after which they start to use first contrasting forms of number and case in derived nouns.

Children begin to use derivatives with patterns that are mostly productive and frequent in modern Russian, including colloquial speech. They are ‘N+affix’, ‘V+suffix’, ‘ADJ+affix’ (see Table 5). Filipp starts with the more frequent ‘N+suffix’ pattern, whereas Kirill uses ‘V+suffix’ and ‘ADJ+suffix’ (see examples in 2.1). Despite different beginnings of the children in relation to the base stem, in both CS corpora the dominant pattern is denominal. Derivatives constructed accordingly are approximately 90% of all derived nouns in Filipp’s CS, and 70% in Kirill’s.

Table 5. Nominal patterns in CS (lemmas/tokens, their % of derived nouns)

Patterns	F	K
N+affix	244/408 (87%/89%)	34/72 (61%/67%)
N+suffix	244/408	32/70
prefix+N	0	1/1
prefix+N+suffix	0	1/1
V+affix	29/35 (10%/8%)	19/33 (34%/30%)
V+suffix	29/35	19/33
ADJ+affix	7/13 (3%/3%)	3/3 (5%/3%)
ADJ+suffix	5/10	3/3
prefix+ADJ+suffix	2/3	0
Total	280/456	56/108

Affixation is not only earlier in CS, but also much more frequent than compounding.¹² Within affixation, explicit suffixation is earlier and the most frequent model. Thus, up to 3;0 Filipp uses 60 suffixes, whereas Kirill uses only 30. The largest variety of suffixes (28) is documented in the ‘N+affix’ pattern (see Table 6). Zero suffixation, which is more typical for deverbal nouns and is also possible within synthetic compounds ‘X+V’, occurs later: e.g. *kus-ak(a)* ‘nipper’ ← *kus-a(t)* ‘bite’ (F. 1;8) is earlier than *pricep-ø* ‘trailer’ ← *pricep-i(t)* ‘attach’ (F. 2;0).

Although suffixation predominates, the children use a few prefixes. Combination of suffixes and prefixes such as *pod-guzn-ik* ‘diaper’ ← *guzn(o)* ‘~goat, tail’ (K. 2;9) or compounding accompanied by suffixation as in *mux+o+mor-ø* ‘amanita, lit. fly killer’ (F. 2;2) ← *mux(a)* ‘fly’ +INTERF+ *mor-i(t)* ‘kill’ occurs later on, as well as non-affix methods of word formation. In particular, syntactic conversion is represented by substantivation: *požarn(yj)* ‘fireman.N’ ← *požarn(yj)* ‘fire.ADJ’ (e.g. in *fire house*) (K. 2;10).

12. First nominal compounds were documented at 1;11 (F.) and at 2;5 (K.). Like first affix derivatives, they are far from perfect, phonologically.

Table 6. Affixes in nominal derivatives (cumulative)

Patterns	No. of affixes	Order of emergence
N+affix	31	
N+suffix	28	<i>-k, -anj, -ok, -ik, -uš, -ux, -nik, -ut, -čik, -ek, -ic, -en(‘)k, -ušk, -yšk, -očk, -c, -ečk, -ix, -ul’, -ess, -n’, -ink, -utk, -av, -ovn, -ič; -aljon, -ovin</i>
prefix+N	1	<i>pra-</i>
prefix+N+suffix	1+1	<i>pod-č-nik</i>
V+suffix	18	<i>-k, -ak, -onok, -ø, -(e)n(i)j, -d, -l, -unok, -ux, -lk, -šk, -v, -nic; -un, -ščik, -(en)ok, -tel’, -(a)nk, -tv</i>
ADJ+affix	9	
ADJ+suffix	7	<i>-ak, -yš, -ic, -ost’; -ik, -atin</i>
prefix+ADJ+suffix	1+1	<i>pod-č-ik</i>

Within denominal patterns the following semantic categories are represented by suffixation (see examples below):

- a. Females (people and animals) are formed with *-ix, -k, -ic, -ess: zajč-ix(a)* ‘hare-F’ ← *zajac* ‘hare’ (F. 2;2), *vnuč-k(a)* ‘granddaughter’ ← *vnuk* ‘grandson’ (K. 2;3);
- b. Baby animals are formed with *-onok: ljaguš-onok* ‘frogling’ ← *ljagušk(a)* ‘frog’ (F. 1;10);
- c. Singulatives are formed with *-in, -ink, -k: goroš-in(a)* ‘one pea’ ← *gorox* ‘peas’ (F. 2;8);
- d. Instruments are formed with *-ušk, -nik: igr-ušk(a)* ‘toy’ ← *igr(a)* ‘game’ (F. 2;0), *gradus-nik* ‘thermometer’ ← *gradus* ‘degree’ (F. 2;2);
- e. Agents are formed with *-nik: *gitar-nik*¹³ ‘guitar player’ ← *gitar(a)* ‘guitar’ (K. 2;7);
- f. Diminutives are formed with *-k, -ik, -ok, -uš, -čik: vagon-čik* ‘small railway carriage’ ← *vagon* ‘railway carriage’ (F. 1;10).¹⁴

Semantic categories of the less frequent deverbal nouns are represented by derivatives with agentive, processual, resultative, instrumental and spatial semantics. For their formation children use the following suffixes:

13. Hereafter “*” will be used for child-created derivatives.

14. Diminutives constitute 65% of all derived nouns in Filipp’s CS, and 44% in Kirill’s (Argus & Kazakovskaya 2018; cf. Protassova & Voeikova 2007). Stylistic (viz. colloquial speech) modifications of words (Švedova 1980: 183) such as *kolen-k(a)* ‘knee’ ← *kolen(o)* ‘knee’ (K. 2;3) are included here within diminutives.

- a. *-un, -ščik, -tel, -ux* for agents: *beg-un* ‘runner’ ← *beg-a(t’)* ‘run’ (K. 2;0), *gon-ščik* ‘racer’ ← *gn-a(t’)* ‘drive very fast’ (K. 2;4), *iska-tel* ‘seeker’ ← *isk-a(t’)* ‘seek’ (K. 2;10);
- b. *-k, -lk, -až, -ø* for processes and activities: *lep-k(a)* ‘molding’ ← *lep-i(t’)* ‘sculpt’ (F. 2;6), *rabot-ø(a)* ‘work’ ← *rabot-a(t’)* ‘work’ (K. 3;0);
- c. *-(e)n(i)j, -ok* for results: *podar-ok* ‘gift’ ← *podar-i(t’)* ‘give’ (K. 2;4), *var-enj(e)* ‘jam’ ← *var-i(t’)* ‘boil jam’ (K. 2;10);
- d. *-lk, -tv* for instruments: *meša-lk(a)* ‘mixer’ ← *meš-a(t’)* ‘mix’ (K. 2;5), *bri-tv(a)* ‘razor’ ← *bri(t’)* ‘shave’ (F. 2;5);
- e. *-ok, -(e)nj, -k, -nk* for locatives: *kat-ok* ‘(ice) rink’ ← *kat-a(t’)-sja* ‘skate’ (K. 2;5), *stoja-nk(a)* ‘parking’ ← *stoj-a(t’)* ‘stand’ (K. 3;0).

The percentage of deverbal derivatives among nouns derived is much higher in Kirill’s speech (34%/30%) (see Table 5).

Rare in CS data, deadjectival nouns are represented by the two patterns ‘ADJ+suffix’ and ‘prefix+ADJ+suffix’ which express the semantics of (a) concrete objects or subjects having definite qualities and (b) abstract qualities. The deadjectival nouns were mainly documented after 2;4. Their percentage among derived nouns is slightly higher in Kirill’s speech (5%/3%). They are formed by the suffixes *-ulj, -ik, -yš*: *grjazn-ulj(a)* ‘a dirty and untidy person’ ← *grjazn(yj)* ‘dirty’ (F. 2;1), *gruzov-ik* ‘truck’ ← *gruzov(oj)* ‘cargo.ADJ’ (K. 2;4). This group also includes nouns with spatial semantics expressed by *-ic* and *-ynj*: *bol’n-ic(a)* ‘hospital’ ← *bol’n(oj)* ‘sick’ (K. 2;10), *pust-ynj(a)* ‘desert’ ← *pust(oj)* ‘empty’ (K. 2;9). The deadjectival nouns meaning abstract qualities are mainly formed by *-ost*: *glup-ost* ‘foolishness’ ← *glup(yj)* ‘foolish’ (F. 2;8). Abstract nouns are documented at the end of the observations.

The prefixes *pod-* and *pra-* are used in denominal and deadjectival patterns. In particular, Kirill uses *pra-* within a denominal model (1/1): *pra-ded* ‘great-grandfather’ ← *ded* ‘grandfather’ (2;4). Within the combining models ‘prefix+N/ADJ+suffix’ both boys use the prefix *pod-* and the suffix *-ik*: *pod-osinov-ik* ‘orange-cap boletus, lit. a mushroom which grows under an aspen’ ← *osinov(yj)* ‘aspen.ADJ’ (F. 2;1). A few examples (3/7 cumulatively) were documented during the third year of life. Their semantics is ‘an object with definite (close to locative ‘under’) qualities’.

3.2 Derivational adjective patterns

Both children start with rare repetitions of adjectives uttered by adults. Filipp repeats two adjectives already at 1;5, in Kirill’s CS they occur seldom only from 1;10. Derived adjectives rarely occur 2–3 months after the first use. First relational adjectives are used without their base nouns: e.g. *ovsj-an(yj)* ‘oat-ADJ’ ← *ovsj* ‘oat’ occurs in Filipp’s CS at 1;8 (about porridge) when he has no idea what ‘oat’ is.

The denominative pattern is the most productive (see Table 7), including possessive adjectives, such as *bab-in* ‘granny’s’ ← *bab(a)* ‘granny’. Denominal adjectives may also express colors, form (see 2.2), or material: *rezin-ov(yj)* ‘rubber-ADJ’ ← *rezin(a)* ‘rubber’ (F. 2;2).

Table 7. Adjective patterns in CS (lemmas/tokens, their % of derived adjectives)

Patterns	F	K
N+affix	56/95 (58%/56%)	40/75 (58%/58%)
N+suffix	55/94	38/72
prefix+N+suffix	1/1	0
V+affix	7/9 (7%/5%)	7/8 (10%/6%)
V+suffix	7/9	7/8
ADJ+affix	19/48 (20%/28%)	18/39 (26%/30%)
ADJ+suffix	13/42	9/27
prefix+ADJ	6/6	8/11
PREP/ADV+suffix	5/8 (5%/5%)	3/4 (4%/3%)
NUM+suffix	2/2 (2%/1%)	0
Compounding+suffix	7/8 (7%/5%)	1/4 (1%/3%)
Total	96/170	67/128

Suffixed patterns predominate absolutely in adjectives (also in adult-directed speech, ADS) (see Table 8).

Table 8. Affixes in adjective derivatives (cumulative)

Patterns	No. of affixes	Order of their emergence
N+affix	11	
N+suffix	7	<i>-an, -n, -ov/-ev, -l, -ij, -k</i>
prefix+N	2	<i>ne-, pre-, pod-</i>
prefix+N+suffix	2	<i>bez-&-ij</i>
V+suffix	4	<i>-t, -č, -n</i>
ADJ+affix	2	
ADJ+suffix	2	<i>-en’k, -usen’k</i>
PREP+suffix	1	<i>-n</i>
ADV+suffix	2	<i>-n, -k</i>
NUM+suffix	2	<i>-ok, -akov</i>

Some expressive forms may also be derived from other adjectives: *mal-jusen’k(ij)* ‘little-DIM’ (K. 3;0) ← *mal(yj)* ‘little’. Rarely adjectives are derived from adverbs or prepositions, such as *pered-n(ij)* ‘frontal’ (1;9) ← *pered* ‘in front of’ (F. 2;2).¹⁵

15. The base preposition is attested later than the derived adjective; however, after a short interval.

The deverbal pattern is used to express properties resulting from the impact on the object: *kolj-uč(ij)* 'thorny, prickly' ← *kol-o(t')* 'prickle'. Prefixed adjectives with *ne-* 'not' are used for antonyms. The percentages of different patterns are similar for both children.

3.3 Derivational verb patterns

3.3.1 General remarks

The only derivational verb pattern used by both children is 'V+affix' (see Table 9). Although 'N+affix' and 'ADJ+affix' patterns exist and are even productive in spoken Russian, they almost never occur in our data. We recorded only three denominal verbs: *po-mest-i(t')* 'place' ← *mest(o)* 'place' (F. 2;2), *zavtrak-a(t')* 'have breakfast' ← *zavtrak* 'breakfast' (F. 2;3), *ozornič-a(t')* 'mess about' ← *ozornik* 'mischief' (F. 2;3); two deadjectival verbs: *čist-i(t')* 'clean-INF' ← *čist(yj)* 'clean' (F. 2;7), *zdravstv-ova(t')* 'be healthy' ← *zdrav(yj)* 'healthy' (F. 2;8), as well as one verb *mjau-ka(t')* formed from onomatopoeic *mjau* 'mew' (F. 2;4).

Table 9. Verb patterns in CS (lemmas/tokens, their % of derived verbs)

Patterns	F	K
V+affix	377/643 (98%/98%)	222/498 (100%)
V+suffix	22/41	34/40
prefix+V	305/529	143/372
V+postfix	45/70	39/70
prefix+V+postfix	4/4	6/16
N+affix	3/7 (> 1%)	0
ADJ+affix	2/8 (> 1%)	0
ONOM+suffix	1/1 (> 1%)	0
Total	382/660	222/498

Verb derivation is limited to slight semantic modifications and/or aspectual distinctions. Earliest derivatives occur at different phases: in Filipp's CS the phonologically opaque past form of the perfective prefixed verb *u-pas(t')* 'PFV-fall down' ← *pas(t')*/*pada(t')* 'fall down' already occurs at 1;5. At 1;8 he starts to use several derivative models: the opposition *ot-da(t')* 'PFV-give back' ← *da(t')* 'give', inchoative *po-exa(t')* 'start going by vehicle' ← *exa(t')* 'go by vehicle', resultative *u-kus-i(t')* 'PFV-bite once' ← *kus-a(t')* 'bite'. Filipp probably still does not realise the connections between derived and basic stems, however, simultaneous occurrence of several verb derivatives shows that he accepted the necessity to modify words. In Kirill's CS the same point was at 2;5 (10 months after the beginning of the observation) when he started

to modify verbs with prefixes: *po-exa(t')* 'PFV-start going by vehicle' and *pri-exa(t')* 'PFV-come by vehicle' ← *exa(t')* 'go by vehicle', or *pere-puta(t')* 'PFV-mix' ← *puta(t')* 'mix.IPFV'.

The number of verbal affixes (see Table 10) is more restricted than nominal ones. Suffix and postfix models, mostly expressing aspectual changes, were all the same for both boys. The only diversity is represented by numerous verbal prefixes.

Table 10. Affixes in verb derivatives (cumulatively)

Patterns	No. of affixes	Order of their emergence
V+affix	25	
V+postfix	1	- <i>sja</i>
prefix+V	15	<i>do-</i> , <i>is-</i> , <i>na-</i> , <i>o-</i> , <i>ot-</i> , <i>pere-</i> , <i>po-</i> , <i>pri-</i> , <i>pro-</i> , <i>pod-</i> , <i>ras-</i> , <i>s-</i> , <i>u-</i> , <i>vy-</i> , <i>za-</i>
V+suffix	5	- <i>va</i> , - <i>iva/-yva</i> , - <i>nu</i> , - <i>i/-a</i>
prefix+V+affix	4	<i>za-&-sja</i> , <i>na-&-sja</i> , <i>po-&-a</i> , <i>vy-&-a</i>
N+suffix	2	- <i>a</i> , - <i>i</i>
ADJ+suffix	2	- <i>i</i> , - <i>stvova</i>
ONOM+suffix	1	- <i>ka</i>

3.3.2 Postfix models of verb formation

The postfix -*sja* expresses different reflexive meanings:

- self-reference: *my(t')*-*sja* 'wash oneself' ← *my(t')* 'wash' (F. 1;9);
- reflexive counterparts from causative transitive verbs: *trjas(ti)*-*s* 'tremble' ← *trjas(ti)* 'shake' (K. 2;4), *kruti(t')*-*sja* 'move around' ← *kruti(t')* 'spin' (F. 1;8);
- autocausative verbs, such as *otloma(t')*-*sja* 'brake off' ← *otloma(t')* 'snap off' (F. 2;0) are formed either by postfix, or by the combination of the postfix with prefixes: *za-gore(t')*-*sja* 'set on fire' ← *gore(t')* 'burn' (K. 3;0);
- absolutive verbs: *kusa(t')*-*sja* 'be bitey (about animals)' ← *kusa(t')* 'bite' (F. 1;9);
- passive (including decausatives and quasipassives): *otkryva(t')*-*sja* 'be opened' ← *otkryva(t')* 'open' (K. 2;8).

Although the children start using reflexive verbs at different times (F. 1;8, K. 2;4), they used this model more often than the suffix models.

3.3.3 Suffix models of verb formation

Verb suffixes usually change the aspect (in this case mostly without change of lexical meaning):

- a. If a verb pair with the same root differs by the alternation of the *-i/-e* vs. *-a* “stem vowel” the *-i/-e* variant is perfective, whereas the *-a* variant is imperfective: *zagar-e(t')* ‘get a tan.PFV’ (K. 2;8, F. 2;4) vs. IPFV *zagar-a(t')*.
- b. The *-iva/-yva*, *-va* suffixes form secondary imperfective counterparts from perfective verbs: *pokaz-yva(t')* ‘show-IPFV’ ← PFV *pokaz-a(t')* (K. 1;8), *otkry-va(t')* ‘open-IPFV’ ← PFV *otkry(t')* (F. 2;2). Semantically, they express doing something several times or being in the process of doing something.
- c. The suffix *-nu* is semelfactive and turns the verb into a perfective: *pryg-nu(t')* ‘jump once-PFV’ ← *pryga(t')* ‘jump-IPFV’ (F. 2;6).

3.3.4 Prefix models of verb formation

The main way of changing the lexical meaning of the verb is adding prefixes expressing aktionsart, or path for verbs of motion. Most of the prefixes have several meanings. The semantic categories are:

- a. motion in a certain direction, e.g. downwards: *s-lez(t')* ‘PFV-peel off’ ← *lez(t')* ‘climb-IPFV’ (F. 2;8); away: *u-exa(t')* ‘PFV-go away by a vehicle’ ← *exa(t')* ‘go by vehicle-IPFV’ (F. 2;2, K. 2;7); out: *vy-lez(ti)* ‘PFV-crawl out’ ← *lez(t')* ‘climb-IPFV’ (F. 2;2);
- b. detachment: *ot-da(t')* ‘PFV-give back’ ← *da(t')* ‘give-IPFV’ (K. 2;3, F. 1;8); division into small parts: *raz-bi(t')* ‘PFV-break into pieces’ ← *bi(t')* ‘beat-IPFV’ (F. 2;2);
- c. transportation or change of place: *pere-vozi(t')* ‘IPFV-transport’ ← *vozi(t')* ‘take by vehicle-IPFV’¹⁶ (K. 2;10);
- d. start of motion: *po-j(ti)* ‘PFV-start going’ ← *id(ti)* ‘go-IPFV’ (F. 2;8);
- e. any telic action: *podo-žda(t')* ‘PFV-wait’ ← *žda(t')* ‘wait-IPFV’ (F. 2;8), *na-risova(t')* ‘PFV-draw’ ← *risova(t')* ‘draw-IPFV’ (K. 2;11).

In general, verb prefixes have either spatial or temporal meaning.

4. Development of derivatives in CS

4.1 Development of nominal derivatives in CS

The word-formation features of early nominal derivation in Russian include the following: (1) the precedence and dominance of suffix models (since suffixes are more salient than prefixes due to the recency effect), (2) the predominance of word-class maintaining derivations, (3) the precedence of overt suffixes over zero

16. Only unidirectional motion verbs are perfectivized by prefixes, verbs of multidirectional motion remain imperfective.

ones (or conversion), (4) the precedence of simple derivatives over complex ones (namely, formed with the help of a further morpheme from other derivatives or compounds). This is justified, in turn, by the degree of cognitive complexity of the derivative, which has at least one more semantic element added to the corresponding simplex (Kazakovskaya 2019: 125). For instance, denominal suffixal derivatives with diminutive semantics can be used at the second year, regardless of the morphemic complexity of their motivating stem (*žuč-oč-ek* ‘beetle-DIM-DIM’ (1;10) ← diminutive *žuč-ok*). (5) Nominations of abstract qualities (*glup-ost* ‘foolishness’ (2;8)) and/or of their owners begin to appear sporadically only by the end of the third year.

A developmental analysis of the semantic categories of early derivatives revealed some sequence types in their occurrence. Thus, after diminutives (1;5–2;3),¹⁷ females (1;6–2;3), actions (1;7–2;0), baby animals (1;8), agents (1;8–2;0), singulatives (2;1–2;5), the results of actions (2;4) and instruments (2;4–2;5) appear. Derived nouns denoting males (2;4), activities and/or occupations (2;4–2;6), as well as derivatives with spatial (2;6) and qualitative (2;8–2;9) semantics complete the list of semantic categories up to 3;0.

The frequency of derivatives belonging to these semantic categories also differs. Within frequent early semantic categories (emergence up to 2;5), there are denominal diminutives, females and baby animals. Suffix singulatives and prefixed names of males, as well as deverbal nouns, are used less frequently. Among the latter are nominations of a process and/or its result, of an agent, of an instrument. Later emerging semantic categories (emergence from 2;6) are represented by the deverbal names of professions and/or activities and locations, deadjectival nominations of locations and abstract qualities. Thus, children acquire semantic categories of noun derivatives, starting from the naming of objects and subjects (people and animals). Abstract categories like the naming of processes and qualities emerge later.¹⁸

The development of the productivity of nominal derivatives begins together with the appearance of their forms of number and case. In particular, at 1;10 in Filipp’s speech three derivatives (*cvetoč-ek* ‘small flower’, *koles-ik(o)* ‘small wheel’, *myš-onok* ‘baby mouse’) have three inflectional forms (i.e. mini-paradigms), and (1) the first derivative chains: *grib* ‘mushroom’ (1;8) → *grib-ok* ‘mushroom-DIM’ (1;10) → *grib-oč-ek* ‘mushroom-DIM-DIM’ (1;10); (2) the first child-created derivative *det-ik* ‘child-DIM’ ← *ditj(a)* (1;7) and (3) zero suffixation occur. The pairs “non-derived noun → derived noun” were documented earlier: *dyr(a)* ‘hole’ (1;5) → *dyr-k(a)* ‘hole’ (1;5),

17. Hereafter the time of the first documentation of the derivative in each CS is presented.

18. Among the main semantic groups of compounds in CS are agents, instruments and locations.

kis(a) ‘pussicat’ (1;5) → *kis-k(a)* ‘pussicat’ (1;8). At 1;11 the first compound emerged. The formation of nominal word families starts at the second year, from 1;6 onwards, mainly consisting of diminutives and hypocoristics. In addition, baby animals and females are their first members: *mys̄* ‘mouse’ → *mys̄-onok* ‘baby mouse’ (1;10) and *mys̄-k(a)* ‘mouse-F’ (2;0); *kot* ‘cat’ → *kotj-onok* ‘baby cat’ (1;8) and *koš-k(a)* ‘cat-F’ (2;0).

4.2 Development of adjectival derivatives in CS

Both children start to utter color, size and evaluative adjectives without demonstrating any difference in the use of derived and non-derived words. Several months later we register the addition of similar suffixes to different stems, e.g. possessive *mam-in* ‘mommy’s’ and *bab-in* ‘granny’s’ (F. 1;11), then relational *apel’sin-ov(yj)* ‘orange-coloured’ and *višnj-ov(yj)* ‘cherry-coloured’ (K. 2;5). In the same recording Kirill has used the prefix *ne-* ‘non’ with three different stems: *ne-vskusn(yj)* ‘unappetizing’, *ne-dovol’n(yj)* ‘unsatisfied’, *ne-krasiv(yj)* ‘ugly’ to form antonyms from *vskusn(yj)* ‘tasty’, *dovol’n(yj)* ‘happy’, *krasiv(yj)* ‘beautiful’. This shows that he can clearly apply some derivational models. Filipp does the same already at 2;1. Although the prefix model is not the most frequent one in adjective formation, it might be acquired early due to its semantic and phonological transparency.

The earliest models do not coincide in both CS corpora (see Table 11). A general idea of relation to an object or person is expressed early, followed by the specification of color, shape or possession and antonyms precede synonyms.

Table 11. Semantic categories of adjectival derivatives in CS

F	evaluation (1;5) → relation to object (1;8) → form, location, possession (1;9) → ability (2;0) → antonyms (2;1) → shape, color (2;2) → material (2;3) → absent feature (2;4) → quality of the part (2;5) → high degree of quality (2;6)
K	relation to object (2;3) → evaluation (2;4) → color, antonyms (2;5) → possession (2;6) → high degree of quality (2;8)

In general, adjectives do not form families in CS. The only case registered was: *bab-in* ‘granny’s’ ← *bab(a)* ‘woman, granny’ (1;11) and *bab-ušk-in* ‘granny-DIM’s’ (F. 2;1).

4.3 Development of verb derivatives in CS

Verbs in CS are only derived from other verbs. Unlike nouns and adjectives, verbs are predominantly built by prefixation that modifies their meaning. There is no evidence that children realise these intricate semantic nuances at the earliest stages: they are most likely using the readily available prefixed verbs that they adopt from CDS. Verb members of aspectual oppositions may greatly differ by their frequency so that one of them has a chance of being acquired earlier, irrespective of their structural properties. The same is true for other lexico-grammatical derivatives, such as reflexive counterparts of the transitive verbs or repetitive action. Table 12 shows the emergence of semantic categories of verb derivatives.

Table 12. Semantic categories of verb derivatives in CS

F	result (1;5) → return back (1;8) → start of movement, reflexive (1;9) → delimitative → repetitive action, arrival, approaching (1;11) → rejection (2;0) → distributive (2;2) → source of movement (2;3)
K	repetitive action (1;8) → reversive action (2;3) → reflexive (2;4–2;5) → start of movement; arrival, approaching (2;5) → exhausting; delimitative (2;6) → result (2;7) → source of movement (2;9) → penetration (2;10) → surface contact (2;11) → inchoative (3;0)

At the end of the second year children start to use mixed models of verb formation involving prefixation accompanied by the postfix. This rise of complexity is presented by three models: (a) deep involvement into action: *za-risova(t')-sja* 'PFV-be absorbed by drawing' ← *risovat'* 'draw.IPFV' (K. 3;0); (b) the beginning of the auto-causative action: *za-gore(t')-sja* 'PFV-start burning' ← *gore(t')* 'burn' (K. 3;0); (c) the combination of direction and reflexivity: *v-reza(t')-sja* 'PFV-bump into' ← *reza(t')* 'cut' (K. 2;11). Mixed methods of verb formation follow elementary affixation.

Verbs of the same semantic class can form pairs and families. This is especially characteristic of verbs of motion: *id(ti)* 'go' → *uj(ti)* 'go away' and *poj(ti)* 'start going' (F. 1;8), *exa(t')* 'go by vehicle' → *priexa(t')* 'arrive' and *poexa(t')* 'start moving by vehicle' (K. 2;5). Other verbs develop aspectual nuances: *pi(t')* 'drink.IPFV' → *popi(t')* (F. 2;3) and *zapi(t')* 'drink after' (K. 2;8). This is the first sign but not necessarily a proof of the productive use of a pattern.

5. CS vs. CDS

Grammatical classes differ by the percentage of derivatives: generally, the percentage of derived adjectives in CDS is higher than that of nouns and verbs (see Table 13).

Table 13. Frequencies of derived noun, verb and adjective lemmas and tokens in relation to noun, verb and adjectives lemmas and tokens in input and output

	MOT F	CHI F	MOT K	CHI K
N lem./tok.	2,881/9,452	874/3,803	1,449/3,455	580/1,518
N DERIV lem./tok.	1,320/4,075	575/1,168	445/901	89/176
% N DERIV lem./tok.	46%/43%	66%/31%	31%/26%	15%/12%
V lem./tok.	2,161/7,347	910/1,920	1,385/3,269	412/1,237
V DERIV lem./tok.	1,132/2,890	382/660	563/1037	222/494
% V DERIV lem./tok.	53%/39%	42%/34%	41%/32%	54%/40%
ADJ lem./tok.	220/1629	143/930	234/810	164/279
ADJ DERIV lem./tok.	145/812	96/170	159/475	67/128
% ADJ DERIV lem./tok.	66%/ 50%	67%/18%	68%/59%	41%/46%

5.1 Input-output relationship in nominal derivation

In both CDS corpora the number of derived nouns is smaller than that of non-derived ones (in lemmas and tokens). It amounted to one third of all nouns in Kirill's input and about half in Philipp's (see Table 13). A comparison of derivatives in CDS and CS shows that there are twice as many derivatives in Kirill's CDS than in Kirill's speech (in lemmas and tokens). Philipp's mother has 20% fewer lemmas than Philipp and more tokens, but not many (12%). A comparison of the derivatives in the both CDS corpora shows that Philipp's mother uses them more than Kirill's mother, and accordingly more semantic categories (14 vs. 12). The relationship of derivatives in both CS corpora coincides. Thus, Philipp has more nominal derivatives of lemmas (four times more) and tokens (two and a half times more) than Kirill. There thus appeared to be more derivatives in whichever boy's input contained more of them.

It is also significant that a strong positive correlation has been revealed between the frequency of the semantic categories of derivatives in CDS and CS (see Table 14). The most frequent categories (besides diminutives, of which twice as many were recorded in Philipp's corpus) for both dyads, are activities and/or their results, instruments, objects, females. However, in the dyad of Philipp, baby animals also had a certain frequency, whereas in the dyad of Kirill, they were agents, locatives and singulatives (more details in Kazakovskaya 2019).

Table 14. Semantic categories of derived nouns (% lemmas/tokens among new derivatives)

	MOT F	CHI F	MOT K	CHI K
DIM	70/77	73/75	37/37	28/33
Activities/Results	6.2/4.3	6/4.4	18/15	16/13
Instruments	4.9/5.4	3/2.3	12/12	19/14.5
Objects (other)	5.2/3.7	5/4.4	7/10	9/7
Females	2.7/2.3	3.5/3.1	3/4	3/1.5
Baby animals	2.7/2.1	4/4.4	2/2	0
Agents	1.6/1.3	0.5/0.3	4.5/5	10/9
Abstract names	1.2/0.5	1/0.5	8/5	1/1
Locatives	1.6/0.8	1/0.5	4/7	9/8
Singulatives	1.2/0.95	2/4.2	4/4	4/5
Males	1.2/0.8	1/0.8	0.4/0.2	1/1
Subjects (other)	0.8/0.9			
Augmentives	0.2/0.1			

A distribution analysis revealed a consistent increase in the proportion of derivatives in CDS (fine-tuning), to which CS correlated. For example, peaks of derivatives were simultaneously found, at 2;3 and sequentially, at 2;10–2;11, in Kirill's corpus.¹⁹ An analysis of the frequency of derivatives in both CDS corpora can thus explain their sequence in CS: in CDS the most frequent patterns, semantic categories and their affixes²⁰ have a greater chance of entering CS.

5.2 Input-output relationship in adjective derivation

The proportion of derived adjectives among all adjective lemmas and tokens fluctuates around 50% and depends rather on the sample size than on the age of the child. Use of adjectives is always optional, whereas repeating them might be deliberate. Thus, we observe more quantitative differences in adjective tokens than in lemmas.

Denominal patterns predominate due to the large proportion of relational and possessive adjectives both in CDS and CS (see Table 15). The most frequent models are relational denominal adjectives expressing colors, possession and relation. The

19. It was symptomatic that compounds developed in a similar way, with both peaks (at 2;5 and 2;11) coinciding in CDS and CS. Such periods of significant increase in the frequency of using a specific language tool indicate the influence of input and shed light on the mechanisms of fine-tuning.

20. Thus, Filipp's mother uses 105 affixes, Filipp 60; Kirill's mother uses 53 affixes, Kirill 30.

most productive model is the derivation of antonyms, with the help of the prefix *ne-*: this prefix is the only one used with several stems both in CS and in CDS. Diminutive suffixes make up the second productive model.

Table 15. Semantic categories of derived adjectives (% lemmas/tokens to derived adjectives)

Semantic category	MOT F	CHI F	MOT K	CHI K
Relation	58/62	56/64	70/73	56/56
Possession	6/9	15/18	2/> 1	5/4
Degree of quality	12/18	16/7	7/16	9/20
Action	15/8	8/7	12/5	18/12
Antonym	6/1	4/4	9/4	12/8
Absence of feature	1/> 1	1/> 1	> 1/1	0

Individual differences between the two dyads are not great: Kirill and his mother use fewer possessive adjectives but a more diverse repertoire of antonyms (especially in lemmas). The proportion of most semantic categories is comparable: adjectives expressing a relationship to different objects prevail in all corpora. This might be due to the relatively small number of frequently used suffixes. Relational adjectives are unavoidable semantic specifiers that occur early. On the other hand, children start to use the less frequent possessive adjectives and antonyms due to their transparent semantics.

5.3 Input-output relationship in verb derivation

Verb derivation in CDS is more diverse than in CS: caregivers do not limit themselves to deverbal models, however, denominal and deadjectival verbs are not numerous. The proportion of patterns and models in CS mirrors that of CDS: rare models are absent in CS and the most frequent ones are dominant. This is especially true for the prefixed verbs. Another reason for this is that the meaning of the majority of aspectual suffixes is too abstract for the child.

Derived verbs make up 40–50% of lemmas and about 30–40% of tokens in CDS. The majority of them are prefixed verbs expressing different semantic and/or aspectual modifications of the motivating verb, e.g. the path in the verbs of motion, the beginning of an action in inchoatives (see Table 16).

With the exception of the reverse correlation in the proportion of derived lemmas, children demonstrate similar trends in the use of verb categories: the proportion of verb derivatives expressing temporal details of the action is about 12% of lemmas and 15–16% of all derived tokens. Reflexive verbs of different semantics

Table 16. Semantic categories of derived verbs (% lemmas/tokens to derived verbs)

Semantic category	MOT F	CHI F	MOT K	CHI K
Path of motion*	12/13	14/16	23/15	42/38
Phase of action	23/22	12/15	28/26	12/16
Reflexive meaning	19/21	22/20	12/16	18/23
Saturative meaning	1/1	> 1/1	3/2	8/7
Repetition of action	29/32	11/11	12/15	6/8
Resultative meaning	16/11	40/37	22/26	14/8

* All prefixes expressing the path of motion like going inside, getting outside, coming/arriving are merged into one semantic category, although these semantic variants are usually treated as separate models.

make up about 20% of all lemmas and tokens in both CS corpora. Filipp prefers verbs expressing resultative actions (*s-varit* ‘PFV-boil’ ← *varit* ‘boil.IPFV’ (2;2)), whereas Kirill uses many verbs stressing the path of motion (*pri-exat* ‘PFV-come by vehicle’ ← *exat* ‘go by vehicle’ (2;5)). These differences do not necessarily reflect the input: Kirill’s CDS contains even more resultatives than Filipp’s (22%/26% vs. 16%/11%), however, their share in Kirill’s speech remains scarce (14%/8%). A big proportion of verbs expressing repetitive actions in the speech addressed to Filipp is, probably, due to the fact that they are obligatory in the negative imperative: *ne otkry-vaj* ‘don’t open-IPFV.IMP’. Thus, children mirror the input indirectly by taking over frequent contexts, rather than frequent single forms.

CS and CDS differ in the following respects: children adopt most semantically transparent models. More abstract semantics, especially tiny grammatical distinctions between perfective and imperfective verbs, are acquired later and do not play as much of a role in CS as in CDS.

6. A brief overview of derivational morphology in the late stages of observation

Further development of derivation was traced only in Kirill’s larger data. In this period (4;2–5;10) nouns (697 lemmas/943 tokens) prevail in lemmas, and verbs (383 lemmas/1162 tokens) are more than twice as frequent as adjectives (158/345). This period is characterised by frequent use of derived adjectives: their percentage of the cumulative number of adjectives is 64% in lemmas and 52% in tokens. However, non-derived adjectives are repeated more frequently than derived ones. The share of tokens and lemmas for non-derived adjectives is 2.9, whereas for derived adjectives it is 1.8. A similar share is registered in verbs (5.8 vs. 1.8) and, with a less striking difference, in nouns (1.7 vs. 1.4). This probably means that Kirill feels more confident with elementary, non-derived words, especially among verbs and adjectives.

Derived nouns have not become any more frequent in CS and non-derived nouns still prevail. The number of word-building pairs (viz. 'N→N') and families is increasing. 3-member families appear: *volšebn(yj)* 'magic' → *volšebn-ik* 'magician', *volšeb-stv(o)* 'magic' (4;2). The proportion of deadjectival nouns is also increasing: *star(yj)* 'old' → *star-ik* 'old man', *star-ušk(a)* 'old lady-DIM' (4;2). New suffixes *-ec*, *-id*, *-arad*, *-ant* begin to be used. The percentage of diminutives decreases to 7–8% in some recordings, while other semantic categories are becoming more diverse. This is especially noticeable in the names of males. The latter include new agents, their professions or activities, nationalities, ages or memberships: *motocikl-ist* 'motorcyclist', *pešot-in/ec* 'infantryman', *japon-ec* 'japanese', *podrost-ok* 'teenager'. After 4;6 the nouns formed by prefixes (*super-dynj(a)* 'supermelon' 5;10)), or complex methods (namely by prefixation and suffixation, such as *na-kolen-nik* 'kneecap') and syntactic conversion (*raboč(ij)* 'working.N' (5;7), *upravljajušč(ij)* 'administering.N' (5;10)) are registered. This period also differs by the frequent use of abstract notions in nouns (*udivl-enij(e)* 'wonderment' (4;6), *skor-ost* 'speed' (5;8)) and adjectives (*ne-klassičesk(ij)* 'non-classic', *svet-sk(ij)* 'secular' (5;10)).

Starting with 4;3, a few derived child-created nouns (1 lemma)²¹ and adjectives (3 lemmas) built with suffixes were found: **podrezyva-tel'* 'undercutter' ← *podrezyva(t')* 'undercut.IPFV', **nem-sk(ij)* (4;3) instead of *nemec-k(ij)* 'German.ADJ' ← *nemec* 'German.N'. Some novel adjectives are colloquial variants of the ADS norm: *afganistan-sk(ij)* (5;10) instead of *afgan-sk(ij)* 'afgani'. In verbs they occur later and are restricted to the prefixed verb **pod-pomo(č')* 'help a bit' (5;10) ← *pomo(č')* 'help' and to the postfixed verb **razdumyva(t')-sja* 'be considering smth' (5;10) with reflexive semantics ← *razdumyva(t')* 'consider', cf. *zadum-yva(t')-sja* 'reflect'. Erroneous noun inflections (mostly of number) occur simultaneously with nominal neologisms, whereas erroneous verb forms occur earlier than verb novel child-created words.

Within verb derivation, development is demonstrated by the regularization of aspectual and voice oppositions. There are numerous 3-member families like *da(t')* 'give.PFV' → *da-va(t')* 'give-IPFV' (4;2) → *dava(t')-sja* 'be given-PASS' (5;7). Kirill overmarks the imperfective with a special suffix: **ukraš-iva(et)* instead of *ukraš(a(et))* 'decorate-IPFV'. These errors are registered only at 4;3 and disappear later, whereas novel words are created later on, at about 6.

21. Two novel child-created words (**babbl-ik* 'bubble-DIM', **gitar-nik* (instead of *gitar-ist*) 'guitar player') were documented earlier.

7. Conclusions

In CS nouns and adjectives are mainly formed from nouns, whereas verb derivatives would all be semantically predictable deverbal derivations due to the formation of aspectual pairs and reflexivisation by semiclititic *-sja*. With nouns and adjectives, suffixation is the main method of word formation whereas for verbs it is prefixation. Prefixes are used for semantic changes, whereas suffixation serves to change grammatical features of verbs.

We can trace two phases in the use of derivatives by children. The first is characterised by sporadic frozen forms. In the second phase both subjects start to follow certain patterns and establish connections between derivatives and their motivating stems. Abstract grammatical distinctions occur in CS later than perceptually demonstrable semantic ones. This probably means that children pick up derivatives without detailed understanding of what they exactly mean. Incorrectly chosen suffixes or stems indicate creative use of a pattern.

The preference for certain patterns is the same for both children in all grammatical classes with reference to CDS and ADS (Švedova 1980). Also, both subjects show a similar development of semantic categories, a similar percentage of derived nouns and verbs among all corresponding tokens (which is also a common tendency in CDS and ADS) and an equally small amount of novel child-created words (in lemmas).

However, there are some individual differences. In Filipp's speech the percentage of derived verbs among all verbs is almost the same as the corresponding percentage of derived nouns, whereas Kirill has almost twice as many derived verbs than nouns. Filipp has a bigger inventory of noun suffixes occurring also in frequent child-created diminutive noun tokens. Kirill has two main models for them (diminutive and agents) and his new nominal derivatives increase towards the end of the observation.

The comparison between CDS and CS shows that the children "mirror" the share of the most frequent and productive patterns of mothers, i.e. that the frequent categories prevail and that the rare models may be absent. There is a positive correlation between CDS and CS in the frequency and diversity of semantic categories. There is also a prevalence of denominal suffix models and semantic categories of objects, as well as activities/results, instruments and agents. Derivational patterns in verbs and adjectives lag behind nominal ones. In the later phases we could also observe a more intensive development of inflectional oppositions in non-derived adjectives and verbs than in derived ones, as if children keep to a certain cumulative (inflectional and derivational) level of morphological complexity. The patterns of productive and morphosemantically transparent derivatives are the first to emerge in all grammatical classes.

References

- Argus, R. & Kazakovskaya, V. V. 2013. Acquisition of compounds in Estonian and Russian: frequency, productivity, transparency and simplicity effect. *Eesti Rakenduslingvistika Uhingu Aastaraamat* 9: 23–42. <https://doi.org/10.5128/ERYa9.02>
- Argus, R. & Kazakovskaya, V. V. 2018. Acquisition of noun derivations in Estonian and Russian L1. *Eesti Rakenduslingvistika Uhingu Aastaraamat* 14: 23–39. <https://doi.org/10.5128/ERYa14.02>
- Argus, R., Kazakovskaya, V. V. & Laalo, K. 2013. Acquisition of compounds in typologically different languages: Evidence from Estonian, Russian, and Finnish. In *XIII Congress of IASCL*. Amsterdam: University of Amsterdam.
- Baudouin de Courtenay, J. N. I. 1904/1963. Jazykoznanie (Linguistics). In *Izbrannye trudy po jazykoznaniju* (Selected works in linguistics), Vol. 2, J. N. I. Baudouin de Courtenay (ed.), 96–117. Moscow: Izdatel'stvo AN SSSR.
- Bauer, L. & Valera, S. 2005. Conversion or zero-derivation: An introduction. In *Approaches to Conversion/Zero-derivation*, L. Bauer & V. Salvador (eds), 7–19. Münster: Waxmann.
- Ceitlin, S. N. 1989. Detskaja reč: Innovacii formoobrazovanija i slovoobrazovanija (Child speech: Inflectional and derivational innovations). Dr. Sc. dissertation, Leningradskij gosudarstvennyj pedagogičeskij institut.
- Ceitlin, S. N. 2009. *Očerki po slovoobrazovaniju i formoobrazovaniju v detskoj reči* (Essays on derivation and inflection in child speech). Moscow: Znak.
- Clark, E. V. 2009. *First Language Acquisition*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511806698>
- Dressler, W. U., Sommer-Lolei, S., Korecky-Kröll, K., Argus, R., Dabašinskienė, I., Kamandulytė-Merfeldienė, L., Ijäs, J. J., Kazakovskaya, V. V., Laalo, K. & Thomadaki, E. 2019. First-language acquisition of synthetic compounds in Estonian, Finnish, German, Greek, Lithuanian, Russian and Saami. *Morphology* 29(3): 409–429. <https://doi.org/10.1007/s11525-019-09339-0>
- Eliseeva, M. B. 2015. *Stanovlenie individual'noj jazykovoju sistemy rebenka* (Establishment of the individual language system of a child). Moscow: Jazyki slavjanskix kul'tur.
- Gvozdev, A. N. 1961. *Voprosy izučeniija detskoj reči* (Issues of the investigation of children's speech). Moscow: Izdatel'stvo Akademii pedagogičeskix nauk RSFSR.
- Jakobson, R. O. 1957[1971]. Shifters, verbal categories, and the Russian verb. In *Selected writings*, Vol. 2, R. O. Jakobson (ed.), 130–148. The Hague: Mouton.
- Janda, L., Endersen, A., Kuznetsova, J., Lyashevskaya, O., Makarova, A., Nessel, T. & Sokolova, S. 2013. *Why Russian Aspectual Prefixes Aren't Empty: Prefixes as Verb Classifiers*. Bloomington IN: Slavica.
- Jurjeva, N. M. 2006. *Problemy rečevogo ontogeneza: Proizvodnoe slovo. Dialog. Eksperimental'nye issledovanija* (Problems of speech ontogenesis: A derivative. A dialogue. Experimental studies). Moscow: Institut jazykoznanija RAN.
- Kazakovskaya, V. V. 2017. Acquisition of nominal compounds in Russian. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketzrez & M. Kilani-Schoch (eds), 63–90. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.04kaz>
- Kazakovskaya, V. V. 2019. Child nominal derivation and parental input: Evidence from morphology-rich Russian. *International Journal of Cognitive Research in Science, Engineering and Education* 7(1): 121–129. <https://doi.org/10.5937/IJRSEE1901121K>

- Kubrjakova, E. S. 1981. *Tipy jazykovyx značenij. Semantika proizvodnogo slova* (Types of linguistic meanings. Semantics of derivatives). Moscow: Nauka.
- Leontjev, A. A. 1965. *Slovo v rečevoj dejatel'nosti* (Word in speech activity). Moscow: Nauka.
- Lepskaja, N. I. 1997. *Jazyk rebenka* (Child language). Moscow: Izdatel'stvo MGU.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Manova, S. & Dressler, W. U. 2005. The morphological technique of conversion in the inflecting-fusional type. In *Approaches to Conversion/Zero-derivation*, L. Bauer & S. Valera (eds), 67–102. Münster: Waxmann.
- Negnevickaja, E. I. & Šaxnarovič, A. M. 1981. *Jazyk i deti* (Language and children). Moscow: Nauka.
- Protassova, E. J. & Voeikova, M. D. 2007. Diminutives in Russian at the early stages of acquisition. In Savickienė & Dressler (eds), 43–73. <https://doi.org/10.1075/lald.43.03pro>
- Savickienė, I. & Dressler, W. U. (eds) 2007. *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Saxarnyj, L. V. 1979. *Slovoobrazovanie v rečevoj dejatel'nosti* (Word formation in speech activity). PhD dissertation, Leningradskij gosudarstvennyj universitet.
- Švedova, N. J. (ed.) 1980. *Russkaja grammatika T. 1* (Russian grammar V. 1). Moscow: Nauka.
- Tambovceva, A. G. 1983. *Formirovanie sposobov slovoobrazovanija u detej doškol'nogo vozrasta v detskom sadu* (Developing of word-formation methods at preschoolers in the kindergarten). PhD dissertation, Moskovskij gosudarstvennyj universitet.
- Tribushinina, E., Voeikova, M. D. & Nocchetti, S. (eds) 2015. *Semantics and Morphology of Early Adjectives in First Language Acquisition*. Newcastle upon Tyne: Cambridge Scholars.
- Ušakova, T. N. 2004. *Reč: istoki i principy razvitija* (Speech: origins and development principles). Moscow: PER SE.
- Voeikova, M. D. 2015. *Stanovlenie imeni: Rannie etapy usvoenija det'mi imennoj morfologii russkogo jazyka* (Establishment of Name: Early phases of the acquisition of nominal morphology by children). Moscow: Jazyki slavjanskix kul'tur.
- Vygotskij, L. S. 1935. *Umstvennoe razvitie detej v processe obučenija* (Mental development of children in the process of training). Moscow: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo.
- Xarčenko, V. K. & Ozerova, E. G. 1999. *Složnye slova v detskoj reči* (Compounds in child language). Belgorod: Belgorodskij universitet.

The acquisition of the Lithuanian derivational system

Laura Kamandulytė-Merfeldienė¹, Ingrida Balčiūnienė^{1,2}
and Ineta Dabašinskienė¹

¹Vytautas Magnus University / ²St. Petersburg State Pediatric
Medical University

This study is based on the longitudinal corpus data of two Lithuanian children (1;7–2;7). The aim of the research is to analyse the production of suffixed and prefixed nouns, verbs, and adjectives in CS and CDS during the early stages of the acquisition of Lithuanian morphology. The acquisition of the Lithuanian derivational system seems to be quite an effortless process. Although a considerable number of errors occurred at the initial stages, the first derivational families emerged already at the age of 1;8–1;9, and the number of errors started to decrease rapidly. The period during which the first derivatives emerge in child speech corresponds to the transition from a premorphological to a protomorphological stage.

Keywords: Lithuanian derivational morphology, verb, noun, prefix, suffix

1. Introduction

Lithuanian belongs to the Baltic branch of the Indo-European language family and is characterized as a morphologically rich and highly inflected language. Endings of words are mostly fusional, and they are the principal means of marking syntagmatic relations between words in a sentence and/or relation between word forms in a paradigm (Ambrasas 1997). In some previous studies on written Lithuanian, a huge discrepancy between the word-formation possibilities of nouns and verbs was observed, i.e. noun derivation employs a large set of diverse affixes, while verb derivation accepts only a very limited number of affixes (Klimas 1975, 1991; Rimkutė, Kazlauskienė & Utkā 2016; Stundžia 2016). Considering the contrast in the semantic range and distribution of derivational affixes in noun and verb formation, it is assumed that the acquisition of the Lithuanian derivational system

presents an interesting object for research. Previous research on the acquisition of verb (Wójcik 2000) and noun derivatives, including diminutives and compounds (Savickienė 2003, 2007; Dabašinskiė & Kamandulytė-Merfeldienė 2017), presents a solid basis for a closer analysis of word-formation areas.

2. Derivational morphology in Lithuanian

In Lithuanian, complex words are mainly formed by means of derivation and composition, the former being much more productive than the latter. Among the means of derivation, suffixed derivatives are the most typical, while prefixed and circumfixal (prefixed-suffixed) derivatives, as well as conversions (inflectional changes), are rarer in the word-formation system of Lithuanian (Stundžia 2016).

In the nominal word-formation system, suffixed derivatives are much more frequent than prefixed ones; in verbal derivation, however, prefixed derivation prevails (Rimkutė, Kazlauskienė & Utkā 2016). In inflectable derived words, suffixes and prefixes, as the main derivation formants, are accompanied by inflections which usually differ from the inflectional paradigm of the base words and, thus, serve as a secondary means of derivation, e.g. *rank-a* ‘an arm’ → *rank-ov-ė* ‘a sleeve’ and *apy-rank-ė* ‘a bracelet’. In word forms (especially in verbs), affixation often causes morphonological changes in the root (vowel or consonant alternation), e.g. *sek-ti* ‘to narrate’ → *sak-y-ti* ‘to say’, *bes-ti* ‘to pierce’ → *bad-y-ti* ‘to prick’ (Ambrazas 1997; Urbutis 2009; Stundžia 2016).

The following subsections provide a brief description of noun, verb, and adjective formation, with the focus on suffixation and prefixation.

2.1 Noun formation

Suffixed derivatives comprise about 600 word-formation types unequal in productivity. Prefixed derivatives are much less frequent and include 24 word-formation types; among them, the prefix *pa-* covering about one third of prefixed nouns is the most productive prefix. Nominal derivatives are classified as denominal, de-adjectival, and deverbal nouns (Ambrazas 1996, 1997; Keinys 1999; Urbutis 2009; Stundžia 2016).

2.1.1 *Denominal nouns*

Denominal nouns fall into several typical semantic groups. Denominal personal nouns (with the exception of some prefixed ones) are derived by means of more than 50 suffixes. Among them, *-iet-is* (e.g. *miest-iet-is* ‘a town-dweller’ from *miest-as* ‘a town’), *-inink-as* (e.g. *dail-inink-as* ‘an artist’ from *dail-ė* ‘an art’), *-ist-as* (e.g.

taks-ist-as ‘a taxi-driver’ from *taks-i* ‘a taxi’), *-ik-as* (e.g. *elektr-ik-as* ‘an electrician’ from *elektr-a* ‘an electricity’), and *-ant-as* (e.g. *muzik-ant-as* ‘a musician’ from *muzik-a* ‘a music’) are the most productive suffixes.

Denominal place nouns are derived by means of ca. 17 prefixes and 17 suffixes. The most productive prefixes include *pa-* (e.g. *pa-jūr-is* ‘a seashore’ from *jūr-a* ‘a sea’), *už-* (e.g. *už-jūr-is* ‘an oversea’), and *prie-* (e.g. *prie-miest-is* ‘a suburb’ from *miest-as* ‘a town’). Usually the prefixed derivation is accompanied by changes of the declension class. The most productive suffixes are *-yn-as* (e.g. *puš-yn-as* ‘a pine forest’ from *puš-is* ‘a pine tree’) and *-in-ė* (e.g. *kav-in-ė* ‘a café’ from *kav-a* ‘a coffee’). Denominal diminutives are derived by means of ca. 15 suffixes. Among them, *-el-is*, *-el-ė* (e.g. *nam-el-is* ‘a small house’ from *nam-as* ‘a house’) and *-ėl-is*, *-ėl-ė* (e.g. *pyrag-ėl-is* ‘a pastry’ from *pyrag-as* ‘a cake’) are the most productive suffixes (Ambrasas 1996; Urbutis 2009; Stundžia 2016). For information on the acquisition of Lithuanian diminutives, see Savickienė (2003, 2007).

Denominal collective nouns (e.g. *gyv-ūn-ij-a* ‘an animal kingdom’ from *gyv-ūn-as* ‘an animal’), status nouns (e.g. *draug-yst-ė* ‘a friendship’ from *draug-as* ‘a friend’), matrimonial status nouns (e.g. *Kazlausk-ien-ė* ‘a wife of Kazlauskas’ from *Kazlausk-as*), sort/origin nouns (e.g. *kiaul-ien-a* ‘a pork’ from *kiaul-ė* ‘a pig’, *uog-ien-ė* ‘a jam’ from *uog-a* ‘a berry’), origin/place of residence nouns (e.g. *kaun-iet-is* ‘a person from/living in Kaunas’ from *Kaun-as*) should also be mentioned.

2.1.2 Deadjectival nouns

Deadjectival nouns comprise mainly quality nouns (these are formed by adding ca. 15 suffixes); among them, *-um-as* (e.g. *ger-um-as* ‘a kindness’ from *ger-as* ‘kind’) and *-yb-ė* (e.g. *lyg-yb-ė* ‘an equality’ from *lyg-us* ‘equal’) are the most productive. Much rarer are personal nouns (e.g. *sen-ol-is* ‘an old man’ from *sen-as* ‘old’), collective nouns (e.g. *bendr-uomen-ė* ‘a community’ from *bendr-as* ‘common’), and place nouns (e.g. *aukšt-um-a* ‘a hill’ from *aukšt-as* ‘high’) (Stundžia 2016).

2.1.3 Deverbal nouns

Deverbal nouns fall into the following semantic groups. Action nouns (one of the most plentiful word-formation categories with respect to the number of derivatives it has) are formed by means of more than 50 suffixes. The suffix *-im-as/-ym-as* (e.g. *ėj-im-as* ‘walking’ from the stem of past tense form *ėj-o* (walk-PST.3)¹ ‘went’) is the most productive (Stundžia 2016). Result nouns are formed by adding more than 40 suffixes, but only some of them are productive (e.g. *krov-in-ys* ‘a load’ from the stem of past tense form *krov-ė* (load-PST.3) ‘loaded’). Agent nouns can be formed by

1. Singular and plural verbs in the form of 3rd person are identical and can be recognized only according to the context.

adding ca. 50 suffixes; the most productive ones are *-toj-as*, *-toj-a* (e.g. *mok-y-toj-as* ‘a teacher’ from *mok-y-ti* ‘to teach’), *-ėj-as*, *-ėj-a* (e.g. *kūr-ėj-as* ‘a creator’ from the stem of past tense form *kūr-ė* (create-PST.3) ‘created’ and *-ėl-is*, *-ėl-ė* (e.g. *pa-bėg-ėl-is* ‘a runaway’ from *pa-bėg-ti* ‘to run away’). Instrument nouns are derived by means of more than 30 suffixes (e.g. *krov-ikl-is* ‘a charger’ from the stem of past tense form *krov-ė* (charge-PST.3) ‘charged’, *trint-uk-as* ‘an eraser’ from *trint-ti* ‘to erase’). Such rare semantic groups as place nouns (e.g. *leid-ykl-a* ‘a publishing house’ from the stem of past tense form *leid-o* (publish-PST.3) ‘published’, celebration/feast nouns (e.g. *vest-uv-ės* ‘a wedding’ from the stem of infinitive *ves-ti* ‘to marry’) may also be mentioned (Stundžia 2016).

2.2 Verb formation

Lithuanian verbs are derived by adding 14 highly productive prefixes and nine basic suffixes, all of them having cumulatively ca. 80 circumfixal (prefixed-suffixed) variants (Jakaitienė 1973; Ambrazas 1996; Keinys 1999; Urbutis 2008b, 2009; Stundžia 2016).

2.2.1 Denominal verbs

Denominal verbs are derived by adding seven basic suffixes. They fall into the following semantic groups: factitive verbs (e.g. *konspekt-uo-ti* ‘to take notes’ from *konspekt-as* ‘notes’), verbs of performance (e.g. *karal-iau-ti* ‘to reign’ from *karal-ius* ‘a king’), inchoative verbs (e.g. *gar-uo-ti* ‘to evaporate’ from *gar-as* ‘a vapour’), instrumental verbs (e.g. *bur-iuo-ti* ‘to sail’ from *bur-ė* ‘a sail’), verbs with the meaning ‘to look for/to collect’ (e.g. *gryb-au-ti* ‘to gather mushrooms’ from *gryb-as* ‘a mushroom’), verbs with the meaning ‘to participate’ (e.g. *puot-au-ti* ‘to feast’ from *puot-a* ‘a feast’), and verbs with the meaning ‘to be consumed/gripped/seized’ (e.g. *liepsn-o-ti* ‘to flame’ from *liepsn-a* ‘a flame’) (Stundžia 2016).

2.2.2 Deadjectival verbs

Deadjectival verbs are derived by adding seven basic suffixes and belong to the following semantic categories: factitive verbs (e.g. *ger-in-ti* ‘to improve’ from *ger-as* ‘good’), verbs of performance (e.g. *švepl-uo-ti* ‘to lisp’ from *švepl-as* ‘lispings’), and inchoative verbs (e.g. *pilk-ė-ti* ‘to turn grey’ from *pilk-as* ‘grey’) (Stundžia 2016).

2.2.3 Deverbal verbs

Deverbal verbs are much more frequent than denominal verbs and, especially, than deadjectival verbs. Deverbal verbs can be derived by adding both prefixes and suffixes. Prefixes usually express that the action, process or state denoted by the base verb was/will be started (e.g. *už-dain-uo-ti* ‘to start singing’ from *dain-uo-ti* ‘to

sing') or finished (e.g. *iš-mes-ti* 'to throw out' from *mes-ti* 'to throw') and highlight specific characteristics of the action expressed by the base word (e.g. *bėg-ti* 'to run' → *i-bėg-ti* 'to run into', *iš-bėg-ti* 'to run out', *nu-bėg-ti* 'to run from', *pa-bėg-ti* 'to run away from', *par-bėg-ti* 'to run back', etc.).

2.3 Adjective formation

Suffixed derivatives with ca. 130 suffixes (most of them low in productivity) dominate among all adjective formations (Ambrazas 1996; Keinys 1999; Urbutis 2008a, 2009; Stundžia 2016).

2.3.1 Denominal adjectives

In denominal adjectives, suffixation is a common device, while prefixation is rather limited. The majority of prefixed adjectives consist of privative adjectives with the prefix *be-* (e.g. *be-darb-is* 'unemployed' from *darb-as* 'a work'). The majority of suffixed denominal adjectives consists of relational adjectives with the highly productive suffix *-in-is*, *-in-ė* (e.g. *darb-in-is* 'work-related' from *darb-as* 'a work') and qualitative adjectives (e.g. *muil-uot-as* 'soapy' from *muil-as* 'a soap') (Stundžia 2016).

2.3.2 Deadjectival adjectives

Deadjectival adjectives are formed by adding both prefixes and suffixes. They fall into the following semantic groups: attenuation adjectives (e.g. *apy-balt-is* 'whitish' from *balt-as* 'white'), evaluative adjectives (e.g. *balt-ut-is* 'very white' from *balt-as* 'white'), and adjectives with the opposite meaning (e.g. *ne-ger-as* 'not good' from *ger-as* 'good') (Stundžia 2016).

2.3.3 Deverbal adjectives

Deverbal adjectives are rarer than denominal and deadjectival ones. They are derived by means of semi-productive suffixes to express both active (e.g. *taik-l-us* 'well-aimed' from the stem of past tense form *taik-ė* (aim-PST.3) 'aimed') and passive (e.g. *tamp-r-us* 'elastic' from *temp-ti* 'to pull') meanings.

3. Acquisition data and methodology

The present study aims to explore the production of suffixed and prefixed nouns, verbs, and adjectives in child speech (CS) and child-directed speech (CDS) during the first stages of the acquisition of Lithuanian morphology. It is based on the longitudinal spontaneous speech corpus data (~45 hours) of two Lithuanian children (a girl Monika and a boy Elvijas) at the ages of 1;7–2;7 (Table 1). Their families are

of a higher socio-economic status; during the period of data compilation, each child was the only child in the family; their parents were students at Lithuanian state universities. The children are monolingual native speakers of Standard Lithuanian without any visual, hearing, intelligence or language impairment, i.e. they represent the typical development of the Lithuanian language.

Table 1. The scope of analysed data

Child's name	Child's age	Total duration of recordings	Number of words (tokens)
Monika (MON)	1;8–2;7	~25 hours	26,449
Elvijas (ELV)	1;7–2;6	~20 hours	28,846

Conversations between the children and their caregivers were recorded at home or another common environment trying to cover as many different situations as possible: games, cooking, eating, bathing, etc. The conversations were recorded by parents at different times of the day, which mostly depended on the child's willingness to communicate. Most of the recordings were interactions between the children and their mothers, some include children's dialogues with their fathers. Our data mainly consists of the dialogues, but few conversations between more than two participants (e.g. the child, his/her mother and father) were recorded as well.

The recordings were grouped by months (trying to maintain a similar size within the material of each month), transcribed, and encoded morphologically by means of the Lithuanian version (Dabašinskienė & Kamandulytė 2009) of CHAT (MacWhinney 2000). Noun, verb, and adjective formations were encoded additionally for an automatized linguistic analysis.

The distribution of the suffixed and prefixed nouns, verbs, and adjectives was analysed in both CS and CDS.

4. Results

4.1 First occurrence

This section discusses the timeline when the Lithuanian children start using derivatives spontaneously as well as applying derivational rules. Previous studies have revealed that Lithuanian children take the first steps in the acquisition of inflectional and derivational systems very early. This concerns, for instance, the first oppositions within the rich and productive diminutive system (Savickienė 2007) and diminutive-simplex oppositions emerged at the same time as inflectional form oppositions, i.e. at the transition phase from pre- to protomorphology (around

1;9–1;10; see Savickienė 2003; Dabašinskienė & Kamandulytė-Merfeldienė 2017). In Lithuanian, noun–noun compounds emerge later, i.e. a few months after the beginning of the protomorphological phase and, therefore, differ greatly from the early emergence of noun inflection, diminutive formation, and verb inflection, as the first spontaneous occurrences of inflection and diminutive formation are already observed in the pre-morphological phase (Wójcik 2000; Dabašinskienė & Kamandulytė-Merfeldienė 2017).

The results of the current study confirm the prediction that diminutives are acquired prior to other word formations (Table 2). In Monika's speech, the first emergence of spontaneous diminutives occurred at the age of 1;8 (*butel-iuk-as* (bottle-DIM-NOM.SG) 'a bottle') and in Elvijus's speech it occurred at 1;7 (*kamuol-iuk-as* (ball-DIM-NOM.SG) 'a ball'). Compounds occurred three months later, i.e. at 1;11 in Monika's speech and at 1;10 in Elvijus's speech.

Table 2. The timeline of derivative emergence in CS

Child	Diminutives	Compounds	Derived nouns		Derived verbs		Derived adjectives	
			Pref.	Suff.	Pref.	Suff.	Pref.	Suff.
ELV	1;7	1;10	1;7	1;8	1;7	1;8	–	1;9
MON	1;8	1;11	1;11	1;10	1;9	1;11	–	2;2

The first derivatives in Elvijus's speech were noted early, at the same time as diminutives, i.e. at 1;7 (Table 2). At this period, spontaneously used prefixed deverbial noun *p̄a-sak-a* 'a tale' (← *p̄a-sek-ti* ← *sek-ti* 'to tell') and a few prefixed deverbial verbs, e.g. *p̄a-ke(l)-k* (PREFIX-pick-IMP.2SG) 'pick up' ← *kel-k* (pick-IMP.2SG) 'pick up', *p̄a-d̄e-k* (PREFIX-put-IMP.2SG) 'put on' ← *d̄e-k* (put-IMP.2SG) 'put', *at-važ-iav-om* (PREFIX-come-PST-1PL) 'came' ← *važ-iav-om* (go-PST-1PL) 'went' were observed in the boy's speech. It is important to mention that underived forms (e.g. *važ-iav-om* (go-PST-1PL) 'went') were also observed at the same time as the derived ones. One month later (1;8) the boy spontaneously used the suffixed nouns *arbat-in-uk-as* 'a tea-pot' (← *arbat-a* 'a tea'), *bagaž-in-ė* 'luggage/baggage carrier' (← *bagaž-as* 'a luggage') and the suffixed verb *skamb-in-a* call-PRS.3 'calls' ← the stem of present tense form *skamb-a* (ring-PRS.3) 'rings'.

In Monika's speech, derivative verbs and derivative nouns emerged later than diminutives (Table 2). At 1;9 several spontaneously used prefixed deverbial verbs, e.g. *iš-im-k* (PREFIX-take-IMP.2SG) 'take of' ← *im-k* (take-IMP.2SG) 'take', and *at-ei-k* (PREFIX-go-IMP.2SG) 'come' ← *ei-k* (go-IMP.2SG) 'go', were observed in the girl's speech, and only at 1;11 the first spontaneously uttered suffixed verb with a prefix and a suffix was noticed, *p̄a-kab-in-ti* 'to hang on' (← *kab-in-ti* 'to hang on' ← *kab-o*

(hang-PRS.3) ‘hangs’). The first derivative noun occurred at 1;10, when one-word token of the suffixed deverbial noun *tep-al-as* ‘a cream’ (← *tep-ti* ‘to spread’) was used by the girl. At that time girl used already trisyllabic words, e.g. *balkonas* ‘balcony’. As in the case of verbs, at 1;11 the first occurrence of both the prefix and the suffix in noun derivation was noticed, i.e. *par-duo-t-uv-ė* ‘a shop’ (← *par-duo-ti* ‘to sell’ ← *duo-ti* ‘to give’).

Adjective derivatives appeared later than noun and verb derivatives. In Elvijos’s speech, they occurred at 1;9 (almost at the same time as compounds; these emerged at 1;10). In Monika’s speech, the first adjective derivatives occurred even later, at 2;2 (three months after the emergence of compounds). In Elvijos’s speech, the first spontaneous adjective derivative was the suffixed denominal adjective *spalv-ot-as* ‘colourful’ (← *spalv-a* ‘a colour’), while in Monika’s speech it was the suffixed de-adjectival adjective *murz-in-as* ‘dirty’ (← *murz-ius* ‘a dirty’). Previous studies have shown that derivative adjectives in child speech and child-directed speech occur rarely (Kamandulytė 2009; Kamandulytė-Merfeldienė 2015), and this explains why only few examples were observed during the period studied.

It can be stated that, as is the case with diminutives, in the child data (Elvijos 1;7, Monika 1;9) the prefixed verbs appear very early, during the premorphological stage. In this stage, the children very often use the premorpheme *a* instead of the relevant prefix, e.g. *a-dė-k* [= *pa-dė-k*] (PREFIX-put-IMP.2SG) ‘put on’ ← *dė-k* (put-IMP.2SG) ‘put’ (MON 1;9); *a-dar-y-k* [= *ati-dar-y-k*] (PREFIX-open-IMP.2SG) ‘open’ ← *dar-y-k* (open-IMP.2SG) ‘open’ (ELV 1;7). However, correct instances are observed as well, e.g. prefixed verbs *pa-kel-k* (PREFIX-pick-IMP.2SG) ‘pick up’ (← *kel-k* (pick-IMP.2SG) ‘pick’); *pa-dė-k* (PREFIX-put-IMP.2SG) ‘put on’ (← *dė-k* (put-IMP.2SG) ‘put’); *at-važ-iav-om* (PREFIX-go-PST-1PL) ‘came’ ← *važ-iav-om* (go-PST-1PL) ‘went’ (ELV 1;7); *iš-im-k* (PREFIX-take-IMP.2SG) ‘take from’ (← *im-k* (take-IMP.2SG) ‘take’); *at-ei-k* (PREFIX-go-IMP.2SG) ‘come’ ← *ei-k* (go-IMP.2SG) ‘go’ (MON 1;9). The first utterances allow us to assume that already at the stage of premorphology children start to differentiate the function of prefixes and to mark the category of aspect, which is closely related to the grammatical past tense.

The following aspects should be discussed at this point. At 1;7, the boy’s father says the one-word utterance *važ-iuoj-am* (go-PRS-1PL) ‘go’, but the boy, having in mind the already accomplished action, produces also the prefixed verb *at-važ-iav-om* (PREFIX-go-PST-1PL) ‘came’. At 1;10, the girl asks the mother to give her a t-shirt, *duo-k* (give-IMP.2SG) ‘give’, and repeats again her request with the prefixed verb when the mother did not react, *pa-duo-k* (PREFIX-give-IMP.2SG) ‘give’. An important fact to emphasise is that in CDS we observe some verbs with prefixes used more often than the simplex: *at-ei-k* (PREFIX-go-IMP.2SG) ‘come’ (99 tokens in the girl’s CDS, but 61 tokens of the simplex *ei-k* go-IMP.2SG ‘go’). In the boy’s CDS, these numbers are 250 and 118 tokens respectively; thus, we assume that the

early emergence of the prefixed verbs in child speech is due to a frequent usage of prefixed verbs in their CDS. Moreover, prefixed verbs were frequently used in both CDS and CS during the entire period of observation. Among verbs, tokens of derivatives represent 24% of all verbs in Monika's speech and 30% in her input. As for Elvijus, derivatives make up 33% in his speech and 31% in his input (Table 3). However, only approximately 1% of all verb derivatives are suffixed; the majority of them are prefixed derived verbs. The infrequent usage of noun and adjective derivatives is demonstrated both in CS and CDS in Table 3. Tokens of all noun derivatives (suffixed + prefixed) represent only 1% of all nouns in Monika's speech and 1% in her input. In Elvijus's corpus, tokens of noun derivatives represent 2% of all noun tokens in CS and 1% in CDS. Among adjectives, the percentage of derivatives is higher: in Monika's corpus, tokens of adjective derivatives represent 4% of all adjective tokens in CS and 5% in CDS; in Elvijus's corpus, they represent 9% of all adjective tokens in CS and 7% in CDS.

Table 3. Frequencies of derived noun, verb and adjective lemmas and tokens in relation to all noun, verb and adjective lemmas and tokens in input and output (the whole corpus)

Speaker	MON CS	MON CDS	ELV CS	ELV CDS
N lemmas/tokens	704/7661	1508/16946	974/7739	1336/7718
N DERIV lemmas/tokens	32/86	51/167	52/166	50/100
% N DERIV lemmas/tokens	5/1	3/1	5/2	4/1
V lemmas/tokens	716/8336	1358/24523	715/6811	948/10141
V DERIV lemmas/tokens	562/2034	1183/7407	486/2234	592/3179
% V DERIV lemmas/tokens	78/24	87/30	68/33	62/31
ADJ lemmas/tokens	60/570	124/1815	75/785	143/942
ADJ DERIV lemmas/tokens	10/25	15/85	19/67	18/69
% ADJ DERIV lemmas/tokens	17/4	12/5	25/9	13/7

The same frequency tendencies can be seen when different lemmas of derivatives are compared. The highest variety was observed in verbs. In Monika's speech, lemmas of verb derivatives comprise 78% among all verb lemmas, in her input – 87% (Table 3). In Elvijus's speech and his input the numbers are 68% and 62%, respectively. The variety of verb derivatives in CS and CDS is great, mainly due to a wide range of prefixes, including the most frequent prefix *pa-* and the frequent prefixes *i-* and *nu-*. To a large extent, these verb derivatives in the children's speech reflect aspectual relations and indicate the action, process or state that is finished. The examples include *mes-ti* 'to throw' → *iš-mes-ti* 'to throw out', *dė-ti* 'to put' → *pa-dė-ti* 'to put on', *skamb-in-ti* 'to call' → *pa-skamb-in-ti* 'to call') and mark the direction (e.g. *ei-ti* 'to go' → *at-ei-ti* 'to come', *lip-ti* 'to climb' → *i-lip-ti* 'to climb/ to come into').

Adjective derivatives were less variable than verb derivatives (in Monika's data, CS – 17%, CDS – 12% of all adjective lemmas; in Elvijus's data, CS – 25%, CDS – 13%). Adjective derivatives in CS and CDS are usually formed from nouns with suffixes (e.g. *vais-ius* 'a fruit' → *vais-in-is* 'fruity', *juok-as* 'a laugh' → *juok-ing-as* 'funny'); the most frequent of these suffixes is *-in-*. In most cases, these derivatives can be attributed to qualitative adjectives (e.g. *purv-in-as* 'dirty') and denote an external property; other derivatives denote an internal property (e.g. *laim-ing-as* 'happy').

Noun derivatives (diminutives were excluded from the analysis) were even less variable than adjective derivatives (in Monika's data, CS – 5%, CDS – 3% of all noun lemmas; in Elvijus's data, CS – 5%, CDS – 4%). Most of them are derived from nouns and verbs by adding various suffixes (e.g. *arbat-a* 'a tea' → *arbat-in-uk-as* 'a teapot', *aus-is* 'an ear' → *aus-in-ės* 'headphones', *kirp-ti* 'to cut' → *kirp-ykl-a* 'a hairdresser's', *lyg-in-ti* 'to iron' → *lyg-in-t-uv-as* 'an iron') and much less often by means of prefixes (e.g. *baig-ti* → *pa-baig-ti* 'to finish' → *pa-baig-a* 'an end', *vart-ai* 'a gate' → *j-vart-is* 'a goal'), with the prefix *pa-* predominating. Almost all noun derivatives in CS fall into a few semantic groups: professions/duties (e.g. *futbol-inink-as* 'a football player'), household items (e.g. *arbat-in-uk-as* 'a teapot'), places (e.g. *par-duo-t-uv-ė* 'a shop'), clothes (e.g. *mezg-t-in-is* 'a sweater'), and food (e.g. *skreb-ut-is* 'a toast'). No abstract nouns were found in CS, even though they are most frequent among all noun derivatives in the adult-directed speech (Ambrazas 1996).

When noun derivatives are compared with other noun formations, it appears that the variety of diminutives is much greater both in CS and CDS, whereas compounds are less variable than noun derivatives (Table 4). In Monika's speech, lemmas of noun derivatives account for 5% of all noun lemmas, while compounds – for only 0.1%. In Elvijus's speech, noun derivatives make up 5% of all noun lemmas, compounds – 2%, and diminutives 22%. The same tendencies are observed in CDS.

Table 4. The variety of noun formations (% of all noun lemmas)

Speaker		Diminutives	Derivatives	Compounds
MON	CS	No data*	5%	0.1%
	CDS	No data*	3%	1%
ELV	CS	22%	5%	2%
	CDS	27%	4%	2%

* Diminutives in Monika's corpus were only partially encoded, and thus they were not included in the analysis.

The results show that the frequency of tokens and lemmas derived from nouns and adjectives is low; however, verb derivatives are used very often and display a great variety. It should be noted that the frequency of derivatives of these parts of speech in CS and CDS could be related to the nature and type of word formation. Almost

all noun and adjective derivatives are formed by means of suffixes (rarely by means of prefixes), e.g. *burt-as* ‘a magic’ → *burt-inink-as* ‘a magician’, *gelež-is* ‘an iron’ → *gelež-in-is* ‘made from iron’. Even though semantically they are related to the base word, such derivatives are acquired as entirely new lexical items. Most verb derivatives are formed by means of prefixation, and a derived verb acquires new nuances of a more specific meaning, but does not change the main meaning of the base word (e.g. *neš-ti* ‘to carry’ → *at-neš-ti* ‘to bring’, *nu-neš-ti* ‘to take away’). Therefore, in the case of verb formation we see that the semantic category is preserved, whereas in the noun formation process the main tendency is change of category.

4.2 Productive usage

As mentioned in the previous section, at the beginning of the period the children used prefixed verbs abundantly, while derivatives of other parts of speech emerged later. The analysis of potential productivity revealed that verb derivatives stand out by their productive usage even during the first months of recordings. In accordance with the mini-paradigm criterion successfully applied to the acquisition of morphology in our group’s previous volumes (e.g. Dressler, Ketrez & Kilani-Schoch 2017; for more on the term mini-paradigm, see Introduction of this volume), a productive usage is the use of one affix with at least three different stems. When investigating verb derivatives, such cases were observed especially early in Elvijus’s speech. During the first month of research, from the premorphological to the protomorphological stage (1;7), two different prefixes were observed to occur in three or more different stems:

- (1) *pa-valg-y-ti* ‘to eat’, *pa-im-ti* ‘to take’, *pa-guld-y-ti* ‘to lay down’, *pa-žiūr-ė-ti* ‘to take a look’, *pa-kel-ti* ‘to pick up’;
- (2) *su-sto-ti* ‘to stop’, *su-duž-ti* ‘to break’, *su-galv-o-ti* ‘to create’.

In Monika’s speech, such cases were observed later, i.e. at the beginning of the stage of protomorphology (1;10):

- (3) *ati-dar-y-ti* [= *a-daj-y-ti*] ‘to open’, *ati-duo-ti* ‘to give’, *ati-trauk-ti* ‘to push’;
- (4) *pa-dar-y-ti* ‘to make’, *pa-žais-ti* ‘to play’, *pa-rod-y-ti* ‘to show’, *pa-dė-ti* ‘to put’, *pa-tep-ti* ‘to spread’, *pa-leis-ti* ‘to let’, *pa-duo-ti* ‘to give’, *pa-žiūr-ė-ti* ‘to take a look’, *pa-bėg-ti* ‘to run away’, *pa-im-ti* ‘to take’.

As seen in Example (4), the prefix *pa-* occurs even in ten different stems.

At the beginning of protomorphology, Elvijus (1;9) already used 5 prefixes (*pa-*, *su-*, *at-*, *į-*, *iš-*, *nu-*) in three or more different stems, and Monika (2;0) used 6 prefixes

(*pa-*, *su-*, *mat-*, *į-*, *iš-*, *nu-*). It is also worth mentioning that at the end of the protomorphology phase (Elvijus 2;6, Monika 2;7) the most productive prefixes in both children's speech are the same, which were used with at least three different stems.

Noun and adjective derivatives do not obtain productive usage in CS. In Monika's speech, less than three different stems with the same prefix or suffix were detected during any of the analysed months. On the other hand, in Elvijus's speech only one case of different adjective stem with the same suffix (*žaisl-in-is* 'toy', *mėt-in-is* 'mint', *vais-in-is* 'fruity' 2;1) and one case of noun derivatives were observed (*pa-kab-a* 'a hanger', *pa-galv-ė* 'a pillow', *pa-dang-a* 'a tire'). This tendency can be explained by a rare use of adjective and noun derivatives in CDS.

The discussion of productivity should also include the first word families in children's speech (for more on the term 'word family', see Introduction of this book). When analysing mini-paradigms of the verb in Elvijus's speech, the first families were already noticed in the second month (1;8):

- (5) *dė-ti* 'to put' → *pa-dė-ti* 'to put on', *su-dė-ti* 'to put together', *į-dė-ti* 'to put inside';
- (6) *mes-ti* 'to throw' → *iš-mes-ti* 'to throw out', *nu-mes-ti* 'to throw down'.

And one month later (1;9) even a 5-member word family was observed:

- (7) *pa-važ-iuo-ti* 'to drive a short distance', *at-važ-iuo-ti* 'to drive back, to come', *nu-važ-iuo-ti* 'to drive to', *į-važ-iuo-ti* 'to drive into', *iš-važ-iuo-ti* 'to drive out of'.

In Monika's speech, the first 3-member verb family was observed one month later than in Elvijus's case (1;9):

- (8) *ei-ti* 'to go', *iš-ei-ti* 'to go out', *at-ei-ti* 'to come'.

However, at the age of 1;10, even 3-member and 4-member word families were already noticeable, e.g.:

- (9) *im-ti* 'to take', *pa-im-ti* 'to take up', *iš-im-ti* 'to take out of', *at-im-ti*, 'to take away';
- (10) *ei-ti* 'to go', *at-ei-ti* 'to come', *iš-ei-ti* 'to leave';
- (11) *dė-ti* 'to put', *į-dė-ti* 'to put in', *pa-dė-ti* 'to put down';
- (12) *dar-y-ti* 'to do', *ati-dar-y-ti* 'to open', *pa-dar-y-ti* 'to finish doing'.

When analysing noun and adjective derivatives, it was observed that the opposition between simplex (base form) and its derivative form emerged early in the speech of both children. In Elvijus's speech, the opposition of *sup-ti-s* 'to swing' – *sūp-yn-ės* 'a swing' was noticed already at the age of 1;7, and at the age of 1;8 two more

oppositions emerged: *arbat-a* ‘a tea’ – *arbat-in-uk-as* ‘a teapot’, *sak-ė* (tell-PST.3) ‘told’ – *pa-sak-a* ‘fairy-tale’. In Monika’s data *tep-ti* ‘to spread’ – *tep-al-as* ‘a cream’ occurred at the age of 1;9 and at 1;10 – *seil-ės* ‘a saliva’ – *seil-in-uk-as* ‘a bib’; however, in general, 3-member families were still very few. The examples from Elvijus’s speech are:

(13) *lenkt-yn-in-is* ‘racing’, *lenkt-yn-iau-ti* ‘to race’, *lenkt-yn-ės* ‘a race’; (2;3)

(14) *kel-ias* ‘a way’, *kel-iau-ti* ‘to travel’, *kel-iau-t-oj-as* ‘a traveller’. (2;4)

Monika’s examples include:

(15) *mok-ė-ti* ‘to know’, *mok-ykl-a* ‘a school’, *mok-γ-ti-s* ‘to learn’; (2;2)

(16) *pri-klij-uo-ti* ‘to stick’, *klij-av-im-as* ‘a glue’, *nu-klij-uo-ti* ‘to stick off’. (2;8)

To sum up, productive usage of derivation emerges first with verbs (in transition from premorphology to protomorphology or at the beginning of protomorphology) and is later followed by a productive use of adjectives and then of nouns. A fast development of verb mini-paradigms, including up to five members at the very beginning of the research period was also observed. Noun and adjective word families are less numerous, although the first oppositions between simplex and derivatives were noticed at the very beginning of the research period.

4.3 Error analysis

It is interesting to observe that even though noun and adjective affixes are low in frequency and productivity in child speech; derivational errors are very infrequent, even at early stages of the acquisition of derivation. Only a few cases of an improper use of noun derivatives were observed in the whole analysed corpus. At the age of (1;8), Elvijus used the erroneous prefix *pa-* instead of *prie-*: *pa-kab-a* [= *prie-kab-a*] ‘a trailer’ (← *pa-kab-ė-ti* ← *kab-ė-ti* ‘to hang’), and at (2;4) the suffixes *-int-is*, *-ink-as* instead of *-inink-as*: *futbol-int-is* [= *futbol-inink-as*] ‘a football player’ (← *futbol-as* ‘a football’), *vart-ink-as* [= *vart-inink-as*] ‘a gate-keeper’ (← *vart-ai* ‘a gate’). It is important to mention that these errors can be phonological in their nature. In Monika’s speech, the erroneous reductive dissimilatory diminutivisation of the derivative was observed at the age of (2;8): *pa-tiek-ėl-į* [= *pa-tiek-al-ėl-į*] (dish-DIM-ACC.SG) ‘a dish’ (from *pa-tiek-al-as* ‘a dish’ ← *pa-tiek-ti* ← *tiek-ti* ‘to serve’). As can be seen, the girl omitted the derivational suffix *-al-* when adding a diminutive suffix.

Some errors might be interpreted as the so-called derivational creativity. They are mainly caused by logical reasoning:

- (17) the derivative *ant-ak-in-iai* [= *ak-in-iai* 'glasses'] was derived by Monika (2;2) from the noun *ak-is* 'an eye' and the prefix *ant-* ('on'); literally, it would mean 'an item which is placed on the eyes'.
- (18) *ėr-iuk-in-as* [= *ėr-iuk-as* 'lamb'] was presumably derived according to an analogical suffixation rule. In Lithuanian, many nouns which mean male animals are derived by adding the suffix *-in-* to nouns indicating male animals (e.g. *kat-ė* (cat-F.NOM.SG) 'a female cat', *lap-ė* (fox-F.NOM.SG) 'a female fox' and *kat-in-as* (cat-M-NOM.SG) 'a male cat', *lap-in-as* (fox-M-NOM.SG) 'a male fox'). In this case, when the mother asked the child: 'Who is the father of *ėr-iuk-as* 'a lamb', Elvijas replied: *ėr-iuk-in-as* (which literally would mean 'a male lamb').

In the children's speech, almost no cases of misuse of adjective derivatives were observed. There were none in Monika's speech (but she produced only 25 tokens/10 types of adjectives during the analysed period), whereas in Elvijas's speech (67 tokens/19 types), two cases of an erroneous suffix (19) and one omitted suffix (20) occurred:

(19) *lentk-yn-el-ė* and *lenkt-yl-in-ė* [= *lenkt-yn-in-ė*] 'racing'; (1;9)

(20) *kar-išk-a* [= *kar-ėiv-išk-a*] 'military'. (1;11)

In contrast to the case of noun and adjective derivatives, an abundant use of erroneous verb prefixes was observed in the data of both children. This can be related to the variety and early high frequency of prefixed verbs. As illustrated in Figure 1, at the beginning of the acquisition of prefixed verbs (in the stage of premorphology), the erroneous use of prefixes accounted for up to 35% of all prefixed verbs of Elvijas (1;7) and for 22% in Monika's (1;9) speech. Later, the number of errors consistently decreased, and at the end of the analysed period they amounted to less than 3%.

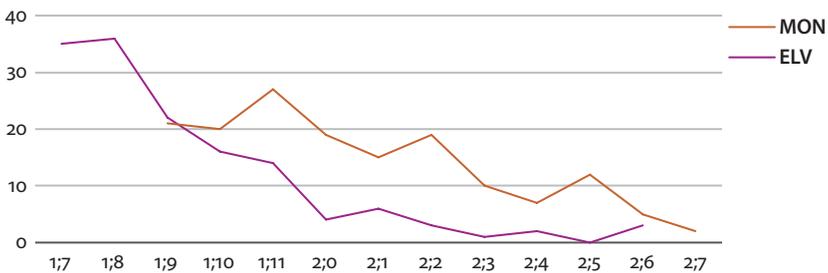


Figure 1. Errors in prefixed verbs

The errors of prefixed verbs include the use of the premorpheme *a* instead of a prefix (49% in Monika's, 50% in Elvijas's speech), prefix omissions (19% in Monika's, 8% in Elvijas's speech), prefix exchanges (7% in Monika's, 2% in Elvijas's speech),

prefix additions (0.5% in Monika's, 1% in Elvijus's speech), and phonotactic errors (23.5% in Monika's, 39% in Elvijus's speech).

Figure 2 shows that most often the children used the premorpheme *a* instead of the prefix; this was the key feature at the beginning of the analysed period in both CS, e.g.:

(21) *a-dė-ti* [= *į-dė-ti*] 'to put inside' (← *dė-ti* 'to put'). (1;10)

(22) *a-kel-ti* [= *už-kel-ti*] 'to put on' (← *kel-ti* 'to put'). (1;11)

In Elvijus's speech, such errors consistently increased from 9 to 40 at the age of 1;7–1;9, whereas later, at the age of 1;10–2;1, they decreased from 15 to 5. In Monika's speech, such errors increased from 11 to 39 at the age of 1;9–2;1, and later, at the age of 2;2–2;5, they decreased from 8 to 5. In Elvijus's speech, the premorpheme *a* of prefixed verbs disappeared at the age of 2;2, whereas in Monika's speech they disappeared only at the age of 2;6.

A significant number of prefix omissions was observed in the data of both children, e.g.:

(23) *Nor-iu dirž-q seg-ti* [= *už-seg-ti*].
 want-PRS.1SG belt-M.ACC.SG stick-INF
 'I want to stick the belt.' (2;4)

The use of an erroneous prefix was also common, e.g.:

(24) *Pa-trau-k* [= *pri-trau-k*] *man, kad būt-ų man.*
 PREF-pull-IMP.2SG I.DAT, that be-COND.3 I.DAT
 'Pull it to me, that I could have it.' (2;6)

Several cases of prefix addition were noticed, but they were few, e.g.:

(25) MOT
Band-y-k.
 try-IMP.2SG
 'try'
 CHI
Pa-band-au [= *band-au*].
 PREF-try-PRS.1SG
 'I am trying.' (2;6)

Phonotactic errors (usually consonant cluster simplifications) also occurred in the children's speech, e.g.:

(26) *į-gėr-iau* [= *iš-gėr-iau*].
 PREF-drink-PST.1SG
 'I drank.' (2;1)

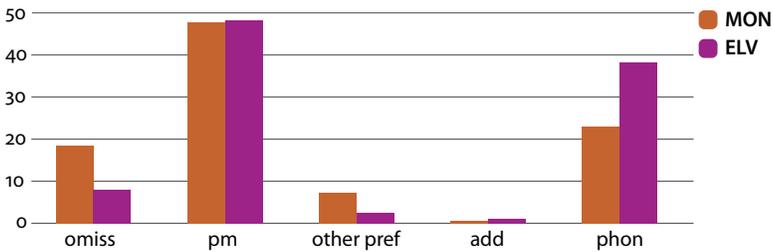


Figure 2. The distribution of different errors in verb prefixation (% among all errors)²

The errors observed in Elvijus's data reflect the general course of morphological development: we observe a greater number of omissions and the use of the pre-morpheme *a* instead of a prefix in premorphology (1;7, 1;8) (see Figure 4). These tendencies are characteristic of premorphology and were already pointed out by Savickienė (2003, 2007) in her analysis of the acquisition of noun morphology and Wojcik (2000) with respect to acquisition of verbs. In transition from pre- to protomorphology the omissions are very few, but the cases of the pre-morpheme *a* are growing as well as instances of correct usage of prefixes (1;9). Along with the acquisition of verb prefixation, the number of pre-morphemes decreased rapidly, and the phonological/phonotactic errors became the most frequent ones. Prefix addition is not characteristic of Elvijus's data as only few instances were registered. At the end of the analysed period (2;6), only very few errors were observed and all of them were phonotactic ones, usually consonant cluster simplifications, e.g. *pi-ėj-o* instead of *pri-ėj-o* (PREF-go-PST.3) 'came up' (Figure 3). At 1;11 we observe a few errors of prefix exchange (see Figure 3) together with a notable decrease of other prefixed errors. Moreover, the number of mini-paradigms increases and a productive use of prefixes is noticed. Therefore, we assume that this period, i.e. the beginning of protomorphology, marks the acquisition of the derivation rule by Elvijus.

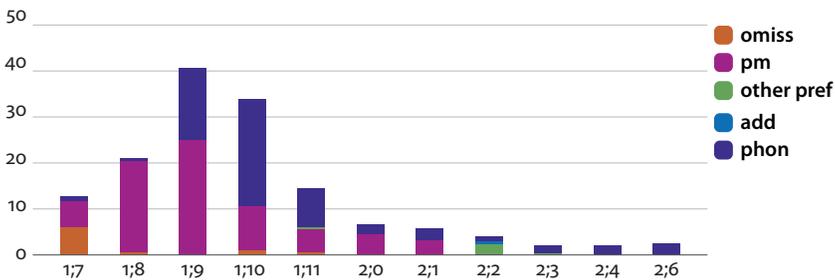


Figure 3. The distribution of errors in prefixed verbs in Elvijus's speech

2. *Omiss* – prefix omissions; *pm* – the use of the pre-morpheme *a* instead of the prefix; *other pref* – prefix exchanges; *add* – prefix additions; *phon* – phonotactic errors.

The acquisition of prefixes in Monika's speech takes more time as we see the dominating use of the premorpheme *a* until 2;1 (Figure 4). Starting from 2;2 on, the use of the premorpheme *a* decreases, but still cases of omissions and exchanges (up to 2;6) are registered (Figure 4). Even though few erroneous productions were noticed up to 2;7, we can conclude that Monika has acquired the derivational rule of prefixed verbs by the end of protomorphology.

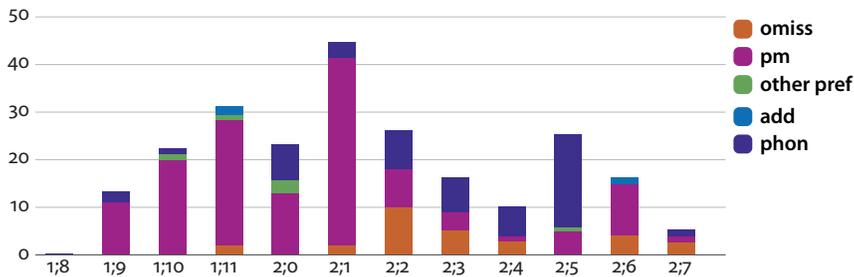


Figure 4. The distribution of errors in prefixed verbs in Monika's speech

A large number of word-formation errors at the beginning of the acquisition of derivative prefixed verbs suggest that the acquisition of the Lithuanian derivational verb system is a complex process. Developmentally, we could establish a path which first starts with the usage of the premorpheme *a*, at the same time (or later) followed by prefix omissions and erroneous usage of prefixes alongside the correct ones. The number of errors tended to decrease throughout the analysed period and were rare or not observed at all at the end of the analysed period.

The study reveals that a characteristic feature of the early period of both children's speech was the use of the premorpheme *a*. Our data demonstrate that prefixes are acquired at a slower speed and later than the case endings. The phenomenon we have discussed above (that is, the omission of a prefix where its presence is obligatory) has been noted in a number of other languages as well.

The dominant prefixed verbs (although mostly produced erroneously in CS) mark an action, process or state that is finished (e.g. *dė-ti* 'to put' → *pa-dė-ti* 'to put on'). The prefixed verbs that denote the direction (e.g. *ei-ti* 'to go' → *at-ei-ti* 'to come', *lip-ti* 'to climb' → *į-lip-ti* 'to climb/to come into') are produced with fewer errors than the others. This could be due to a relatively smaller variety and a more frequent usage of the same lemmas.

To summarize this section, it can be stated that two groups of derivatives are clearly distinguished. The first group is noun and adjective derivatives. Although they are rarely used and their affixes are not productive, only a few errors occurred during the whole analysed period. Presumably, noun and adjective derivatives are acquired during the investigated period as lexicalised items. The second group is

verb derivatives. Although during the first months prefixation errors were quite frequent, at the stage of protomorphology the children were able to perceive the main function of verb prefixes, that is, of modifying the meaning of a given verb. Such frequent use of prefixed verbs appear due to several reasons: (a) prefixed verbs are also frequent in CDS; (b) some forms of prefixed verbs are more often used in CDS with prefixes than without them; (c) in many cases, in their form verb prefixes correspond to prepositions which are homophonous with the prefixes (e.g. *į-ei-ti į namą* ‘to enter the house’, *su-si-tik-ti su draugais* ‘to meet the friends’), and our data evidenced that children produced these prepositions at the same time as the prefixed verbs.

The results of this study suggest that derivatives of nominal (noun and adjective) and verbal parts of speech are acquired not only at different ages but also in different ways. To conclude, it can be stated that the transition from an individual lexical acquisition of deverbal verb formation to the acquisition of morphological patterns occurs early in the investigated period; as concerns noun and adjective derivation, this happens, presumably, only after the investigated period, as we did not observe a clear application of derivational rules.

5. Conclusions

The analysis presented here focused on the children’s onset of morphological development, which is characterised in both children by a steep rise of lexical diversity in nouns, verbs and other words and by the emergence of mini-paradigms in derivation as well as form oppositions in inflection. These results further corroborate the hypothesis that there exists a close relationship between lexical and morphological development (“critical mass hypothesis”, see Marchman & Bates 1994). The observed period provides strong evidence for the relevance of the concept of protomorphology, first, as the stage where children detect morphology and, second, the influence of morphological richness in the input for morphology detection. It is at this period that the system of morphological grammar starts to develop. We have in mind such phenomena as the application of patterns of rules and the word spurt. Language use at this period is marked by creativity, and it does not come as a surprise that some deviations from adult speech norms are noticed.

Moreover, already at the premorphological stage both children use a great variety of verbal prefixes and the first oppositions between base nouns and noun diminutives together with the first manifestations of the inflectional system occur. At this early phase, relevant prefixes or suffixes were substituted with the premorpheme *a*, followed by the emergence of the grammatically correct markers. All the forms, however, have been used interchangeably for some time. Likewise, the

phoneme *a* frequently occurs in many morphemes (including suffixes, inflections and prepositions, see Savickienė 2003), especially at the early phases due to the phonetic simplicity of the sound.

The findings of this study allow us to make claims regarding some phenomena already identified in previous studies on the acquisition of Lithuanian morphology. Thus, we posit that the category-preserving character of deverbal verb formation is one reason for its early acquisition, as is the case with category-preserving diminutive formation, which contrasts with other types of denominal noun formation that do not preserve categories but just the word class.

One more note is in order before we close. This study revealed interesting tendencies in the domain of verbal prefixation; that is, in the diversity of prefixes, their order of acquisition, interrelationship with prepositions and so on. Moreover, we presume that inflectional suffixes cognitively are less precise than positionally salient. However, no far-reaching generalisations with respect to the acquisition of Lithuanian derivation and morphology in general can be put forward at this stage. This is due to the fact that our study has analysed the data of only two children. More data and more research are needed to offer more general conclusions on the issue.

References

- Ambrasas, V. (ed.) 1996. *Dabartinės lietuvių kalbos gramatika* (A grammar of Modern Lithuanian). Vilnius: Mokslo ir enciklopedijų leidykla.
- Ambrasas, V. (ed.) 1997. *Lithuanian Grammar*. Vilnius: Baltos lankos.
- Dabašinskienė, I. & Kamandulytė, L. 2009. Corpora of spoken Lithuanian. *Estonian Papers in Applied Linguistics* 5: 67–77. <https://doi.org/10.5128/ERYa5.05>
- Dabašinskienė, I. & Kamandulytė-Merfeldienė, L. 2017. The early production of compounds in Lithuanian. In Dressler, Ketrez & Kilani-Schoch (eds), 145–163.
- Dressler, W. U., Ketrez, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61>
- Jakaitienė, E. 1973. *Veiksmažodžių daryba* (priesagų vediniai) (Verb formation (suffixation)). Vilnius: Vilniaus V. Kapsuko universiteto leidybinis skyrius.
- Kamandulytė, L. 2009. Lietuvių kalbos būdvardžio įsisavinimas: Leksinės ir morfosintaksinės ypatybės (Acquisition of Lithuanian adjective: Lexical and Morphosyntactic Features). PhD dissertation, Vytautas Magnus University.
- Kamandulytė-Merfeldienė, L. 2015. The acquisition of Lithuanian adjectives: Lexical and morphological features. In *Semantics and Morphology of Early Adjectives in First Language Acquisition*, E. Tribushinina, M. D. Voeikova & S. Noccetti (eds), 313–346. Newcastle upon Tyne: Cambridge Scholars.
- Keinys, S. 1999. *Bendrinės lietuvių kalbos žodžių daryba* (Word formation in Modern Lithuanian). Šiauliai: Šiaulių universiteto leidykla.

- Klimas, A. 1975. Word-formation in Lithuanian. In *Flexion und Wortbildung* (Inflection and Word Formation), H. Rix (ed.), 130–147. Wiesbaden: Reichert.
- Klimas, A. 1991. Prefixes and suffixes in the Lithuanian verbal system. *Lituanus. Lithuanian Quarterly Journal of Arts and Sciences* 37(4). Retrieved from <http://www.lituanus.org/1991_4/91_4_06.htm> (2 May 2021).
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Marchman, V. A. & Bates, E. 1994. Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language* 21: 339–366. <https://doi.org/10.1017/S0305000900009302>
- Rimkutė, E., Kazlauskienė, A. & Utkā, A. 2016. Morphemic structure of Lithuanian words. *Open Linguistics* 2: 160–179. <https://doi.org/10.1515/opli-2016-0008>
- Savickienė [= Dabašinskienė], I. 2003. *The Acquisition of Lithuanian Noun Morphology*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften. <https://doi.org/10.1553/oxo0027c5d>
- Savickienė [= Dabašinskienė], I. 2007. Form and meaning of diminutives in Lithuanian child language. In *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43], I. Savickienė & W. U. Dressler (eds), 13–41. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.02sav>
- Stundžia, B. 2016. Word-formation in the individual European languages. Indo-European. Lithuanian. In *Word-formation: An International Handbook of the Languages of Europe*, Vol. 5, P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (eds), 3089–3106. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110424942-001>
- Urbutis, V. 2008a. Būdvardžių daryba (Adjective formation). In *Lietuvių kalbos enciklopedija* (Encyclopedia of Lithuanian language), V. Ambrazas & K. Morkūnas (eds), 97–98. Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- Urbutis, V. 2008b. Veiksmažodžių daryba (Verb formation). In *Lietuvių kalbos enciklopedija* [Encyclopedia of Lithuanian language], V. Ambrazas & K. Morkūnas (eds), 593–594. Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- Urbutis, V. 2009. *Žodžių darybos teorija* (Theoretical background of word-formation). Vilnius: Mokslo ir enciklopedijų leidybos institutas.
- Wójcik, P. 2000. *The Acquisition of Lithuanian Verb Morphology: A Case Study*. Krakow: Universitas.

Acquisition of noun and verb derivation in Estonian

Reili Argus

Tallinn University

The aim of the study is to describe the acquisition of derivation in Estonian. The analysis is based on recorded spontaneous speech of three Estonian children between the ages of 1;3 and 3;1. The role of productivity, transparency, semantics and frequency of suffixes and patterns, as well as the learning strategy (from simple to complex), have been analysed. The Estonian children started to use the first derivatives at the same time as the first compounds emerged and noun inflection started to develop. The number of different suffixes used by the children was quite large, but not all suffixes were used productively. The role of productivity and frequency can be observed in the order of the acquisition of the suffixes.

Keywords: first language, acquisition of derivation, noun derivation, verb derivation, Estonian

1. Introduction

Derivation and composition are important in the Estonian language as they are actively used in spoken and written language. Although compounding is still probably the most productive method for forming words (Kerge 2016), derivation is also a frequent word-formation strategy. The system of Estonian derivation was studied in depth in a comprehensive approach of word formation (Kasik 2015). There is also a dictionary of word families (Vare 2012) for Estonian and there have been other studies of word formation (e.g. the overview of Estonian word formation by Kerge 2016). Still, the acquisition of derivation is not a well-studied area in first-language research based on Estonian data. However, some works on the acquisition of word formation do exist: these works are mostly devoted to causatives (Argus 2012) and word formation by children with SLI (Padrik 2010). The only work on the acquisition of different derivational suffixes in Estonian among typically developing

children is a comparative approach to the acquisition of derivation in Estonian and Russian (Argus & Kazakovskaya 2018).

Typologically, Estonian is an agglutinating Finnic language, but is more fusional and analytic than other languages in the northern branch of the group (Erelt 2003: 7). According to Kasik (2015) and Vare (2012), there are over 60 derivational patterns and about 100 subpatterns (suffixes). Considering both form and meaning, only a few of the derivation patterns can be described as regular and productive. The target system that children start to acquire differs in the number of productive and transparent patterns in two subsystems: the derivation of verbs is more transparent than the system of noun derivation, which can be characterized as opaque, consisting of a large number of different suffixes with overlapping meanings and various degrees of productivity. It could be that children faced with a system where the derivation in different parts of speech have different degrees of productivity follow different paths in the acquisition of these subsystems.

Consequently, alongside the main aim of the study, which is the description of the acquisition of derivation in Estonian, the role of productivity of different subsystems and patterns and other factors, such as transparency and availability/frequency of different affixes in the input, will also be examined. Besides all of the above, the main research question is what is most important in the acquisition of derivation in Estonian: factors that have been emphasized in usage-based approaches (Dressler 2005; Dressler, Kilani-Schoch & Klampfer 2002), such as productivity, transparency and frequency, as well as semantics of different suffixes, or more general factors, such as learning strategies (from simplex to complex). The main theoretical claims when discussing the acquisition of derivatives have been that children acquire derivatives while mastering them via affixation (Clark 2014: 425–439), and the opposite view that derivatives may be acquired as single items (Gleason 1958: 176, and Nagy, Diakidoy & Anderson 1993: 45). The analysis of derivatives on the basis of their belonging to word families could clarify these claims. More specifically, the following research questions will be investigated: Does the occurrence of simplex stems precede the occurrence of derivatives, and does the child use derivational suffixes first with simple stems/roots and only afterwards with complex words (compounds and derived stems)? What role is played by different factors, such as productivity, transparency and frequency, in the acquisition of derivation? What are the main differences in the acquisition of nominal versus verbal derivation?

In this chapter, the development of Estonian derivation will be analysed on the basis of spontaneous speech data of three Estonian-speaking children from age 1;3 to 3;1. The chapter focuses on the emergence and development of verbal and nominal derivation in Estonian.

The analysis begins with a description of the emergence of the first derivatives, and then continues with an analysis of the development of derivational patterns and

their productivity, transparency and the semantic categories of derivatives (always in comparison with child-directed speech (CDS)). The comparison with CDS data is presented in every subtopic. The analysis of the emergence of simplex words vs. derivatives and the frequency of usage of suffixes with stems of different complexity, as well as derivatives belonging to different word families, will shed light on the general acquisition strategies followed by children.

2. The system of derivation in the Estonian language

The formal devices for derivation in the Estonian language are suffixes, zero-derivation, and some prefixoids which share some common features with prefixes (more abstract meaning than in the case of words, e.g. *eba* ‘non-’) and content words (they can attach to other suffixes and form new derivatives or occur independently, e.g. *mitte* ‘not’). Derivation includes the transformation of lexical units from (1) one part of speech to another (e.g. the nominalization of verbs, or the verbalization of nouns: *ela-ma* ‘to live’ → *ela-mine* ‘living’ or *kasu* ‘use/benefit’ → *kasu-ta-ma* ‘to use’), and (2) from one semantic subcategory of verbs or nouns to another (e.g. *hüppa-ma* ‘to jump’ → *hüpp-le-ma* ‘to hop’). Often the boundary between the base and the suffix is vague, e.g. the verb *rõõmusta-ma* may be based on *rõõm*: *rõõmu* ‘joy’ or on *rõõmus* ‘joyful, merry’, thus formed either with the suffix *-sta* or *-ta* (Kerge 2016).

The systems of verb and noun derivation differ in terms of semantic transparency and the number of suffixes. There are 9 semantic subcategories (e.g. frequentatives and causatives for verbs and action nouns for nouns) of derivational suffixes, each having 1–8 phonologically determined allomorphs of suffixes: around 30 suffixes for verbs and some 40 suffixes for nouns. Although there are fewer derivational suffixes for verbs than for nouns, most verbs are derivatives.

Verbal derivational suffixes can be divided into two main groups. The first group consists of valency-changing suffixes: (a) causatives, e.g. *kee-ma* → *kee-t-ma* ‘boil-CAUS-INF’; (b) reflexives, e.g. *peit-ma* ‘to hide something or somebody’ → *peit-u-ma* ‘to hide yourself’; (c) resultatives, e.g. *aur* ‘steam’ → *auru-sta-ma* ‘evaporate’; (d) translatives (verbs with the meaning of some resultative process) *pragu* ‘crack’ → *pragu-ne-ma* ‘to crack’; (e) instrumentatives, e.g. *õli* ‘oil’ → *õli-ta-ma* ‘to oil’; (f) imitatives, e.g. *niuts* ‘squeak’ → *niuts-u-ma* ‘to squeak’, *rag-ise-ma* ‘crackle’, usually formed from onomatopoeic stems; (g) state verbs, e.g. *rivaal* ‘rival’ → *rivaali-tse-ma* ‘to rival’; etc.

The second group are modifying suffixes; by adding a suffix, the change influences the nature of the event and does not affect the syntactic properties of a verb. Suffixes modifying the nature of the event are: (a) frequentatives, e.g. *karga-ma* ‘jump’ → *karga-le-ma* ‘hop, bounce’; and (b) momentatives (verbs for punctual

activities), e.g. *köhi-ma* ‘cough’ → *köha-ta-ma* ‘to cough once’ (Kasik 2015). Among verbal suffixes the formation of causatives, instrumentatives, and imitatives can be considered as productive and many novel derivatives can be found in different texts (see Kasik 2015: 108–135).

Similar to verbal suffixes, noun suffixes can be divided into suffixes that change the meaning of a noun and those that simply modify the meaning of a noun. In case of the modifying derivation the essential meaning of a word remains the same and the suffix just adds one additional feature to the meaning. Typical modifying suffixes are diminutives and suffixes denoting gender. The regular modification of nouns concerns mainly two types: female and diminutive suffixes: (1) *-nna* (less frequently also *-tar*) used for female gender in a number of nouns, e.g. *näitleja* ‘actor’ → *näitleja-nna*, *näitleja-tar* ‘actress’; (2) *-ke[ne]*: *-kese*, e.g. *lipp* ‘flag’ → *lipu-ke* ‘small flag’, and *-u*, which is widely used in child-directed speech, e.g. *nukk* → *nuk-u* ‘doll’, *lutt* → *lut-u* ‘pacifier’. Diminutives do not always have the morphosemantic meaning of smallness (see also Dressler & Merlini Barbaresi 1994); they are also used in pragmatically diverse contexts: For example, affection may be expressed in some contexts, e.g. (*onu* → *onu-ke* ‘uncle’), while irony is the most typical motive in the use and interpretation of such types, e.g. *proua* ‘madam/lady’, *proua-ke* ‘pretentious lady’ (see also Kerge 1991). As is typical for diminutives (see Savickienė & Dressler 2007), diminutives in *-u* are easy to inflect, as this word type lacks the stem (quantity) alternations otherwise typical of many frequent noun types.

Around 40 suffixes used for changing the meaning of a noun can be divided into some semantically clear categories: (1) denominal suffixes: (a) person suffixes, e.g. *-lane*, which refers to belonging to any group related to the base, e.g. *tallin-lane* ‘resident of Tallinn’, and *-line*, which indicates either an occasional relation to the base or to the temporary nature of implicit action (*töö-line* ‘worker’); (b) collective suffixes, e.g. *-kond* (*õpetaja* ‘teacher’ → *õpetajas-kond* ‘teaching staff’) and *-stik* (*sõna* ‘word’ → *sõna-stik* ‘dictionary’); (2) deverbal suffixes denote state, process or action (hardly distinguished from acts or results) and agent, actor, instrument, object or result (hardly distinguished from each other). The most regular and productive suffix for process nouns is *-mine* (e.g. *jooksma* ‘run’ → *jooks-mine* ‘running’) (see Kerge 2003, Sahkai 2011). The deverbal suffix *-ja* is used for agent and instrument nouns (e.g. *jooksma* ‘run’ → *jooks-ja* ‘runner’, *vahustama* ‘mix’ → *vahusta-ja* ‘mixer’) and the suffix *-e* for action and result nouns (*hüppama* ‘jump’ → *hüp-e* ‘leap/jump’). The formation of *e*-derivatives is regular but not productive, and many *e*-derivatives are lexicalized and used only as modifiers in compounds.

In addition to these relatively transparent categories, there are suffixes with very opaque and abstract meanings; some suffixes are polyfunctional, e.g. *-is* can form both (mostly pejorative) agent nouns and nouns that express the result or instrument of an action, e.g. *eputama* ‘put on airs’ → *eput-is* ‘poseur’ or in *riputama*

'sprinkle' → *riput-is* 'pendant (or scrap/crumb, usually for cake)'. Many of these suffixes have only a very general meaning of an entity or object; their meaning is nothing but the category of noun, e.g. *-ik* for *kuus-ik* 'spruce forest' and *kumm-ik* 'rubber boot', or *-kas* for *vilt* 'felt' → *vildi-kas* 'felt-tip pen' and *-nd* for *pakkima* 'wrap' → *pake-nd* 'wrapping' or *vahe* 'difference' → *vahe-nd* 'tool'. Most of these suffixes have only restricted productivity and many derivatives are lexicalized.

In addition to suffix derivation, there is also conversions in Estonian derivation,¹ e.g. zero-derivation of verbs from nouns (mostly instrumentatives), e.g. *liim* 'clue' → *liimi-ma* 'clue-INF=to clue', and nouns from verbs (mostly process verbs), e.g. *kõndi-ma* 'walk-INF=to walk' → *kõnd* 'walking'.

3. Data and method

The analysis in the present paper is based on the speech of three typically developing monolingual Estonian-speaking children between the ages of 1;3 and 3;1. There were 16.6 hours of recordings of spontaneous speech of one child, Martina, 15 hours of another child, Andreas, and 22 hours of Linda, in sum 53.6 hours of recordings made in everyday situations. Each recording session lasted about 60 minutes (except Martina's recordings, which lasted 90 minutes each) and there was usually one session per month. All of the data used were transcribed with the program Chat² and form a part of the Estonian database in CHILDES (subcorpora Kapanen, Vija and Zupping).³

Among the first derivatives, there were a lot of amalgams and omissions of syllables, e.g. *takis* [takist-us] 'impede-us=obstacle', *mistetas* [meister-da-s] 'master-da-3SG.PST=constructed', *puata* [puhas-ta-b] 'clean-CAUS-3SG.PRS=clean', *ollatäis* [üllat-us] 'surprise-us=surprise', etc. (MAR 1;5). Words which were used as derivatives by children but did not consist of phonologically clearly detectable suffixes were excluded from the analysis.

Derivatives were analysed according to their first occurrences, productive use (a suffix occurring with at least 3 different stems and simplex in the data), percentage of derivatives among nouns or verbs, number of different suffixes of nouns or verbs, and the semantic categories of the nouns and verbs.

1. The question of conversion and zero-derivation in Estonian has been addressed in Kasik (2012) and the Estonian noun-verb conversion has been described as zero-derivation.

2. The Chat conventions for transcription are part of the Childes Project (MacWhinney 2000) <<http://childes.psy.cmu.edu>>.

3. <<http://childes.psy.cmu.edu/browser/index.php?url=Other/Estonian/>> (20 May 2019).

The main issue was to identify what factors most frequently influenced the acquisition of derivation by Estonian-speaking children, and to form hypotheses concerning the paths through which the derivational system and subsystems of verb and noun derivation might develop.

4. Results

4.1 General overview of data

The absolute numbers of lemmas and tokens and the percentages of derivatives among nouns and verbs are presented in Table 1.

The number and the percentage of derivatives varied greatly in different recordings but did not exceed 14.5% of all verb tokens (Andreas' speech at age 2;1) and 19% of nouns (Linda's speech at age 2;5). The total number of derivative tokens started to increase in the speech of all children beginning at age 2;0; it increased moderately during the observation period, with no sudden spurts.

Generally, there were more derivative tokens in CDS than in child speech (CS) before age 2;0; in recordings made after 2;0, the percentage of derivatives was similar in the children's and their caregivers' speech and it was higher in children's speech in some recordings (see Table 1). The higher number of noun derivatives in Linda's recording made at age 2;5 may have resulted from the relatively small number of nouns (and speech in general) in that recording, and the frequency of some derivatives could have had an impact on the percentage of derivatives.

The percentages of noun and verb derivatives were very similar in the sense that the general number of derivatives did not exceed 19% of all tokens and CDS did not consist of many more derivatives than in the case of CS.

4.2 Emergence of the first suffixes

Estonian children start to use their first derivatives at the same time as the first compounds emerge and noun inflection starts to develop (Argus 2009, 2017; Argus & Kazakovskaya 2018). Still, there was only one derived noun lemma occurring in one inflectional form in the recording session, at the beginning of the observation period (Martina at age 1;3, Andreas at age 1;7, and Linda at age 1;5–1;7). The children began to use more than one form of the same noun only 6–8 months after they had used the first nouns.

The first derivatives to emerge in the children's speech were diminutives: Martina used the diminutive *nuk-u* 'dolly' (1;3); there was the diminutive *mähk-u* 'nappy' in Linda's speech (age 1;5), and *jänk-u* 'bunny' (1;7) in Andreas' speech. All

Table 1. Total number and percentage of noun and verb derivatives in the three corpora

Age	Martina				Andreas				Linda			
	Noun derivatives	% of all noun tokens	Verb derivatives	% of all verb tokens	Noun derivatives	% of all noun tokens	Verb derivatives	% of all verb tokens	Noun derivatives	% of all noun tokens	Verb derivatives	% of all verb tokens
1;3	1/1	0.3	2/2	2.1					0/0	0	0	0
1;4									0/0	0	0	0
1;5	4/30	6.3	2/7	3.4					1/4	2.7	0	0
1;6	2/5	1.2	2/2	1.6					1/2	1.2	0	0
1;7					1/1	0.4	0/0	0	1/1	0.4	1/10	8.4
1;8					2/4	0.6	0/0	0	2/3	3.3	¼	3.4
1;9	4/9	2.2	3/5	1.2	0/0	0	1/1	2.3	5/6	4.6	8/13	6.1
1;10	4/16	2.2	4/10	2.6	3/3	0.9	3/6	4.2	4/7	6.5	1/10	0.7
1;11	6/21	5.8	3/6	3.4	3/6	3.3	3/7	13.7	3/3	3.2	2/3	6.5
2;0					9/16	4.2	13/34	12	4/7	6.5	3/3	1.7
2;1	10/29	6.3	4/40	8	1/1	0.4	9/37	14.5	1/1	2.3	2/2	2
2;2									2/3	7.9	3/6	7.1
2;3	13/33	7.2	5/6	2.1	7/12	8.6	7/11	5.3	5/13	13	4/14	8.7
2;4	2/4	2.2	3/7	1.3	4/5	1.8	9/15	4.3	3/3	4.3	2/2	9
2;5					7/9	5.6	10/26	9.2	7/12	19.4	5/9	6.2
2;6					8/13	4.6	6/12	3.9	6/17	9.7	9/21	8.3
2;7	13/25	5.1	10/19	3.9	9/9	4.7	5/11	2.8	6/11	11.6	4/5	3.2
2;8					12/22	10	6/9	2.9	2/4	6.6	5/10	10
2;9									5/9	7.8	6/14	10.2
2;10									6/13	9.5	5/11	6.5
2;11									7/11	8.3	9/15	7.7
3;0	6/8	2	7/18	4.8	15/18	5.2	7/64	12.5	6/8	9.5	4/4	3.3
3;1					15/43	8.6	10/19	2.9				

the children's first diminutives were formed with the suffix *-u*, which only forms disyllabic words of an inflectional class which doesn't involve any stem alternations.

The next suffix the children started to use was also a diminutive: *-ke*: *linnu-ke* 'bird.GEN-DIM' (Martina 1;5), *kohu-ke* 'rise-DIM=sweet made from cottage cheese' (Andreas 1;10). The derivative *kohu-ke* can be considered to be a lexicalized word (*kohuke* is a brand name irregularly derived from *kohu+piim* 'curd, cottage cheese', which, in turn, is a compound based on the infrequent verb *kohu-ma* 'rise-INF'), but the child still used the suffix *-ke* productively, i.e. with several stems, and formed derivatives not present in the caregiver's speech (*saali-ke* 'hall-DIM', at age 2;3). All of the children also used the stylistically most marked diminutive suffix *-kene* (e.g. *linnu-kene* 'bird-DIM', Linda at age 1;9) with 3–6 different stems.

Other suffixes started to emerge shortly after the first diminutives and differed between the speech of all of the children. Martina used the suffix with a pejorative meaning *-is* (*mmeirmut-is* [hernehirmut-is] 'scarecrow') at age 1;5. At age 1;10, Andreas used a compound consisting of a derivative as a modifier, *puhk-e+päev* 'holiday' ← *puhka-ma* 'rest-INF',⁴ and a compound consisting of a derivative as a head, *saja+jalg-ne* 'centipede' ← *jalg* 'foot'. The third child, Linda, used a derivative with the suffix *-us* (e.g. *kingi* 'gift.GEN' → *kingi-t-us* 'present') at age 1;7.

Verb derivatives emerged approximately at the same age as noun derivatives did. The first verb derivative suffix in the speech of all three children was the causative suffix *-ta/-sta*, which is one of the most productive verb derivations in Estonian. Two children, Martina (age 1;3) and Andreas (age 1;9), used the suffix *-ta* with the same verb, *näi-ta-ma* 'show-CAUS-INF' ← *näge-ma* 'see-INF'. The third child, Linda, used the suffix *-sta*, which has a factitive meaning (*jooni-sta-d* 'line-FACT-2SG=draw', at age 1;7).

The emergence and frequency of the first noun and verb derivative suffixes in child speech correspond to input frequency to a great extent, e.g. the most frequent suffixes in the input to Martina, the noun suffix *-u* (40 tokens) and verb suffix *-ta/-t* (162 tokens), emerged first in the speech of the child. In the first recordings, the diminutive suffixes occurred most frequently both lemma and token-wise in the caregiver's speech. Diminutives even made up 2/3 of all noun derivative tokens in Linda's CDS at age 1;3. The suffix *-ta* was also the most frequent verb suffix in CDS for all children.

4. The suffix *-e* regularly forms compound modifiers of verbal meaning (see Kerge 2016: 3243).

4.3 Further development of derivation: The productive use of first suffixes and the emergence of different formation patterns and suffixes

The first diminutive suffixes did not become productive immediately upon emergence: Andreas first used the diminutive suffix *-u* at age 1;8 with two different roots: *jänk-u* ‘rabbit-DIM’, and *nuk-u* ‘doll-DIM’, and at age 2;1 with the third root *kät-u* ‘hand-DIM’. At age 1;9, he produced the simplex of one of his first diminutives (*jänes* ‘rabbit’). Linda’s first productive noun derivative suffix was the other diminutive suffix, *-ke*. First she used the simplex *Paul* (at age 1;9), then the suffix *-ke* emerged with other roots (*poti-ke* ‘pot-DIM’ at age 1;10, *kiisu-ke* ‘cat-DIM=kitty’ at age 2;3), and a derivative from a simplex used 3 months before (*Paul*) emerged at age 2;0, *Pauli-ke* ‘Paul-DIM’. The first productive suffix type in the speech of Martina was not the diminutive. She used the noun suffix *-us* three times with different roots (*ärat-us+kell* ‘awakening-*us*+clock=alarm clock’, *sala-dus* ‘secret’ in one recording made at age 2;1, *üllat-us* ‘surprise’ at age 2;3). At age 2;1, the simplex from *ärat-us* ‘awakening’, *ärata-ma* ‘wake-INF’ was also found. On the whole, the period between the emergence and the productive usage of a noun derivation suffix was 6 to 10 months.

The period between the first emergence and productive usage of verb suffixes also lasted up to 10 months in the case of Martina: she used the suffix *-ta* in her first recording at age 1;3 (*näi-ta-ma* ‘show-CAUS-INF’), and after that it appeared with different roots (e.g. *kee-t-ma* ‘boil-CAUS-INF=to make boil’ at age 1;5, *ehma-ta-ma* ‘startle-CAUS-INF’ at age 1;11), but both the derivative and the simplex from the same root emerged only at age 2;1 (*ärka-ma* ‘wake-INF’, *ära-ta-s* ‘wake-CAUS-PST’).

The two other children (Andreas, Linda) did not have such long periods between the first emergence and productive usages of verb suffixes. The causative suffix *-ta* emerged in Andreas’ speech at age 1;9 (*näi-ta* ‘show-CAUS’, *lõpe-ta* ‘finish-CAUS’) and 2;0 (*küipse-ta* ‘bake-CAUS’) and the simplex *lõppe-s* ‘finish-PST’ emerged after four months, at age 2;1. In Linda’s speech the period between first usage and productive usage was four months in length. The suffix *-ta* was registered in her recording first at age 1;8 (*mõista-ta-ma* ‘understand-CAUS-INF=guess’) and both the derivative and simplex from the same root appeared at age 1;9 (*uju-ta-vad* ‘swim-3PL’, *uju-ta-vad* ‘swim-CAUS-3PL=float’ at age 1;9); the third causative emerged in her speech at age 2;0 (*mahu-ta* ‘contain-CAUS’).

According to the productivity criteria among noun suffixes, only the diminutive can be considered to be productively used. Other suffixes were usually used with only one or two different lemmas or the corresponding simplex was not found in the data. Among verbal suffixes, in addition to causatives resultatives (e.g. *pildi-sta* ‘picture-SUFF=take a photo’) and imitatives (e.g. *ur-ise-ma* ‘growl-SUFF-INF=growl’) can be considered as being used productively by the children.

Child neologisms were not common in the recordings (only four). Linda used verb derivations, e.g. *liug* 'slip' → *liug-ata-ma* 'slide' at age 1;8, and *kleeps* 'sticker' → *kleeps-u-ma* 'to put on a sticker' at age 1;9. Andreas had two self-invented noun derivatives: the derivative *kulist-us* [=kurist-us] 'something bad' was probably formed from *kuri* 'bad, angry' by analogy of *pahand-us* 'mischief' or *karist-us* 'punishment', and *mustu-kas* 'black alien' (from *must* 'black' and *tulnu-kas* 'alien') at age 2;6.

The most frequent pattern of noun derivatives in the speech of two of the children, Martina and Andreas, was 'noun+suffix' (e.g. *foto-kas* 'photo-SUFF=camera'). The third child, Linda, formed more nouns according to the pattern verb+suffix, e.g. *ära-ta-ma* 'wake-CAUS-INF' → *ära-t-us* 'wake-CAUS-SUFF=awakening'. Although there were other patterns, such as verb, adjective, adverb and even a numerical stem or onomatopoeic word + derivational suffixes (see Table 2) in the speech of the children, noun+suffix and verb+suffix were the most frequent patterns. The high frequency of diminutives might have an impact on the frequency of the 'noun+suffix' pattern among noun derivatives.

Among verb derivatives, the pattern 'verb+suffix' was clearly the most frequent pattern and the denominal pattern 'noun+suffix' was used less. The pattern 'onomatopoeic word+suffix' was not one of the first verb derivation patterns used by the children and emerged only after age 2;0–2;3. The number of these derivatives increased with age in the speech of all the children. These imitative verbs were formed by adding different suffixes to onomatopoeic stems, usually imitating some kind of sound, e.g. *torr* → *tor-ise-ma* 'grumble', Linda, at age 2;3; *sups* → *sups-a-ma* 'smash' (Martina, 2;3); *köhh* → *köh-i-ma* 'cough' (Andreas, 2;5); *raks* → *ragi-sta-ma* 'grab' (Andreas, 2;1); *plaks* → *plaksu-ta-ma* 'clap' (Andreas, 2;3), and *nuuks* → *nuuks-u-ma* 'sob' (Linda, 2;9).

Table 2. Number of lemmas formed according to different derivational patterns

Patterns	Nouns			Verbs		
	Martina	Andreas	Linda	Martina	Andreas	Linda
N+suffix	22	29	21	6	16	11
V+suffix	17	24	29	18	17	22
ADJ+suffix	1	4	2	2	10	5
ADV+suffix	5	2	6	1	0	0
NUM+suffix	1	0	0	0	0	0
ONOM+suffix		1	1	2	12	7
N+0 (conversion)				4	5	1
ADJ+0 (conversion)				0	0	1
Total	46	60	59	33	60	47

While the pattern ‘adverb+suffix’ was more frequently used for noun derivatives, the pattern ‘adjective+suffix’ was more common in forming verbs: *kallis* ‘dear’ → *kallis-ta-ma* ‘dear-FACT-INF=hug’, *kuiv* ‘dry’ → *kuiva-ta-ma* ‘dry-FACT-INF=to make dry’, Andreas at age 2;6.

Conversion was used by children only in the derivation of verbs. The most frequently used converted verb was *sodi-ma* ‘scribble-INF’ (formed from the noun *sodi* ‘mess’), which appeared in Andreas’ speech 10 times in 4 recordings. The verb *värvi-ma* ‘paint-INF’ occurred in the recordings of two children, Martina and Andreas (6 times in total); other converted verbs were used usually only once and then only by a single child.

The number of different noun suffixes was very similar in the speech of the children. The adults had a little bit more variety in their use of suffixes, but, looking more closely at the frequency of suffixes in CDS (Table 3), we can see that although the order of frequency did not exactly match the order of the emergence of suffixes in the speech of the children, still the more frequent suffixes (e.g. *-u* (DIM), *-kene*, *-us*) tended to be among the first suffixes to emerge. The early emergence of diminutives has been reported in several languages, e.g. in Lithuanian, where the first morphemes the child in question used were diminutive suffixes (Savickienė 2007: 36), or in Russian, where diminutives were among the means to generalize the use of grammatical markers (Protassova & Voeikova 2007: 68).

Table 3. Noun suffixes in CS and CDS

Child	Suffixes, examples in order of emergence	No. of suffixes	Suffixes, examples in order of frequency	No. of suffixes
	CS		CDS	
Martina	<i>-u</i> , <i>-us</i> , <i>-is</i> , <i>-ke</i> , <i>-ik</i> , <i>-kene</i> , <i>-ja</i> , <i>-e</i> , <i>-kas</i> , <i>-i</i> , <i>-k</i> , <i>-mine</i> , <i>-ne</i> , <i>-nd</i> , <i>-lane</i>	15	<i>-u</i> (DIM), <i>-us</i> , <i>-mine</i> , <i>-e</i> , <i>-k</i> , <i>-ja</i> , <i>-ik</i> , <i>-u</i> , <i>-kene</i> , <i>-kas</i> , <i>-I</i> , <i>-lane</i> , <i>-m</i> , <i>-nd</i> , <i>-is</i> , <i>-ke</i> , <i>-line</i> , <i>-t</i> , <i>-er</i> , <i>-stik</i> , <i>-o</i> → <i>u</i> , <i>-ne</i> , <i>-t</i> (1)	22
Andreas	<i>-u</i> (DIM), <i>-ke</i> , <i>-e</i> , <i>-ne</i> , <i>-ja</i> , <i>-line</i> , <i>-us</i> , <i>-kas</i> , <i>-mine</i> , <i>-kene</i> , <i>-ur</i> , <i>-ik</i> , <i>-i</i> , <i>-is</i> , <i>-k</i> , <i>-nd</i> , <i>-ks</i> , <i>-stik</i>	18	<i>-us</i> , <i>-u</i> (DIM), <i>-kas</i> , <i>-ja</i> , <i>-a</i> , <i>-ik</i> , <i>-mine</i> , <i>-u</i> , <i>-kene</i> , <i>-k</i> , <i>-ke</i> , <i>-e</i> , <i>-us</i> , <i>-ne</i> , <i>-ur</i> , <i>-m</i> , <i>-ti</i> , <i>-is</i> , <i>-ps</i> , <i>-line</i> , <i>-lane</i>	21
Linda	<i>-us</i> , <i>-i</i> , <i>-u</i> (DIM), <i>-kene</i> , <i>-kas</i> , <i>-ik</i> , <i>-ke</i> , <i>-e</i> , <i>-is</i> , <i>-mine</i> , <i>-ja</i> , <i>-ts</i> , <i>-ks</i> , <i>-s</i> , <i>-tis</i> , <i>-a</i> , <i>-u</i>	17	<i>-u</i> (DIM), <i>-kene</i> , <i>-k</i> , <i>-ke</i> , <i>-us</i> , <i>-e</i> , <i>-kas</i> , <i>-ik</i> , <i>-ti</i> , <i>-ps</i> , <i>-mine</i> , <i>-is</i> , <i>-u</i> , <i>-o</i> → <i>u</i> , <i>-ja</i> , <i>-u</i> , <i>-line</i> , <i>-la</i> , <i>-a</i> , <i>-m</i> , <i>-in</i>	20

Table 4. Verb suffixes in CS and CDS

Child	Suffixes, examples in order of emergence	No. of suffixes	Suffixes, examples in order of frequency	No. of suffixes
	CS		CDS	
Martina	<i>-ta, -sta, -t, -du, -a, -nda, -le, -tse, -ne, -ise</i>	10	<i>-ta</i> (120), <i>-sta, -t, -le, -u, -nda, -i, -rda, -ata, -stu, -lsa, -tse, -lda, -a</i> (1)	13
Andreas	<i>-t, -rda, -ise, -i, -sta, -ta, -nda, -a, -eeri, -u</i>	10	<i>-ta</i> (99), <i>-sta, -i, -ise, -nda, -le, -t, -rda, -u, -a, -lda, -eeri</i>	12
Linda	<i>-sta, -ata, -ta, -u, -le, -t, -a, -ne, -i, -o</i>	10	<i>-ta</i> (146), <i>-sta, -le, -ise, -u, -a, -rda, -nda, -lda, -tse, -ata, -eeri, -t</i>	13

There were fewer verb suffixes than nominal ones in the speech of the children and in CDS and, as with the verbs, the number of suffixes was a little bit higher in CDS than in CS (see Table 4). Although the verb suffixes occurring in the children's speech differed to some extent, the number of suffixes was exactly the same in the speech of all of the children. The most frequent suffix in CDS was among the first suffixes in the speech of the children. The most frequent verb suffixes occurred approximately 100 times in the data of one child, while infrequent suffixes only occurred once or twice (see Table 4).

The order of the semantic categories of nouns presented by different suffixes was quite similar in the speech of all the children. In the cases of the pattern 'N+suffix', besides the three diminutive suffixes (*-u, -ke, -kene*, which were used with several stems), the children used several suffixes for deriving nouns with the very broad meaning of an object/entity or person: *-ik* (*homm-ik* 'tomorrow-ik=morning'), *-k* (*lennu-k* 'flight.GEN-k=plane'), *-nd* (*pake-nd* 'pack-nd=package'), *-kas* (*foto-kas* 'photo-kas=camera'), *-lane* (*mesi-lane* 'honey-lane=bee'), *-ne* (*saja+jalg-ne* 'hundred+leg-ne=centipede'), *-line* (*küla-line* 'village-line=visitor'), and *-ks* (*jän-ks* 'rabbit-ks'). There was only one suffix for collective nouns, *-stik* (*ilu+tule-stik* 'beauty+fire-stik=firework'), in the speech of Andreas.

The most frequent suffix for forming deverbal nouns indicated a process or a result (often not distinguishable), e.g. *-us* (*joonist-us* 'draw+us=drawing'), which was used with 18 different stems. The suffix *-ja* was used both for indicating an agent (*õpeta-ja* 'teach-ja=teacher') and an instrument (*kruvikeera-ja* 'screwdriver'). The suffix *-mine* (*rääki-mine* 'speak-mine=speaking') for indicating actions was not frequent in CS and emerged with only 5 different verbal bases. The only suffix with a pejorative meaning (*-is, herne+hirmu-t-is* 'pea+fear-CAUS-is=scarecrow') was used by Martina at age 1;5. Some patterns with a verbal base can be used with suffixes that also combine with noun stems: the suffix *-is* (see above mentioned example), as well as the suffixes *-k* (*söö-k* 'eat-k=food'), *-e* (*kast-e* 'dip/moisten-e=dressing'),

and *-kas* (*tulnu-kas* ‘come-*kas*=alien’). Two different deverbal suffixes, in addition to the previously mentioned suffix *-ja*, were used to indicate the meaning of an instrument: e.g. *-ur* (*ved-ur* ‘pull-*ur*=locomotive’) and *-i* (*arvut-i* ‘calculate-*i*=computer’).

The occurrence of suffixes with adverbial and adjectival stems was quite rare in CS: most suffixes were used only with one or two stems. Children used two suffixes with adjectival stems (*-us*, e.g. *ulak-us* ‘naughty-*us*=naughtiness’, *-kas*, e.g. *sini-kas* blue-*kas*=bruise’), and three suffixes with adverbial stems, e.g. *-kas* (*alu-kas* ‘down-*kas*=panties’), *-ik* (*ümbr-ik* ‘around-*ik*=envelope’) and *-is* (*täid-is* ‘full-*is*=filling’). The pattern with a numeral stem emerged only once in one noun derivative: *-ik* (*kaks-ik* ‘two-*ik*=twin’). Concerning the productive use of different noun suffixes, it can be concluded that 4 suffixes corresponded to the productivity criteria (diminutive suffixes *-u*, *-ke*, *-kene*; noun suffixes with the general meaning of process or result *-e*). The other 15 suffixes were not used productively by the children.

The order of emergence of different semantic categories of noun derivatives can be summarized as follows: the children started with diminutives. This result agrees with previous results from other languages, e.g. English (Clark 2014: 426). The next suffixes used can be grouped into a very general semantic category, that of objects (e.g. *-is*, *-k*, *-ik*, *-kas*, *-e*, *-u*), results (often indistinguishable from objects, e.g. suffix *-nd*), or people (e.g. *-ne*, *-lane*, *-line*). The category of an instrument is represented by 3 suffixes: *-ja*, *-i*, and *-ur*. The next category to emerge was agent, and the suffix used was the same as already used for instruments (*-ja*). This order differs somewhat from results found in English, where agentive and instrumental nouns emerged early (Clark 2014: 426). The difference can be explained by availability: there are many derived nouns for objects in everyday Estonian. The very general category of processes or results (often with overlapping meaning) was expressed using the suffix *-us* by two children. The suffix *-mine* was used by them to express the meaning of an action. The suffix *-stik* for collective nouns emerged late and was used only once (*ilutule-stik* ‘firework’) by one child, Andreas at age 3;0.

Of the 17 different verb suffixes, only three were used productively by the children: suffixes were those with a causative/factitive meaning (*-t*, *-sta*, *-ta*). Therefore, it can be stated that before age 3;0 the Estonian children productively used only valency changing verbal suffixes.

The order of the semantic categories of verbs presented by different suffixes was very similar in the speech of all the children. Deverbal verb derivatives emerged early: causatives at age 1;3–1;9 and factitives at age 1;5–1;10. Other semantic categories emerged later and at different ages and in different orders in the speech of the children.

The first momentatives, e.g. *liug-ata-ma* ‘slide’, emerged in the speech of Linda at age 1;8, and the first imitatives, e.g. *sups-a-ma* ‘quick jumping motion’, in the speech of Martina at age 2;3. Reflexives were represented by only three derivatives:

maand-u-ma ‘land’ (Martina, at age 1;9), *murd-u-ma* ‘break down’ (Linda, 2;8), and *tund-u-ma* ‘seem’ (Linda, 2;9). Verbs with continuative meaning emerged at different ages and were relatively infrequent, e.g. *til-ise-ma* ‘ring’ (Martina at age 3;0, but Andreas at age 1;10) and *ur-ise-ma* ‘growl’, Andreas at age 2;5). Verbs with frequentative meaning emerged later in the speech of all three children and were not frequent; all of these verbs were formed from onomatopoeic words, e.g. *toks* → *toks-i-ma* ‘knock’ (Andreas at age 2;0) and *köhh* → *köh-i-ma* ‘cough’ (at age 2;5). Even more infrequent (represented by only one derivative) were translatives, e.g. *kogu-ne-ma* ‘gather’ (Martina, at age 3;0).

The order of emergence of different semantic categories of verb derivatives can be summarized as follows: the children started with verbal suffixes changing the verb valency of the verb, i.e. with causatives and factitives. These suffixes were very frequent also in CDS. After that they started to use suffixes that modified the meaning of the verb, i.e. translatives, frequentatives, momentatives, and continuatives.

4.4 The general characteristics of the process of the acquisition of derivation: Rising complexity and word families

According to the idea of rising complexity during acquisition (see Zurek 1990; Dziubalska-Kończyk 2014, and for inflectional morphology Dressler 2011), with respect to the acquisition of derivatives, children should start to use derivatives only after they have already acquired corresponding simplex stems. To test the validity of this hypothesis, all lexemes emerging first as simplex, and only afterwards as derivatives in CS and vice versa were analysed (see Table 5 and Figure 1).

Table 5. Derivative lemmas in the speech of the children

	Simplex before derivative	Simplex at the same time as derivative	Derivative before simplex	Derivative without simplex	Total number of derivative lexemes
Martina					
Nouns	16	6	6	19	47
Verbs	10	2	3	13	28
Andreas					
Nouns	23	2	9	22	56
Verbs	11	14	3	31	59
Linda					
Nouns	9	6	7	27	49
Verbs	13	4	5	26	48

The total number of derivative lexemes was quite similar in the speech of all the children. Martina used fewer verb derivatives than the other children. In the speech of the other children, the number of verb and noun derivatives was almost the same (see Table 5).

One-third of all derivatives appeared for the first time as simplex and in some cases the derivative and simplex occurred simultaneously, within the same recording of CS. The percentage of these lexemes was approximately 10%. The other 10% of all derivatives emerged first as derivatives and only afterwards as simplex and, what is even more important, more than half of all derivatives did not have a simplex match in the data. It would be assumed that most of these words are lexicalized words and not productively formed during speech production (see Figure 1).

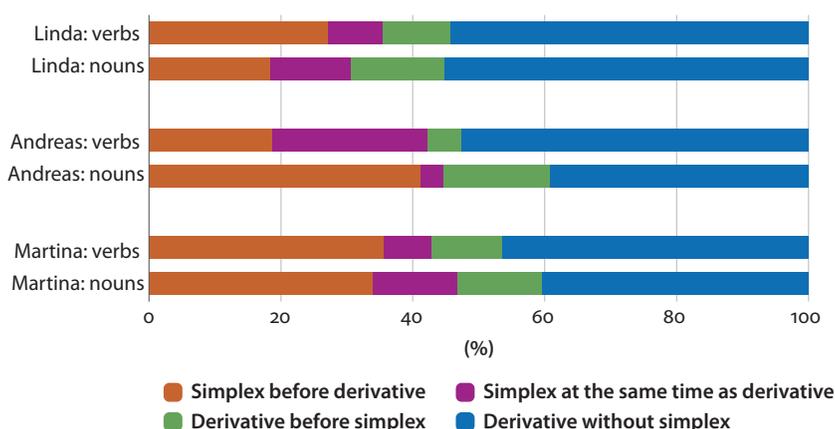


Figure 1. Order of emergence of derivatives according to the rising complexity model

The next question was whether the children would use derivational suffixes first with simple stems/roots and only afterwards with complex words (i.e. with compound or derived stems). The first suffixes (within the diminutive semantic category) were added to simplex stems. Shortly after that, the children started to use derivational suffixes inside compounds, i.e. the derivative could first have occurred as the modifier: *puhk-e+päev* ‘rest-SUFF+day=day of rest’ (Andreas 1;11), with the simplex stem with the same suffix only emerging after (*kast-e* ‘water-SUFF=dip/dressing’ 2;0), and then later being used with several stems. This also supports the view that a child will first use the suffix as a part of a lexicalized word, not productively. This result is similar to one reported by Clark (2014: 426) “...the first uses of affixes tend to be unanalysed uses, where the affix is simply part of the word”.

Some suffixes (especially deverbals) occurred first, not only with derived verb stems as modifiers of compounds: *ära-t-us+kell* ‘alarm clock’ ← *ära-ta-ma*

‘awaken-CAUS-INF’ (Martina 2;1), but also as heads of compounds: *piima+vahu-sta-ja* ‘milk+foam-CAUS-*ja*=milk foamer’ (Martina 1;9) (see also Dressler et al. 2019). Thus, there is still no clear evidence that in Estonian the same derivational suffix occurs first with simplex and only afterwards with complex stems.

A significant number of derivational formations in Estonian are not transparent (more nouns than verbs); therefore, it is important to study “word families”, i.e. classes of derivational formations, compounds and simple lexemes that share a word or stem. This makes it possible to study not only the relations among derivative categories, but also the development and organization of the lexicon. The number and size of word families is expected to be smaller in the speech of children than in the speech of adults.

Approximately 60% of all of the derivatives occurred in word families in the children’s speech data. There were slightly more nouns than verbs among the derivatives occurring in word families (see Table 6). All of the children started with 2-member word families, i.e. pairs, and the percentage of such pairs was quite high in the children’s speech. The most common noun pair consisted of two nouns (2/3 of all word pairs), mostly a diminutive and simplex, e.g. *lill* ‘flower’ – *lille-kene* ‘flower-DIM’ (Martina), or a noun derivative and simplex verb, e.g. *ist-e* ‘seat’ – *istu-ma* ‘sit-INF’ (Andreas). In the case of bigger word families, one or two members were usually compounds: *söö-ma* ‘eat’ – *söö-t-ma* ‘eat-CAUS-INF=to feed’ – *söö-k* ‘eat-*k*=food’ – *kommi+söö-dik* ‘candy.GEN+eat-*dik*=candy eater (sweet tooth)’ (Andreas). The children used only a few word families bigger than 3 members. In the case of verb derivatives, the 2-member word families mostly consisted of causative and non-causative pairs, e.g. *jää-ma* ‘stay-INF’ – *jä-t-ma* ‘stay-CAUS-INF=leave’ (Linda). Other combinations, such as a verb derivation and a noun (*pildi-sta-ma*

Table 6. Word families in the speech of the children (percentages)

	Derivatives not occurring in word families	Derivatives occurring in 2-member word families	Derivatives occurring in 3-member word families	Derivatives occurring in 4-member word families
Martina				
Nouns	44.7%	27.6%	21.2%	6.5%
Verbs	36.7%	46.7%	10%	6.6%
Andreas				
Nouns	38.4%	46.6%	13.3%	1.7%
Verbs	55.9%	32.2%	10.2%	1.7%
Linda				
Nouns	47.4%	52.6%	0%	0%
Verbs	58.3%	35.4%	6.3%	0%

‘picture-CAUS-INF=take a photo’ and *pilt* ‘picture/photo’) or a verb derivation and simplex adjective (*sooje-nda-ma* ‘warm-FACT-INF=warm up’ – *soe* ‘warm’, Andreas), were not common. The biggest word families had 4 members and consisted of one simplex noun, one simplex verb, one noun derivative and one noun compound, e.g. *tud-u-mine* ‘sleep-DIM-mine=sleeping’ – *tud-u-le* ‘sleep-DIM-ALL=(go) to sleep’ – *tud-u-b* ‘sleep-DIM-3SG’ – *tud-u-lapi* ‘sleep-DIM-napkin=binky ((child’s)security blanket)’ (Martina).

There was no significant difference between the number of nouns compared to the number of verbs occurring in word families – with two children the number of verbs occurring in word families was bigger, and with one child it was smaller.

Table 7 gives an overview of the percentage of derivatives in word families in the speech of the adults. An analysis of the target system revealed that approximately 63% of all derivatives occurred in word families in CDS, i.e. the difference between the children and their caregivers was unexpectedly small. The adults used more word families consisting of three and four members than the children did and there were even two 5-member word families in the CDS to Martina, *puhas* ‘clean’ – *puhas-ta-b* ‘clean-CAUS-3SG’ – *puhas-ta-mine* ‘clean-CAUS-NSUFF=cleaning’ – *põranda+puhas-ta-ja-d* ‘floor-clean-CAUS-NSUFF-PL=floor cleaners’ – *puhas-t-us+vahend* ‘clean-CAUS-NSUFF+tool=cleaning product’. The smaller word-families in CS are indicative of the developing system of derivation of children compared with adults, but at the same time, they could just be the result of smaller vocabulary of children in general.

Table 7. Word families in CDS (percentages)

	Derivatives not occurring in word families	Derivatives occurring in 2-member word families	Derivatives occurring in 3-member word families	Derivatives occurring in 4-member word families	Derivatives occurring in 5-member word families
Martina					
CDS					
Nouns	26.4%	48.8%	21.5%	2.5%	0.8%
Verbs	40.5%	41.9%	13.5%	2.7%	1.4%
Andreas					
CDS					
Nouns	24.4%	57.5%	13.8%	4.3%	0%
Verbs	42%	43.6%	10.1%	4.3%	0%
Linda CDS					
Nouns	46.8%	33.3%	17.8%	2.1%	0%
Verbs	45.4%	37.3%	14.3%	3%	0%

5. Conclusion

The acquisition of Estonian derivation constitutes a considerable task so it may take until the end of the third year of life for children to acquire the most productive parts of derivational patterns and suffixes. The children in the study started to use the first derivatives at an early age, i.e. at the same time as the first compounds emerged and noun inflection started to develop. Still, the first derivatives occurred in only one inflectional form and the first suffixes with only one lemma at the beginning of the observation period. A more transparent system of verb derivation did not have an impact on the first emergence of different derivatives, and verb derivatives emerged approximately at the same age as noun derivatives in the speech of all of the children.

The percentage of noun and verb derivatives among all nouns and verbs was very similar: the general number of derivatives did not exceed 15% of all tokens in the children's or the adults' speech.

There was a considerable number of different suffixes in the speech of the children (10 suffixes used for verb derivatives and 15–17 suffixes used for noun derivatives). The smaller number of suffixes in the case of verbs points to the more compact system of verb derivation compared to noun derivation. Although the number of different suffixes used by the children was large, not all suffixes were used productively by the children. By the end of the observation period, the children had productively used less than half of the noun suffixes (4 of 15) and only three out of 17 of the verb suffixes. Productive noun suffixes were diminutives and one suffix with the meaning of object. Causatives and resultatives were the productive verb suffixes, and it is clear that by the end of the observation period the children used only valency-changing verbal suffixes productively. The productivity and frequency of suffixes are related, since these suffixes are very frequent in CDS and colloquial Estonian in general.

Following the idea of increasing complexity during acquisition, the emergence of derivatives before simplex forms was analysed. The results revealed that one third of all derivatives appeared for the first time as simplex and in 10% of cases the derivative and simplex occurred simultaneously, within the same recording of CS. At the same time, 10% of all derivatives emerged first as derivatives and only afterwards as simplex, and more than half of all derivatives did not have a simplex match in the data at all. Therefore, it can be assumed that a large number of derivatives are lexicalized words and are not productively formed during speech production and can therefore be supposed to be acquired as lexical words.

The first suffixes emerged in simplex stems in the speech of the children but very shortly after that some suffixes were used inside compounds. Some suffixes also emerged first with derived stems as bases. Thus, there is still no clear evidence that in Estonian the same derivational suffix occurs first with simplex and only afterwards with complex stems. This tendency shows also that in Estonian at least some suffixes can be acquired as lexicalized units. Some of the “lonely” derivatives

illustrate the unproductive part of Estonian derivation. Approximately 60% of all derivatives in CS and 63% of all derivatives in CDS occurred in word families. The occurrence of larger word families in CDS is indicative not only of the children's developing system of derivation but also their generally smaller vocabulary compared to that of adults.

The initial hypothesis that children who are faced with a system where different parts have different degrees of productivity would follow different paths in the acquisition of these subsystems was not confirmed by the study. Children did not acquire verb derivation faster than noun derivation, the number of productive suffixes was not larger in the case of verbs and the verbs did not occur in larger word-families than nouns.

References

- Argus, R. 2009. The early development of case and number in Estonian. In *Development of Nominal Inflection in First Language Acquisition. A Cross-Linguistic Perspective*, U. Stephany & M. D. Voeikova (eds), 111–152. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110217117.111>
- Argus, R. 2012. Kausatiivisuse omandamisest eesti keeles (Acquisition of causatives in Estonian). *Eesti Rakenduslingvistika Ühingu Aastaraamat* (Yearbook of Estonian Applied Linguistics) 8: 5–20. <https://doi.org/10.5128/ERYa8.01>
- Argus, R. 2017. Acquisition of noun compounds in Estonian. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketrez & M. Kilani-Schoch (eds), 165–189. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61>
- Argus, R. & Kazakovskaya, V. V. 2018. Acquisition of noun derivation in Estonian and Russian L1. *Eesti Rakenduslingvistika Ühingu Aastaraamat* (Yearbook of Society of Estonian Applied Linguistics), 14: 23–40. <https://doi.org/10.5128/ERYa14.02>
- Clark, E. 2014. Acquisition of derivational morphology. In *Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 424–442. Oxford: OUP.
- Dressler, W. U. 2005. Morphological typology and first language acquisition: some mutual challenges. In *Morphology and Linguistic Typology*, G. Booij, E. Guevara, A. Ralli, S. SgROI & S. Scalise (eds), 7–20. Bologna: Università degli studi di Bologna.
- Dressler, W. U. 2011. The rise of complexity in inflectional morphology. *Papers and Studies in Contemporary Linguistics* 47: 159–176. <https://doi.org/10.2478/psicl-2011-0013>
- Dressler, W. U., Kilani-Schoch, M. & Klampfer, S. 2002. How does a child detect morphology? Evidence from production. In *Morphological Structure in Language Processing* [Trends in Linguistics. Studies and Monographs 151], H. Baayen & R. Schreuder (eds), 391–425. Berlin: Mouton de Gruyter.
- Dressler, W. U. & Merlini Barbaresi, L. 1994. *Morphopragmatics*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110877052>
- Dressler, W. U., Sommer-Lolei, S., Korecky-Kröll, K., Argus, R., Dabašinskienė, I., Kamandulytė-Merfeldienė, L., Johansen Ijäs, J., Kazakovskaya, V. V., Laalo, K. & Thomadaki, E. 2019. First-language acquisition of synthetic compounds in Estonian, Finnish, German, Greek, Lithuanian, Russian and Saami. *Morphology* 29(3): 409–429.
<https://doi.org/10.1007/s11525-019-09339-0>

- Dziubalska-Kolaczyk, K. (ed.) 2014. Special volume on Linguistic Complexity. *Papers and Studies in Contemporary Linguistics* 50(2).
- Erelt, M. (ed.) 2003. *Estonian Language* [Linguistica Uralica Supplementary Series]. Tallinn: Estonian Academy Publishers.
- Gleason, J. B. 1958. The child's learning of English morphology. *Word* 14(2–3): 150–177. <https://doi.org/10.1080/00437956.1958.11659661>
- Kasik, R. 2012. Nulltuletus ja konversioon eesti keeles (Zero-derivation and conversion in Estonian). *Keel ja Kirjandus* 11: 793–806.
- Kasik, R. 2015. *Sõnamoodustus* (Word formation). Tartu: Tartu Ülikooli Kirjastus.
- Kerge, K. 1991. Morfopragmaatilised seigad eesti tuletuses. (Morphopragmatics of Estonian derivation). In *Eesti keele ja kirjanduse kateedri töid 2*, J. Õispuu & K. Kerge (eds), 26–30. Tallinn: TPedI.
- Kerge, K. 2003. Keele variatiivsus ja *mine*-tuletus allkeelte süntaktilise keerukuse tegurina (Linguistic varieties and regular nominalization as a factor of their syntactic complicity). *TPÜ humanitaarteaduste dissertatsioonid* (Dissertations in Humanities of Tallinn Pedagogical University) 10. Tallinn: TPÜ kirjastus.
- Kerge, K. 2016. Word-formation in the individual European languages: Estonian. In *Word-Formation. An International Handbook of the Languages of Europe* [Handbooks of Linguistics and Communication Science 40.5], P. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds), 3228–3259. Berlin: De Gruyter. <https://doi.org/10.1515/9783110424942-009>
- MacWhinney, B. 2000. *CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum Associates.
- Nagy, W. E., Diakidoy, I. N. & Anderson, R. C. 1993. The acquisition of morphology: Learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading Behaviour* 25(2): 155–170. <https://doi.org/10.1080/10862969309547808>
- Padrik, M. 2010. Word-Formation Skills in Estonian Children with Specific Language Impairment [Dissertationes Pedagogicae Universitatis Tartuensis 8]. Tartu: Tartu Ülikooli Kirjastus.
- Protassova, E. & Voeikova, M. 2007. Diminutives at the early stages of acquisition. In Savickienė & Dressler (eds), 43–72. <https://doi.org/10.1075/lald.43.03pro>
- Sahka, H. 2011. *Teine grammatika. Eesti keele teonimede süntaksikonstruktsioonipõhised perspektiivis* (The syntax of Estonian action nominalizations from the constructional perspective). Tallinn: TLÜK.
- Savickienė, I. 2007. Form and meaning of diminutives in Lithuanian child language. In Savickienė & Dressler (eds), 13–42. <https://doi.org/10.1075/lald.43.02sav>
- Savickienė, I. & Dressler, W. U. (eds) 2007. *The Acquisition of Diminutives* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Vare, S. 2012. *Eesti keele sõnapered I* (Word families in Estonian). Tallinn: Eesti Keele Sihtasutus.
- Zurek, W. H. (ed.) 1990. *Complexity, Entropy and the Physics of Information*. Redwood City CA: Addison Wesley.

Derivation in Finnish child speech and child-directed speech

Klaus Laalo

University of Tampere

This chapter examines the early phases of the acquisition of Finnish derivational morphology: what kind of derivational types are used in early child speech and child-directed speech? Which types emerge first and why these types? The analysis is based on recorded and transcribed material of two Finnish-speaking children and their caregivers. In addition, some diary data are used to illustrate the emergence of innovative derived words and the productivity of different derivational types. The study shows the importance of productivity and frequency to the emergence of Finnish derivational types can best be observed in verbs and adjectives: the derivational types which are most frequently used and productive in Finnish emerge early and are frequently used in child speech.

Keywords: morphology, child language, derivation, Finnish, neologism

1. Introduction

This study examines the early phases of the acquisition of Finnish derivational morphology. Both derivation and composition are productive word-formation processes in Finnish. Until now, derivation in Finnish child language has been studied mostly from the perspective of innovative word formation: Lieko (1998), Vänntilä (1998), Laalo (2011: 254–268), Nygrén (2019), Savonen (2019) and Silvennoinen (2014) have published material about the neologisms of their own children. Besides neologisms, word formation in Finnish child language has been studied from the perspective of nominal compounding (Laalo 2017) and diminutives (Laalo 2007).

Finnish is a Finno-Ugric language with a rich morphology, mainly suffixation. The language is synthetic rather than analytic; it is semi-agglutinative with morphophonological variation in many inflectional patterns and categories, both stem-types and suffixes, and it has a wide variety of derivational morphemes. Both derivation and compounding are very productive word-formation processes in spoken and written Finnish.

Finnish nouns have about 70 derivational suffix types including agents (jA-suffix,¹ e.g. *aja-ja* ‘driver’ ← *ajaa* ‘drive’, *heittä-jä* ‘thrower’ ← *heittää* ‘throw’), and instruments (deverbal suffix *-in* : *-ime-*, e.g. *ava-in* : *ava-ime-n* ‘key, opener’ [nominative : genitive] ← *avata* ‘open’). Verbs have about 40 derivational suffix types, among them such very productive ones as the causative *-TTA*, frequentative *-ELE* and reflexive *-U*. Adjectives have more than 10 derivational suffix types, and some of them are complex. The most productive of them is *-inen*, which is also included in many adjective-noun complex derivational suffixes. Others are *-isA*, *-kAs*, *-kkA*, *-vA*, *-tOn*, and *-hkO*. Adverbs have about 10 derivational suffix types: the most productive is *-sti*, which expresses manner and can be added to all adjectives; others include *-lti*, *-itse*, *-(i)ttain*, etc. (ISK 2004). Finnish derivation is mostly additive or modificatory, but there is also some conversion between nouns and verbs.

The main research questions are:

1. Which derivational patterns emerge and are acquired first?
2. What does the early emergence depend on? Does it depend on the type or token frequency in the input? On productivity? On the transparency of the derivational patterns? On the simplicity or shortness of the forms?
3. What is the relation between derivation, compounding and inflection in terms of emergence and productivity?²

2. Data and method

The data for the quantitative analysis is based on the recorded and transcribed material of two Finnish-speaking children, a girl called Mari and a boy called Tomi, and their caregivers. The girl’s transcripts are from age 1;7–3;4 and the boy’s from age 1;7–3;1. In addition, some diary data from these two children are used to illustrate the emergence and use of innovative derived words; the neologisms are helpful in evaluating the onset of the productive use of derivatives and the productivity of different derivational patterns. Furthermore, both published (Lieko 1998; Laitsaari & Laitsaari 1996; Vänntilä 1998) and unpublished (Nygrén 2019; Savonen 2019; Silvennoinen 2014) comparative material from other Finnish-speaking children including mostly neologisms is taken into account for the sake of illustration.

1. Capital letters in the derivational suffix indicate allomorphic variation: *-jA* represents the allomorphs *-ja* and *-jä*; *-U* the allomorphs *-u* and *-y*; *-ELE* the allomorphs *-ele*, *-ile*, *-el* and *-il*; *-TTA* the allomorphs *-tta*, *-ttä*, *-ta* and *-tä*; etc.

2. The criterion of productivity used was as follows: at least three combinations of a derivational suffix and a base must occur in the corpora.

The duration of Mari's monthly recordings varies from 30 minutes (recordings made at 1;10 and 2;3) to 120 minutes (1;8 and 2;1), and Tomi's from 30 minutes (1;6, 1;7 and 2;1) to 60 minutes (his other recordings). Mari's diary data consists of about 9500 utterances and Tomi's of about 6700 utterances.

It is difficult to give the exact number of derived words, because there are many borderline cases of several types both in the data of this study as well as in Finnish in general. First, there is a continuum from clear participles to lexicalized adjectives – from these only the clearly lexicalized ones are presented as adjectives here. Second, there are several opaque derived words such as *eläin* 'animal' (cf. *elää* 'live') and *aukko* 'opening' (cf. *auki* 'open [adverb]', *avata* 'open [verb]'); many derivation etymologies are not certain, especially in the case of opaque words. Third, for certain derivation types, such as *-eA*, there is no clear base word, for example *vihreä* 'green' and *korkea* 'high'; these are not included as derivatives in the analysed data.

To evaluate the onset of the productive use of derivational suffixes and thus to compare derivation with the onset of inflection and compound use, it is interesting to consider the many neologisms in the diary material. There are only a few neologisms in the recordings (e.g. *kiehu-ttaa* 'to make boil' ← *kiehua* 'boil'), and some of them are derived adjectives as compounds such as *heikko+jalka-inen* 'weak-legged' and *vahva+jalka-inen* 'strong-legged'. The full repertoire of the child is not realized in recordings and thus the emergence and the productive use – the criterion for productivity: at least three combinations of a base and affix – can be attested only with the most frequent derivative suffixes.

3. The acquisition of derivational categories in Finnish

3.1 Derivational categories of nouns

There are about 30–40 noun derivative suffix types in the recordings; the exact number depends on the grouping criteria, for example whether the patterns in Section 3.1.5. should be counted as one pattern ending in *-Vs* and having several subtypes (as has been done in this article) or as many separate derivative patterns.

3.1.1 *Agent nouns*

Finnish agent nouns are derived from verbs mostly with the *jA*-suffix. The agent nouns in the recordings are presented in Table 1 for Mari and in Table 2 for Tomi. The first occurrences in child speech (CS) and the first recorded occurrences in child-directed speech (CDS) are shown.

Table 1. Mari's agent nouns in the recordings

Agent noun (← base verb)	First rec. occurrences	
	CS	CDS
<i>valmistaja</i> 'producer' (← <i>valmistaa</i> 'produce')	–	1;8
<i>pötköttäjä</i> 'sprawler' (← <i>pötköttää</i> 'sprawl')	–	2;3
<i>piirtäjä</i> 'drawer' (← <i>piirtää</i> 'draw')	–	2;4
<i>ratsastaja</i> 'rider' (← <i>ratsastaa</i> 'ride')	–	2;4
<i>ohjaaja</i> 'driver' (← <i>ohjata</i> 'steer')	3;0	3;0
<i>kasvis+syöjä</i> 'vegetarian' (← <i>kasvis</i> 'vegetable' + <i>syödä</i> (eat.INF) 'to eat')	3;2	–
<i>lihan+syöjä</i> 'meat eater' (← <i>liha-n</i> (meat-GEN) 'meat' + <i>syödä</i> (eat.INF) 'to eat')	3;2	–

In Mari's early recordings all derived agent nouns appear only in CDS. Mari's first derived agent noun, namely *ohjaa-ja* 'driver' (← *ohjata* 'steer') emerges at 3;0. In the recording at 3;2, Mari uses two agent nouns, namely compounds with the second part consisting of the agent noun *syö-jä* 'eat-er'.

Table 2. Tomi's agent nouns in the recordings

Agent noun (← base verb)	First rec. occurrences	
	CS	CDS
<i>ampuja</i> 'shooter' (← <i>ampua</i> 'shoot')	2;7	–
<i>ohjaaja</i> 'driver' (← <i>ohjata</i> 'steer')	2;9	2;11
<i>hoitaja</i> 'nurse' (← <i>hoitaa</i> 'take care of')	2;11	2;11

For Tomi, the first agent noun is an occasional innovative derived word which is used only in CDS in the recording at 2;1: *lauttaili-ja* 'somebody on the ferry' (*lautta* 'ferry' → *lautta-illa* 'to be on a ferry' → *lauttaili-ja*). Six months later, more agent nouns are used, and the first one (*ampuja* 'shooter') in CS is in the recording at 2;7. In CDS, there are some more such as *aja-ja* 'driver' (← *ajaa* 'drive') and *lähtijä* 'goer' (← *lähteä* 'go') at 2;7, *kuljettaja* 'driver' (← *kuljettaa* 'transport') at 3;0, and *metsästäjä* 'hunter' (← *metsästää* 'hunt') at 3;1.

In the diary data, there are some neologisms showing that the agent nouns are in productive use relatively early, in Mari's diary data already at 2;1 when there are not yet any CS-tokens in the recordings: Mari uses *ilostu-ja* 'one who gets delighted' (← *ilostua* 'be delighted' ← *ilo* 'joy') at 2;1, and Tomi uses *kuoletta-ja* 'killer' (← *kuolettaa* 'kill' (neologism) ← *kuolla* 'die') at 2;11 instead of the established agent noun *tappa-ja* 'killer' (← *tappaa* 'kill').

3.1.2 Instrument nouns

The most productive derivational suffix for instrument nouns is *-in* (the variant *-ime-* in most inflectional forms, also in the plural nominative *-ime-t*). It is attached to verb stems: *avata* ‘open’: *avaa* ‘opens’ → *ava-in* ‘key’. With other derivative suffixes, for example *-Uri* and *-ke*, instrument nouns can be derived from noun bases. The first occurrences of instrument nouns in the recordings are presented in Tables 3 and 4.

Table 3. Mari’s instrument nouns in the recordings

Instrument noun (← base verb)	First rec. occurrences	
	CS	CDS
<i>avain</i> ‘key’ (← <i>avata</i> ‘open’)	1;9	1;9
<i>nauhuri</i> ‘recorder’ (← <i>nauha</i> ‘tape’)	3;0	2;4
<i>puhelin</i> ‘telephone’ (← <i>puhella</i> ‘talk’)	3;2	3;2

In CDS, there are some additional instrument nouns: *istu-in* ‘seat’ (← *istua* ‘sit’) at 2;3, and *soit-in* ‘musical instrument’ (← *soittaa* ‘play’) and *puserr-in* ‘squeezer’ (← *pusertaa* ‘press, squeeze’) at 3;0.

Table 4. The instrument nouns in Tomi’s recordings

Instrument noun (← base verb)	First rec. occurrences	
	CS	CDS
<i>avain</i> ‘key’ (← <i>avata</i> ‘open’)	1;10	1;10
<i>helistin</i> ‘rattle’ (← <i>helistää</i> ‘jingle’)	2;4	2;4
<i>puhelin</i> ‘telephone’ (← <i>puhella</i> ‘talk’)	2;11	2;11
<i>ohjain</i> ‘control unit’ (← <i>ohjata</i> ‘steer’)	3;0	–
<i>mittari</i> ‘thermometer’ (← <i>mitata</i> ‘measure’)	2;4	2;4

Most of the instrument nouns in the recordings have been derived with the *in*-suffix but there are others also such as (*kuume*)*mittari* ‘thermometer’ derived from the verb *mitata* ‘measure’ using the *Ari*-scheme, occurring first in Tomi’s recording 2;4 once in CDS and three times in CS and once in CS at 3;0. The noun-based *haarukka* ‘fork’ (← *haara* ‘branch’) belongs to another types of instrument noun; it occurs in Tomi’s recording at 1;8 with seven tokens in CDS and two tokens in CS (and once in Mari’s recording at 1;8 but only in CDS). Also plausible is *teline* ‘rack’ (← *tela* ‘roller, cylinder’) in Tomi’s recording at 1;10 with four tokens in CDS and one token in CS.

In the diary data, there are innovative instrument nouns showing that the pattern is in productive use in Mari's speech from the age of 2;4 on and in Tomi's speech from the age of 2;8 on. Mari's first attempted instrument noun was *leikka-mi-t* 'scissors' (derived from the verb *leikata* 'cut; the correct formation would be *leikka-ime-t*) at 2;3. Her correctly derived innovations were *jarrut-in* 'an instrument for braking' (← *jarruttaa* 'brake') at 2;4, *huitais-in* 'an instrument for swiping' (← *huitaista* 'swipe') at 2;8, *aja-in* 'vehicle' (← *ajaa* 'drive') and *pyörit-in* 'an instrument for rotating' (← *pyörittää* 'rotate') at 2;9, *silit-in* 'an instrument for ironing' (← *silitää* 'iron') at 3;0, and *perj-in* 'an instrument for cleaning fish' (← *perata* 'clean fish') at 3;9.

Tomi produced his first innovative instrument nouns a little later: *silit-in* 'an instrument for ironing' (← *silitää* 'iron') and *ampu-mus* 'gun; cannonball' (← *ampua* 'shoot') at 2;8, *kahva-ime-t* 'handles of the mixer' (= contamination: *kahva* 'handle' + *vatka-in* 'mixer') at 3;0, *kuulut-in* 'loudspeaker' (← *kuuluttaa* 'announce') at 3;5, *pöly-n+pyyh-in* 'dust wiper' (← *pöly* (dust.GEN) 'dust' + *pyyhkiä* 'wipe' + *in*) at 3;6, *kuul-ime-t* 'headphones' (Standard Finnish *kuulokkeet* ← *kuulla* 'hear') at 3;8, and *pyörit-in* 'gyroscope' (← *pyörittää* 'make rotate') and *auto+aja-mus* 'car driving device' at 3;9.

Tomi's innovation *pysy-ke* 'holder' (← *pysyä* 'hold fast'; cf. Standard Finnish *pidi-ke* 'holder') was in use for a long time, from the age of 3;3 to 3;10. His other innovation *paini-ke* 'small weight' (← *painaa* 'weight; press; print') at 3;10 belongs to the same derivational pattern. It is semantically innovative: the meaning 'small weight' is derived from the base *paino* 'weight', but there is also an established word *painike* which has the meaning 'button'.

The derivational pattern *ava-in*, *helist-in* for instruments is used in the neologisms of many children, but they do not belong to the very early derivatives. Examples include Elina's *lakas-in* 'brush' at 2;5, *hiiht-ime-t* 'skis' at 3;4, and *veiva-in* 'shaft' at 4;10 (Lieko 1998: 553); Alina's *pyörit-in* 'propeller' at 3;0 and *punn-in* 'scales' at 3;11 (Riionheimo 2002: 433); Evert's *valost-ime-t* 'lightening devices' at 6;0 (Silvennoinen 2014: 33); Noora's *ruuv-in* (← *ruuvimeisseli*) at 2;10 and Sonja's *sout-ime-t* at 5;3 (Nygrén 2019: 54); Sallamari's *lent-ime-t* 'wings' at 3, Matti's *puhut-in* 'microphone' at 4 (Laitsaari & Laitsaari 1996).³

3.1.3 Derived nouns expressing the result or the action/process

The derivational suffixes *-O* and *-U* are used for nouns expressing the end result or the action or process. They are attached to verb stems. Table 5 presents Mari's data and Table 6 Tomi's data.

3. In Laitsaari & Laitsaari (1996), age information is given only by year, not by month.

Table 5. The first *O*- and *U*-derivatives in Mari's recordings

Derivative (← base)	First rec. occurrences	
	CS	CDS
<i>keitto</i> 'soup' (← <i>keittää</i> 'cook')	1;8	1;8
<i>pesu</i> 'washing' (← <i>pestä</i> 'wash': <i>peseä</i> 'washes')	1;9	2;4
<i>hoito</i> 'care' (← <i>hoitaa</i> 'take care of')	2;5	2;5
<i>meno</i> 'going' (← <i>mennä</i> 'go': <i>meneä</i> 'goes')	3;0	2;11
<i>peitto</i> 'quilt' (← <i>peittää</i> 'cover')	3;4	1;7

These suffixes are used also in compounds: *lent-o+kone* 'aeroplane' (literally 'flying+machine') in CS and CDS at 1;9, *sout-u+vene* 'rowing-boat' in CS at 1;11, etc.

The type frequency of the derivational suffixes *-O* and *-U* in CDS is high – for example, *heitt-o* 'throwing' (← *heittää* 'throw'), *jatk-o* 'continuation' (← *jatkaa* 'continue'), *käsittel-y* 'handling' (← *käsitellä* 'handle'), *laul-u* 'song' (← *laulaa* 'sing'), *lep-o* 'rest' (← *levätä* 'rest'), *tek-o* 'action' (← *tehdä* 'do') etc. The token frequency is low: many suffixes occur only once in the recordings.

Table 6. The first *O*- and *U*-derivatives in Tomi's recordings

Derivative (← base)	First rec. occurrences	
	CS	CDS
<i>hyppy</i> 'jump' (← <i>hypätä</i> : <i>hyppää</i> 'jump')	1;10	1;10
<i>alku</i> 'beginning' (← <i>alkaa</i> 'begin')	2;8	2;1
<i>lepäily</i> 'rest' (← <i>lepäillä</i> 'rest': <i>lepäilee</i> 'rests')	2;9	2;9
<i>ajelu</i> 'driving' (← <i>ajaa</i> ~ <i>ajella</i> 'drive')	2;11	–

Used also in compounds: *lent-o+kone* 'aeroplane' in CS and CDS at 1;10, *kein-u+tuoli* 'rocking chair' in CS and CDS at 2;1, and *syött-ö+tuoli* 'high chair' (lit. feeding+chair) in CS and CDS at 2;2.

In Tomi's recordings, the type frequency in CDS is also high, for example *alk-u* 'beginning' (← *alkaa* 'begin'), *heitt-o* 'throwing' (← *heittää* 'throw'), *jatk-o* 'continuation' (← *jatkaa* 'continue'), *kylp-y* 'bath' (← *kylpeä* 'take a bath'), *laul-u* 'song' (← *laulaa* 'sing'), *lep-o* 'rest' (← *levätä* 'rest'), *nous-u* 'rising' (← *nousta* 'rise'), *näyttel-y* 'exhibition' (← *näyttää* 'show'), *pallottel-u* 'catches' (← *pallotella* 'play ball'), *soitt-o* 'playing' (← *soittaa* 'play'), *vet-o* 'pulling' (← *vetää* 'pull') etc. The token frequency is mostly only one or two tokens cumulatively in all recordings.

These types of noun derivatives examined in Subsection 3.1.3 are not very productive: there are no neologisms in the data.

3.1.4 Nouns derived with *-e* from verbs and adjectives

Mari's first *e*-derivative is the deadjectival *kuum-e* 'temperature' (← *kuuma* 'hot') in CS and CDS at 1;7. Her next *e*-derivatives are deverbal: *peit-e* 'cover(ing)' (← *peittää* 'cover'; synonym: *peitt-o*) is used in CS and in CDS at 1;9, and *sad-e* 'rain' (← *sataa* 'rain') at 2;6. There are some opaque *e*-derivatives as well: *kone* 'machine' at 3;0 and *esine* 'object' at 3;2.

There are only a few neologisms of this pattern in the diary data, such as Mari's *päivä+nokkareet* 'afternoon nap' (cf. the established word *päivä+nokoset* 'afternoon nap', literally 'day sleeping') at 1;9. Both *nokoset* and *nokkareet* are plural forms.

In Tomi's recordings *e*-derivatives emerge only later: *liik-e-nne* 'traffic' (← *liik-kua* 'move') at 2;3, and *korist-e* 'decoration' (← *koristaa* 'decorate') and 2;8 *maust-e* 'spice' (← *maustaa* 'spice, season') at 2;8.

3.1.5 Deverbal nouns formed with *-s/-kse*⁴

A very productive morpheme for deriving nouns from verbs is *-s/-kse-* (stem-final *-kse-* > *-s* at the end of the syllable, for example *ohja-us* 'steering', genitive *ohja-ukse-n*, partitive *ohja-us-ta*). There are several subtypes with different phonemes preceding the *-(k)s(e)*, for example *-Os* (*kierr-os* 'round'), *-Us* (*ohja-us* 'steering'), and *-mUs* (*sopi-mus* 'agreement'); the morpheme has several functions.

One function is to express the result or process expressed by the verb, for example, *uupu-a* 'get tired' → *uupu-mus* 'exhaustion, fatigue'. In CS the *mUs*-suffix is often used instead of the longer and more complex *minen*-suffix which is very productive in adult language. The *mUs*-suffix is used also in children's neologisms: Lieko (1998: 553, 562) mentions Elina's five neologisms, for example, *seiso-mukset* 'pedestals' at 3;7; Savonen (2019: 40) Nooa's *hyppi-mys* 'jumping' at 3;8, and Väänttilä (1998: 116) Erkkä's *töötty-mys* 'horn' and *ampumus* 'shooting' at 2;9, among others.

Because the *-s/-kse-* derivatives are three-syllabic, in Mari's recordings the first *-s/-kse-* derivatives emerge only after the trochaic stage (that is, the tendency to use only of mostly two-syllabic word-forms) was over: *muhenn-os* 'stew' (← *muhentaa* 'stew, mash') in CDS and CS at 1;11, *sukell-us* 'diving' (← *sukeltaa* 'dive') in CDS and CS at 2;4, and *rakenn-us* 'building' (← *rakentaa* 'build') in CDS at 2;4.

The first *-s/-kse-* derivatives in Tomi's recordings were *ohja-us* 'steering' (← *ohjata* 'steer') in CDS at 1;8, *kierr-os* 'circle, round' (← *kiertää* 'move around') at 1;9, and *sopi-mus* 'agreement' (← *sopia* 'agree') at 2;2. Tomi himself used this pattern only later: *kat-os* 'shelter' at 2;6, *ilotulit-us* 'fireworks' at 2;7, *ohja-kse-t* 'reins' at 2;9, and *hälyt-ys* 'alarm' and *kaver-ukse-t* 'companions' at 3;0.

4. The dashes are used to indicate the distribution of the variants in the derivative suffix: *-s* is used in the final position in the nominative and *-kse-* is used in other case forms before the inflectional suffixes.

The neologisms from the diary data show the productivity of this type: Mari's *tuuppa-ukse-t* 'pushings' at 2;5 and *kiltte-yks-i-ä* 'nice actions' at 2;8. Interestingly enough, the *mUs*-derivatives that are not very productive usually in present-day Finnish are productive in CS – for example, Mari's *etsi-mys+matka* 'seeking trip' at 3;0 and Tomi's *pelaa-mus+pallo* 'playing ball' and *paijaa-mus* 'cuddling' at 2;4, *ampu-mus* 'shooting' at 2;9, and *sylke-mys* 'spitting' and *piirtä-mys+paperi* 'drawing paper' at 3;3. It is also interesting that Tomi's first *mUs*-neologism is in the diary data already at the age of 2;4 but his first conventional *s*-derivative in the recordings, *katos* 'shelter', occurs only at 2;6.

3.1.6 *-la* attached typically to noun stems to express location

The word *mummi-la* 'home of the grandmother', derived from *mummi* 'grandmother', is found in the recordings of both children (Mari in CS and CDS at 1;11, Tomi in CDS at 2;9). Tomi uses the word *vanki-la* 'prison' (← *vanki* 'prisoner') in a play situation in the recording at 3;0. In the diary data Mari has such neologisms as *pimento-la-an* 'to a dark place' (← *pimeä* 'dark') at 2;3, *Olli-la* 'Olli's home' and *Mörkö-lä* 'the place where Mörkö lives' at 2;6, and *Juuso-la* 'Juuso's place', *Tuuti-la* 'Tuuti's place', and *Tuuli-la* 'Tuuli's place' at 2;8. Tomi uses two neologisms, namely *vaari-la* 'home of the grandfather' (← *vaari*, father's father) at 2;9 and *ukki-la* 'home of the grandfather' (← *ukki*, mother's father) at 2;10 to stress that grandfathers also live in a place as well as grandmothers.

3.1.7 Deverbal *-minen/-mis(e)-*

The most productive derivative suffix to form nouns from verbs is *-minen* (for example *rakentaa* 'build' → *rakenta-minen* 'building process'). It can be attached to all verb stems to express the process expressed by the verb. But because the meaning is abstract, only two are used in the recordings, one in CS and one in CDS. They both are compounds in Tomi's recording at 2;4, namely *ajaa-mis+homma* (*ajaa* 'drive' -*minen+homma* 'activity') 'driving activity' in CS and *siirtä-mis+paikka* 'moving place' (*siirtää* 'move' -*minen+paikka* 'place') in CDS.

In the diary data, Mari has some derivatives of this pattern, all of them in compounds: *soutu-mis+käsi* 'the hand for rowing' (*soutaa* 'row'), *tukotta-mis+lelu* 'a toy for blocking' (*tukkia* 'block'), and *syö-mis+kepit* 'eating+sticks' (for *syö-mä+puikot* 'chopsticks'; *syö-dä* (eat-INF) 'to eat') at 2;4.

Children often use *-mUs* instead of *-minen*; see the examples in 3.1.5.

3.2 Derivational categories of verbs

3.2.1 Verbs derived with the derivative morpheme *-TTA*

One of the most productive suffixes for deriving verbs is *-TTA*, generally used also in such new words as *konmari-ttaa* ‘arrange things using the konmari method’ (← Marie Kondo), *whatsapi-ttaa* ‘use WhatsApp’ etc. The *TTA*-suffix has many functions (points a-e below) and it is so productive that some polysemy has emerged, for example the causative *nuku-ttaa* (derived from *nukkua* ‘sleep’) has such different meanings as ‘lull (the baby) to sleep’ and ‘anaesthetize (the patient)’; moreover, there is the emotional *TTA*-verb *nuku-ttaa* ‘feel sleepy’ which is homonymous with the causative *nuku-ttaa*. The *TTA*-suffix can be attached to stems of different parts of speech:

- a. very productive deverbal patterns: *-TTA* as a transitivizer
 - a1. causatives
 - (1) *lennättää* ‘make to fly’ (← *lentää* ‘fly’)
 - (2) *sammuttaa* ‘put out, turn out’ (← *sammua* ‘go out (fire, lights)’)
 - (3) *sytyttää* ‘light a fire, light a lamp’ (← *syttyä* ‘catch fire, light up’)
 - a2. factitives
 - (4) *pesettää* ‘have washed’ (← *pestä* : *pesee* ‘wash : washes’)
- b. *TTA*-derivatives expressing emotion, both denominal and deverbal
 - (5) *harmittaa* ‘annoy’ (← *harmi* ‘annoyance’)
 - (6) *väsyttää* ‘feel tired’ (← *väsyä* ‘get tired’)
- c. descriptive *TTA*-derivatives (expressing sounds, etc.); the base is often an interjection
 - (7) *höpöttää* ‘speak nonsense’ (← *höpö höpö*, descriptive expression for nonsense)
 - (8) *kaakattaa* ‘cackle, cluck, make a *kaak*-sound (like a hen)’ (← *kaak*)
 - (9) *koputtaa* ‘knock’ (← *kop*, interjection for knocking)
 - (10) *pötköttää* ‘sprawl’ (no clear base)
- d. factitive *TTA*-verbs based mainly on adjective stems
 - (11) *kuivattaa* ‘let dry, get dry, dry up’ (← *kuiva* ‘dry’)
 - (12) *lämmittää* ‘warm up’ (← *lämmiin* ‘warm’)
 - (13) *sulattaa* ‘melt’ (← *sula* ‘unfrozen, melted’)

e. denominal; very productive in neologisms

(14) *auttaa* ‘to help’ (← *apu* ‘the help’)

neologisms in CS:

(15) *viuluttaa* ‘play violin’ (← *viulu* ‘violin’)

Tables 7 and 8 present the first *TTA*-verbs in the recordings.

Table 7. Mari’s first *TTA*-verbs in the recordings

<i>TTA-verb</i> (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>auttaa</i> ‘help’ (← <i>apu</i> ‘help’ [noun])	1;7	1;7	4/4	3/5
<i>kopu[ttaa]</i> ‘knock’ (← <i>kop</i> , interjection for knocking)	1;7	1;7	2/2	3/5
<i>näyttää</i> ‘show’ (← <i>näkyä</i> ‘be visible’)	1;8	1;8	2/8	4/6
<i>harmi[ttaa]</i> ‘annoy’ (← <i>harmi</i> ‘annoyance’)	1;9	1;9	2/3	2/5
<i>pyörittää</i> ‘roll, whirl’ (TR) (← <i>pyöriä</i> ‘go round, rotate’ (INTR))	1;11	1;7	1/1	1/2
<i>herättää</i> ‘wake up’ (TR) (← <i>herätä</i> ‘wake up’ (INTR))	–	1;8	–	1/1
<i>kiehuttaa</i> ‘boil, cook’ (TR) (← <i>kiehua</i> ‘boil’ (INTR))	2;1	–	1/1	–
<i>pelottaa</i> ‘be afraid’ (← <i>pelko</i> ‘fear’ [noun])	2;1	3;2	2/9	1/1
<i>syöttää</i> ‘feed’ (← <i>syödä</i> (eat.INF) ‘to eat’)	2;1	1;11	3/5	3/6

Because developmentally Mari was in a strong trochaic stage, that is she used only or mostly two-syllabic word-forms, she shortened the first trisyllabic *TTA*-derivatives so that, for example *harmi-ttaa* was realized phonetically as *hammi* and *kopu-ttaa* as *kopu*. More trisyllabic verbs were used only after the trochaic stage was over, such as *pyöri-ttää* ‘rotate’; in a recording at 1;11, Mari uses both the base *pyörii* ‘rotates’ and the causative *pyöri-ttää* ‘to make rotate’. Mari used the base verb *pyöriä* ‘rotate’ already at the age of 1;6.

In Mari’s recordings there are altogether 43 verbs derived with *-TTA*. 30 of them in CS and 41 in CDS.

The productivity of this derivation pattern is manifested in Mari’s neologism, *kiehu-ttaa* ‘boil, cook’ (← *kiehua* ‘boil’ (INTR)) at 2;1, which she used when asking whether the fish had been cooked (instead of the established transitive verb *keittää* ‘boil, cook’):

- (16) Mari 2;1 *on-ko nii-tä kieh-u-te-ttu ?*
 be-Q they.PL-PARTIT boil-CAUS-PTCP ?
 Have they been cooked?

Table 8. Tomi's first *TTA*-verbs in the recordings

<i>TTA-verb</i> (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>auttaa</i> 'help' (← <i>apu</i> 'help' [noun])	1;8	1;8	3/4	5/14
<i>lennättää</i> 'fly' (TR) (← <i>lentää</i> 'fly' (INTR))	1;9	1;9	2/2	1/2
<i>sammuttaa</i> 'turn out' (← <i>sammua</i> 'go out' (about lights))	1;9	1;9	2/10	2/5
<i>ylittää</i> 'reach' (← <i>yli</i> 'over')	1;9	1;9	2/3	2/4

There is one *TTA*-verb in Tomi's first recording at 1;8 (*auttaa* 'to help'). The suffix is in productive use already in his next recording at 1;9 with four verbs if the criterion for productivity is that there are at least three base+affix combinations. With three verbs *-TTA* is in principle (if attempted verbs are counted) productive already in Mari's first recording but she shortened the three-syllabic verbs to trochaic forms: *kopu(tta-)*, *pötkö(ttä-)* [pökkö].

There are plenty of *TTA*-neologisms in the diary data: Mari uses *nouse-ttaa* 'raise' (← *nousta* 'rise') and *kiive-ttää* 'help to climb' (← *kiivetä* 'climb') at 1;11, *valo-ttaa* 'put the lights on' (← *valo* 'light') at 2;0, *iso-ttaa* 'make bigger' (← *iso* 'big') at 2;2, *viulu-ttaa* 'play the violin' (← *viulu* 'violin') at 2;7, etc. Tomi uses *pinko-ttaa* 'make a ping-sound' at 1;10, *hypä-ttää* 'to make jump' (← *hypätä* 'jump') at 2;1, etc.

3.2.2 Reflexive-passive verbs derived with *-U*, *-UTU* etc

There are many derivative suffixes for reflexive verbs. These suffixes are intransitivizers and the products are reflexive and passive verbs. The shortest one is *-U* (e.g. *kuul-u-u* 'is heard' ← *kuule-e* 'hears'; *paist-u-u* 'is roasted' ← *paista-a* 'fry, roast'; lengthening of the final vowel is the suffix of 3rd person singular). The reflexives produced with *-U* fit the trochaic pattern when added to bisyllabic stems, and they are used already in the first recordings in both CS and CDS. A longer derivative suffix used for morphophonemic reasons in certain stem types is *-UTU* (e.g. *laske-utu-a* 'descend' ← *laske-a* 'lower'; the final *a* is the suffix of the infinitive). The *U*-reflexives are presented in Tables 9 and 10.

Table 9. Mari's first reflexive-passive *U*-verbs in the recordings

<i>U</i> -verb (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>kaatua</i> 'fall down' (← <i>kaataa</i> 'turn over, tilt, pour')	1;7	1;7	3/4	2/2
<i>kääntyä</i> 'turn' (INTR) (← <i>kääntää</i> 'turn' (TR))	1;7	1;7	1/1	1/1
<i>kuulua</i> 'is heard' (← <i>kuulla</i> 'hear')	1;9	1;9	4/16	3/20
<i>kallistua</i> 'lean' (INTR) (← <i>kallistaa</i> 'lean' (TR))	1;11	1;11	1/1	1/1
<i>näkyä</i> 'be visible' (← <i>nähdä</i> 'see')	2;2	2;2	1/2	1/1
<i>ylettyä</i> 'reach' (← <i>yli</i> 'over')	2;6	–	1/1	–
<i>innostua</i> 'get inspired' (← <i>innostaa</i> 'inspire')	2;11	–	1/1	–
<i>päätyä</i> 'end up' (← <i>päättää</i> 'end')	2;11	–	1/1	–
<i>rauhoittua</i> 'calm down' (← <i>rauhoittaa</i> 'calm' (TR))	3;0	3;0	1/1	1/1

Mari used the base verb *kaataa* 'turn over' already at 1;6; this demonstrates the derivational character of the reflexive *kaatua* at 1;7.

Table 10. Tomi's first reflexive-passive *U*-verbs in the recordings

<i>U</i> -verb (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>kaatua</i> 'fall down' (← <i>kaataa</i> 'turn over, tilt, pour')	1;8	1;8	7/41	5/25
<i>laskeutua</i> 'descend' (← <i>laskea</i> 'lower')	1;9	1;8	1/1	2/6
<i>pelästyä</i> 'be frightened' (← <i>pelätä</i> 'be afraid of')	1;10	1;10	1/3	2/8
<i>kääntyä</i> 'turn' (INTR) (← <i>kääntää</i> 'turn' (TR))	2;2	2;4	2/4	1/1
<i>löytyä</i> 'be found' (← <i>löytää</i> 'find')	2;2	1;8	2/5	2/8
<i>maistua</i> 'taste' (INTR) (← <i>maistaa</i> 'taste' (TR))	2;6	2;6	1/1	1/1
<i>siirtyä</i> 'move' (INTR) (← <i>siirtää</i> 'move' (TR))	2;6	2;6	2/2	2/2
<i>unohtua</i> 'be forgotten' (← <i>unohtaa</i> 'forget')	2;7	–	1/1	–
<i>sulkeutua</i> 'get closed' (← <i>sulkea</i> 'close')	2;11	–	1/1	–

There are some early reflexive verb neologisms in the diary data: Tomi uses *ilma-stu-a* 'float' (← *ilma* 'air') at 1;11 and Mari uses *poltt-u-a* 'get burned' (← *polttaa* 'burn') at 2;9.

3.2.3 Verbs derived with the derivative morpheme *-ELE*

-ELE is a productive derivative suffix for frequentative-continuative verbs (e.g. *hyp-el-lä* ‘keep jumping’ ← *hypätä* ‘jump’); the base can belong to different parts of speech, but most common are verb bases. The typical meaning of deverbal *ELE*-derivatives is frequentative – for example *puh-el-la* ‘chat, converse’ (← *puhua* ‘speak’) – but this is not always the case; there can be other kinds of semantic change; for example *kuunn-el-la* ‘listen’ (← *kuulla* ‘hear’) and *pelot-el-la* ‘scare’ (← *pelottaa* ‘be afraid’). When *-ELE* is used to derive verbs from nouns, adjectives or root stems, there is often no clear frequentative meaning – for example, *pyörä-il-lä* ‘ride a bicycle’ (← *pyörä* ‘wheel’), *ilke-il-lä* ‘be malicious’ (← *ilkeä* ‘malicious’), and *ulko-il-la* ‘take outdoor exercise’ (← *ulko-* ‘out’). Whereas *TTA*-verbs typically express the active actions of an agent, *ELE*-verbs often express a more passive state such as *hygge-il-lä* ‘to relax’ (← Danish *hygge* ‘cosy’). The first *ELE*-derivatives emerge later than the first *TTA*- and *U*-derivatives. The first *-ELE*-verbs in the corpora of the two children are presented in Table 11 (Mari) and Table 12 (Tomi).

Table 11. Mari’s first *ELE*-verbs in the recordings

<i>ELE</i> -verb (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>jutella</i> ‘talk’ (← <i>juttu</i> ‘story’)	1;11	1;9	1/1	3/4
<i>heitellä</i> ‘keep throwing’ (← <i>heittää</i> ‘throw’)	2;3	2;3	2/10	2/4
<i>kävellä</i> ‘walk’ (← <i>käydä</i> ‘walk; visit’(verb))	2;3	2;4	3/8	2/8
<i>hyppellä</i> ‘keep jumping’ (← <i>hyppiä</i> ‘jump’)	2;4	2;4	1/1	1/1
<i>katsella</i> ‘look, watch’ (← <i>katsoa</i> ‘look’)	2;6	1;11	3/3	5/5
<i>sukseilla</i> ‘ski’ (neologism) (← from the noun <i>suksi</i> ‘ski’)	2;11	2;11	1/1	1/1

In the recording at 2;11, Mari uses the denominal neologism *sukse-ilee* (← *suksi* ‘the ski’) instead of the established verb *hiihtää* ‘to ski’. The neologisms in the diary data show the productiveness of the type: Mari uses *istu-ilee* ‘keeps sitting’ (← *istua* ‘sit’) at 1;11, *ilo-ilee* ‘is happy’ (← *ilo* ‘joy’) at 2;2, and *nurkka-ilee* ‘is standing in the corner’ (← *nurkka* ‘corner’) at 2;5.

There are two neologisms in Table 12: at 2;4 Tomi produced both *vilkut-ella* ‘keep waving’ (← *vilkuttaa* ‘wave’) and *purjehd-ella* ‘sail’ using *purjehtia*, the conventional verb for sailing, as a base. The productivity of the type is reflected in the fact that there are more neologisms in the diary data: Tomi uses *loikk-elee* ‘keep jumping’ (← *loikkia* ‘jump’) at 2;8 and *jätski-ttelee* ‘eat ice cream’ (← *jätski* ‘ice cream’) at 4;10.

Table 12. Tomi's first *ELE*-verbs in the recordings

<i>ELE</i> -verb (← base)	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>kävellä</i> 'walk' (← <i>käydä</i> 'walk, visit')	1;9	1;9	3/10	3/5
<i>hypellä</i> 'keep jumping' (← <i>hypätä</i> 'jump')	1;10	1;10	1/1	2/2
<i>katsella</i> 'watch' (← <i>katsoa</i> 'look')	2;2	1;10	1/1	5/12
<i>piileskellä</i> 'hide' (← <i>piillä</i> 'be hidden')	2;2	2;2	1/5	1/1
<i>heräillä</i> 'wake up slowly' (← <i>herätä</i> 'wake up')	2;4	–	1/1	–
<i>huudella</i> 'keep shouting' (← <i>huutaa</i> 'shout')	2;4	–	1/1	–
<i>purjehdella</i> 'keep sailing' (neologism) (← <i>purjehtia</i> 'sail')	2;4	2;4	1/1	1/1
<i>vilkutella</i> 'keep waving' (← <i>vilkuttaa</i> 'wave')	2;4	–	1/1	–
<i>käännellä</i> 'keep turning' (TR ← <i>kääntää</i> 'turn')	3;0	3;0	1/2	1/1
<i>kääntyillä</i> 'be tossing' (INTR ← <i>kääntyä</i> 'turn')	3;0	3;0	1/3	1/1

3.2.4 Momentaneous verbs formed with *-Ahta* and *-Aise*

Since the semantic range of momentaneous verb derivatives is not as broad as that of the *ELE*-derivatives, the former are less frequently used and less productive. There are two productive patterns of momentaneous verbs: *-Ahta* (e.g. *nuk-ahtaa* 'fall asleep' ← *nukkua* 'sleep', *horj-ahtaa* 'sway' ← *horjua* 'stagger') and *-Aise* (e.g. *potk-aisee* 'gives a kick' ← *potkia* 'kick').

In Mari's early recordings there are momentaneous verb derivatives only in CDS, all three in the same recording at 2;5: *napp-aista* 'snatch quickly' (cf. *napata* 'snatch'), *nuuhk-aista* 'sniff quickly' (← *nuuhkia* 'sniff') and *potk-aista* 'give a kick' (← *potkia* 'kick'). In the later recordings there are some also in CS, namely *lös-ähtää* 'flop' (descriptive verb, CS and CDS) at 3;0, and *nuk-ahtaa* 'fall asleep' (cf. *nukkua* 'sleep') and *säik-ähtää* 'be scared' (both only in CS) at 3;4.

In Tomi's recordings there are 10 momentaneous verb derivatives in CDS and only one in CS, namely *liik-ahtaa* 'stir' (← *liikkua* 'move'); this verb is not used in CDS in the recordings. In Tomi's diary data there are a few neologisms, for example, *auk-ahta-a* 'open quickly' (← *aueta* 'to open' : *aukeaa* 'opens') at 2;9.

3.2.5 Conversion

There are a few stems in Finnish that can be used both as nouns and verbs, for example *sylki* 'spit', *usko* 'belief', *tahto* 'will', and *toivo* 'hope'. One of them is used many times in the recordings, namely the verb *tahto-a* (want-*INF*) : *tahdo-n* (want-*1SG*), cf. the noun *tahto* 'will' (NOM) : *tahto-a* (will-*PARTIT*) : *tahdo-n* (will-*GEN*). This verb is frequently used in modal expressions (Laalo 2021).

3.3 Derivational categories of adjectives

The suffix types of the derived adjectives in the recordings are *-inen*, *-vA*, *-kAs* and *-isA*. It is possible to derive from the same base adjectives using different derivative suffixes, and thus produce slightly different meanings, such as *paino* ‘weight’ → *paina-va* ‘heavy’, *paino-kas* ‘emphatic’ and (GEN +) *paino-inen* ‘weighting as much as’. Examples include *kilon paino-inen* (kilo-GEN *paino-inen*) ‘weighting one kilo’; *terä* ‘blade’ → *terä-vä* ‘sharp’, *kaksi terää* ‘two blades’ → *kaksi+terä-inen* ‘with two blades’, *käyrä terä* ‘curved blade’ → *käyrä+terä-inen* ‘with a curved blade’.

Adjectives derived with the very productive suffix *-inen* emerge already in Tomi’s recording at 1;9 and are productive with three lemmas cumulatively in the recording at 1;10. In Mari’s recordings they are at first truncated to fit the trochaic pattern; full forms are used from the recording at 1;11 onwards. Adjectives derived with *-vA* emerge a little later, others clearly later.

3.3.1 *-inen/-is(e)-*

The most productive morpheme for deriving adjectives is *-inen*, meaning ‘such as expressed by the base’; this suffix type has also longer variants such as *-llinen*, etc.; the base can be

- a. a noun: *ilo* ‘joy’ → *iloinen* ‘joyful’, *suru* ‘sorrow’ → *surullinen* ‘sad’, *veri* ‘blood’ → *verinen* ‘bloody’, *villa* ‘wool’ → *villainen* ‘made of wool’
- b. a compound: *mielen+kiinto* ‘interest’ → *mielen+kiintoinen* ‘interesting’
- c. a NP: *iso koko* ‘big size’ → *iso+kokoinen* ‘having a big size’, *kaksi osaa* ‘two parts’ → *kaksi+osainen* ‘bipartite’, *vanha aika* ‘old times’ → *vanhan+aikainen* ‘old-fashioned’

The first *inen*-adjectives are presented in Tables 13 and 14.

Some of these adjectives have transparent bases, such as *surullinen* ‘sad’ (← *suru* ‘sadness’) and *tavallinen* ‘usual’ (← *tapa* ‘manner’). For some others, the base is not in frequent use but is used in certain expressions, such as *puna poskilla* ‘redness, glow on the cheeks’ (cf. *puna-inen* ‘red’), or in compounds, such as *valko+peippi* ‘white dead-nettle’ and *valko+pyykki* ‘white laundry’ (cf. *valko-inen* ‘white’).

Because developmentally Mari was in a strong trochaic stage, she shortened three-syllabic adjectives to two syllables at first. Her compound adjectives at the age of 2;2 and 2;3 are neologisms used to characterize some of her toy animals.

The productivity of these adjectives is demonstrated by the neologisms in the diary data: Mari uses *reikä-inen* (← *reikä* ‘hole’) ‘full of holes’ at 3;9, *piraija-inen vesi* (← *piraija* ‘piranha’) ‘water with piranhas’ at 4;4, and *krokotiil-inen vesi* (← *krokotiili* ‘crocodile’) ‘water with crocodiles’ at 4;10.

Table 13. Mari's first *inen*-adjectives in the recordings

Adjective	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>puna[inen]</i> 'red'	1;7	1;7	11/29	14/26
<i>suru[llinen]</i> 'sad'	1;7	1;7	2/4	4/4
<i>kelta[inen]</i> 'yellow'	1;8	1;8	12/30	11/32
<i>pikkuinen</i> 'little'	1;11	1;8	6/14	8/24
<i>sininen</i> 'blue'	1;11	1;8	7/20	11/26
<i>valkoinen</i> 'white'	1;11	1;11	6/10	8/17
<i>tavallinen</i> 'usual'	2;1	–	2/2	–
<i>pitkä+kaulainen</i> 'long-necked' (← <i>pitkä</i> 'long' + <i>kaula</i> 'neck'+ <i>inen</i>)	2;2	–	1/1	–
<i>heikko+jalkainen</i> 'weak-legged' (← <i>heikko</i> 'weak' + <i>jalka</i> 'foot'+ <i>inen</i>)	2;3	2;3	1/3	1/6
<i>vahva+jalkainen</i> 'strong-legged' (← <i>vahva</i> 'strong' + <i>jalka</i> 'foot'+ <i>inen</i>)	2;3	2;3	1/1	1/2

Table 14. Tomi's first *inen*-adjectives in the recordings

Adjective	Age of first occurrence		All occurrences (types/tokens)	
	CS	CDS	CS	CDS
<i>punainen</i> 'red'	1;9	1;8	8/53	10/45
<i>pikkuinen</i> 'little'	1;10	1;8	5/43	9/26
<i>väsynen</i> 'tired' (neologism, standard expression for 'tired' is the participle <i>väsynyt</i> ← <i>väsyä</i> 'get tired')	1;10	1;10	1/2	1/2
<i>keltainen</i> 'yellow'	1;11	1;11	3/9	3/9
<i>sininen</i> 'blue'	1;11	1;11	6/23	6/17
<i>valkoinen</i> 'white'	1;11	1;8	2/2	3/8
<i>iloinen</i> 'happy'	2;2	1;9	1/1	3/4
<i>vihainen</i> 'angry'	2;3	2;3	2/6	2/5
<i>viimeinen</i> 'the last one'	2;7	2;6	1/3	1/1
<i>vaaleanpunainen</i> 'pink'	2;8	3;1	1/1	1/1

3.3.2 -vA

The *vA*-suffix expresses that something has qualities expressed by the base, for example, *ete-* ‘front-’ → *ete-vä* ‘talented’ and *mahti* ‘power’ → *mahta-va* ‘powerful, mighty’. Furthermore, there are *vA*-adjectives that are lexicalized present participles: *paina-va* ‘heavy’ (cf. *painaa* ‘weight’), *seuraa-va* ‘next’ (cf. *seurata* ‘follow’) and *taita-va* ‘skillful’ (cf. *taitaa* ‘master’).

Mari has four *vA*-adjectives in the recordings, etymologically participles, namely *sopi-va* ‘suitable’ (← *sopia* ‘fit, suit’) at 1;7, *paina-va* ‘heavy’ (← *painaa* ‘weight’) at 1;11, *syötä-vä* ‘edible’ (← *syödä* ‘to eat’) at 2;6, and *elä-vä* ‘living’ at 3;0. In CDS, there are eight other *vA*-adjectives.

Tomi has two *vA*-adjectives in the recordings: *sopi-va* ‘suitable’ at 1;11 and *paina-va* ‘heavy’ at 2;8. In CDS there are eight other *vA*-adjectives, for example *ete-vä* ‘eminent’, *jännittä-vä* ‘exciting’, *kaare-va* ‘arched’, and *taita-va* ‘skillful’.

At 3;0, Mari produced one neologism belonging to *vA*-adjectives, namely *kelta-va* (following with the established adjective *kellertä-vä*) ‘yellow-like’ based on the same stem as the adjective *kelta-inen* ‘yellow’ when speaking about trees in the autumn:

- (17) Mari 3;0:
ne on kelta-v-i-a, ne on
 they be.3PL yellow-DERIV-PL-PARTIT, they be.3PL
kelle-rtä-v-i-ä
 yellow-DERIV-PL-PARTIT (-v- = participle or derivative element)
 ‘They are yellowish, they are yellow like’

3.3.3 -kAs and -isA

The suffixes *-kAs* and *-isA* express an abundance of qualities expressed by the base, for example

- (18) *nero* ‘genius’ → *nero-kas* ‘ingenious’
 (19) *touhu* ‘bustle’ → *touhu-kas* ‘busy, energetic’
 (20) *vitsi* ‘joke’ → *vitsi-käs* ‘funny’
 (21) *valo* ‘light’ → *valo-isa* ‘well-lighted, sunny’

There are only a few adjectives in the recordings that have these suffixes, and only one in CS: Mari uses *valo-isa* ‘well-lit’ at 2;1 (also in CDS). In Mari’s recording there is one in CDS, namely *nero-kas* ‘ingenious’ at 2;6. In Tomi’s recordings, there are four *kAs*-adjectives in CDS.

The productivity of these suffixes is demonstrated by neologisms in the diary data, for example Mari’s *raivo-kas* ‘furious’ (the corresponding established adjective is *raivo-isa*) at 2;10.

4. Derivational morphology acquired at later stages

The very productive *-minen*: *-mis(e)-* can be attached to all verb stems, and it has the same meaning as the English *-ing* suffix. Because of the abstract meaning, it is more used in written than in spoken Finnish to derive action nouns from verbs – it is sometimes even overused in writing to replace verbs, which is called “paper-tasting noun sickness”. In CS it is used often as a non-head in compounds such as Tomi’s *aja-mis+homma* ‘driving activity’ (*ajaa* ‘drive’-*mis* + *homma* ‘activity’) and *siirtä-mis+paikka* ‘moving place’ (*siirtää* ‘move’-*mis* + *paikka* ‘place’) at 2;4, and Mari’s *katso-mis+nälkä* ‘watching hunger’ (*katso* ‘watch’-*mis* + *nälkä* ‘hunger’) at 3;6, when she wanted to watch the television. Sometimes this kind of compound is used when the child does not know or remember the established word: for example, Mari’s *syö-mis+keppi* ‘eating stick’ (*syö* ‘eat’-*mis* + *keppi* ‘stick’) instead of the established compound *syö-mä+puikko* ‘chop stick’ (*syö* ‘eat’-*mä* + *puikko* ‘stick’) at 2;10 and *siirtä-mis+vehje* ‘moving device’ (*siirtää* ‘move’-*mis* + *vehje* ‘device’) instead of *atulat* ‘tweezers’ at 3;4. Instead of *-minen*, children often use *mUs*-derivatives, some of which are presented in 3.1.e.

The adjectives derived with *-(mA)tOn* express a lack of qualities expressed by the base; these are found first in later recordings. The very first recorded token is in Mari’s recording at 3;0 but this is only an attempt, targeting the compound adjective *asiaan+kuulu-maton* ‘irrelevant, inappropriate’ (matter+belong-NEG) but changed to *asia-ttomaan kuuluva* (matter-NEG belonging). In the same recording there is another adjective with this suffix in CDS: *käsittä-mätön* ‘incomprehensible’ (← *käsittää* ‘comprehend’). The first correct token is Mari’s *levo-ton* ‘restless’ (← *lepo* ‘rest’) at 4;1. In the diary data, there are some neologisms with this suffix.

5. Word-class changing vs. maintaining derivations

Nouns are derived from verbs with such suffixes as *-ja* for agents, *-in* for instruments, etc. The first deverbal suffixes for noun derivation in the data are *-O*, *-U* and *-s/-kse-*. In adult Finnish, the most productive suffix to derive nouns from verbs is *-minen*, but it has an abstract meaning and is used in CDS only infrequently and very seldom in CS; instead, children use one subtype of the derivational type *-s/-kse-*, namely *-mUs*. One reason might be that the inflection of *-minen* is quite complicated (*-minen/-mise-/mis-*). With the *e*-suffix nouns are derived from both verbs (*sataa* ‘rain’ → *sad-e* ‘rain’) and adjectives (*kuuma* ‘hot’ → *kuum-e* ‘fever’).

Verbs are derived from both nouns and adjectives with *-TTA* (*apu* ‘the to help’ → *auttaa* ‘help’, *lämmin* ‘warm’ → *lämmittää* ‘warm up’), and this productive suffix is used also in children’s neologisms, such as *viulu* ‘violin’ → *viuluttaa* ‘play

the violin'. A few verbs are also derived from nouns with *-ELE*, for example *sukseilee* 'is skiing' from *sukset* 'skis'.

Adjectives are derived from nouns above all with the suffix *-inen*, but a few are also derived with the suffixes *-vA*, *-kAs*, and *-isA*.

The most important word-class maintaining derivational suffixes for verbs are the causative suffix *-TTA*, the reflexive-passive suffixes *-U* and *-UTU*, and the frequentative suffix *-ELE*; all are productive in CS. The momentaneous verb suffixes *-Ahta* and *-Aise* are used only a few times and mostly in CDS.

For nouns, the word-class maintaining derivational suffixes are few, but for example *-IA* is used for locations (e.g. *mummi* 'grandmother' → *mummi-la* 'grandmother's home'). For adjectives, there are a few word-class maintaining derivational suffixes, such as the moderative suffix *-hkO* (e.g. *vanha* 'old' → *vanha-hko* 'rather old'), but these are not found in the data.

To sum up, nouns are derived from verbs with several suffixes, both productive and less productive ones. Verbs are derived from both nouns and adjectives with a few productive suffixes. Adjectives are derived from nouns with a few suffixes, only one of them being very productive. The word-class maintaining derivations are productive in verbs and to a certain extent also in nouns.

6. Conclusion

6.1 Research question 1: Which derivational patterns emerge and are acquired first?

The first derivational patterns to emerge are *U-* and *TTA-*verbs (Mari at 1;7, Tomi at 1;8), adjectives derived with the *inen-*suffix (Tomi at 1;9, Mari at 1;7 in truncated forms, full forms at 1;11), and such bisyllabic nouns as *kuume* (Mari at 1;7), *keitto* (Mari at 1;8), and *hyppy* (Tomi at 1;10) derived with *-e*, *-O*, and *-U*, respectively. Many diminutives emerge early as well, and they are used both in CS and CDS; unlike many other noun types, the diminutives have a simple and transparent inflection (Laalo 2007). The first noun and verb derivations emerge about the same time.

The first verb derivatives to emerge are transparent bisyllabic reflexive verbs derived with the *U-*suffix: *kaatuu* 'falls down' is used by Mari at 1;7 and Tomi at 1;8, and Mari uses *kääntyy* 'turns' at 1;7. Tomi uses trisyllabic reflexive verbs *laskeutuu* 'descends' already at 1;9 and *pelästyy* 'is frightened' at 1;10; Mari only after the disyllabic phase: *kallistuu* 'leans' at 1;11, *innostuu* 'gets excited' at 2;11, and *rauhoittuu* 'calms down' at 3;0. In the recordings Mari has three base-affix combinations at 1;9 (in the diary data already at 1;7), and Tomi at 1;10 (in the diary data at 1;8).

The first *TTA-*derivative to emerge is the disyllabic *auttaa* 'to help' (Mari at 1;7, Tomi at 1;8); still, this derivative is only semantically transparent but

morphologically opaque because the phoneme *p* of the short stem is deleted in the derivation: *apu* ‘the help’ + *TTA* → *auttaa* ‘to help’. Most *TTA*-derivatives have at least three syllables and emerge somewhat later, such as Tomi’s causative verbs *lennättää* ‘make to fly’ and *sammuttaa* ‘put out a lamp’ at 1;9. Because Mari had a strong trochaic stage, her first trisyllabic *TTA*-derivative emerged first at 1;11 *pyörittää* ‘make rotate’; this is a very clear causative verb derivative because Mari also uses the intransitive base *pyöriä* ‘rotate’ in the same recording. In the recordings, Tomi has four *TTA*-derivatives at 1;9; Mari has two disyllabic *TTA*-derivatives at 1;8 (and two shortened forms) but three full forms first at 1;11 due to her strong trochaic stage.

The third type of derived verbs to emerge consists of *ELE*-verbs. Tomi’s first were *kävelee* ‘is walking’ at 1;9, *hyppelee* ‘is jumping’ and *juttelee* ‘is talking’ at 1;10, and *katselee* ‘is watchig’ and *piileskelee* ‘is hiding’ at 2;2. Mari’s first *ELE*-verbs were *juttelee* at 1;11 and *katselee* at 2;1; in CDS *katselee* is the most frequently used *ELE*-verb.

Conversion appears only in one verb, namely *tahto-* ‘want to’, which is frequently used in CS in modal utterances already at an early stage (Laalo 2021). There are no neologisms created by conversion in the data. This is in contrast to Clark’s (1993) observation that English-speaking children create very early neologisms using conversion.

The most productive derivational suffix for adjectives is *-inen*. In Tomi’s data it emerges first in CDS at 1;8 with nine *inen*-adjectives having 18 tokens, and in CS soon after that: *puna(i)nen* ‘red’ at 1;9, *pikku(i)nen* ‘small’ at 1;10, and three more colour adjectives at 1;11. Mari attempts to use them already at 1;7 and produces two adjectives occurring also in CDS but shortened in CS to fit the trochaic pattern. Mari produces trisyllabic full forms of *inen*-adjectives only at 1;11, but then these adjectives develop into a fast-growing lexical group.

Of the other adjective suffix types, *vA*-adjectives (e.g. *sopiva* ‘suitable’) are used both by Mari and Tomi. There is only one *isA*-adjective in the recordings, namely Mari’s *valoisa* ‘well-lit’ at 2;1.

One early derivational group of nouns consists of diminutives which emerge at the age of 1;7 in both Mari’s and Tomi’s speech (Laalo 2007: 269). The diminutives are morphologically more transparent than their simplex counterparts and thus facilitate inflection. Their use in Finnish CDS and CS is discussed in detail in Laalo (2007).

Other derivative suffix types of nouns to emerge early are *-e* (Mari’s *kuume* ‘fever’ at 1;7) and *-O/-U*: Mari’s *keitto* ‘soup’ at 1;8 and Tomi’s *hyppy* ‘jump’ at 1;10. Other productive derivative suffix types are used later in CS:

-in (instrument nouns) Mari’s *avain* ‘key’ at 1;9, Tomi’s *avain* at 1;10; later many neologisms;

-*jA* (agent nouns) Mari's *kasvis+syöjä* 'vegetarian' and *lihan+syöjä* 'carnivore' at 3;2, Tomi's *ampuja* 'shooter' at 2;7;
 -*s* Mari's *muhennos* 'stew' at 1;11 and Tomi's *katos* 'shelter' at 2;6.

The first derivative types of verbs, adjectives and nouns are presented in Table 15.

Table 15. Emergence and productive use of certain important early derivative suffixes/types (cumulatively productive)

Derivational type	Mari		Tomi	
	Emergence	Prod. use	Emergence	Prod. use
<i>TTA</i> -verbs	1;7	1;8/1;11*	1;8	1;9
<i>U</i> -verbs	1;7	1;11	1;8	1;10
<i>ELE</i> -verbs	1;11	2;3	1;9	2;2
<i>inen</i> -adjectives	1;7/1;11*	1;8/1;11*	1;9	1;10
<i>O</i> - and <i>U</i> -nouns	1;8	1;9**/2;5	1;10	2;1**/2;8
<i>e</i> -nouns	1;7	2;6	2;3	2;8

* Mari's trochaic stage: at first the *inen*-adjectives and *TTA*-verbs are truncated; these word types are used with different stems at 1;8, but they are realized as shortened trochaic forms; full forms emerge first at 1;11.

** Productive use first in compounds (at 1;9 and 2;1), and as independent words only later.

6.2 Research question 2: What does the early emergence depend on?

Frequency, especially type frequency in CDS, is important for early emergence: from the first derivative types both *TTA*-verbs and *inen*-adjectives have a high type and token frequency in CDS; *U*- and *ELE*-verbs as well as *O/U*-nouns have a high type (but not token) frequency in CDS.

Derivative suffixes that are not frequent in CDS emerge only late if ever in CS, for example, momentaneous verbs derived with *AhtA* and *Aise*, as well as adjectives derived with *-kAs* and *-isA*.

Productivity also plays a role for early emergence: verbs derived with the productive suffixes *-TTA*, *-U*, and *-ELE* emerge early, in the same way nouns derived with *-O/-U* and adjectives derived with *-inen*, whereas less frequent and less productive adjectives with *-vA* emerge only later, and *-kAs* and *-isA* are used mostly in CDS; these suffix types are not as productive, which is reflected in their low frequency. There are some productive but not very frequently used suffix types:

- jA*-agent nouns: in the recordings of Mari at 3;0 and Tomi at 2;7, but the first neologisms in Mari's diary data are already at 2;1; Tomi's first are at 2;11;
- in*-instrument nouns: in the recordings both Mari (at 1;9) and Tomi (at 1;10) have as their first instrument noun the word *avain* 'key'; the first neologisms in the diary data appear at 2;4 for Mari and at 2;8 for Tomi.

Productivity alone is not very important for early emergence. For example, the suffix for moderative adjectives *-hkO* is in principle very productive (it can be attached to all adjectives) but was not in use in the recordings, either in CS or in CDS; in general, moderativeness is usually expressed in spoken Finnish with the construction *melko* + ADJ ‘rather + ADJ’.

Transparency also has some influence on early emergence: the derivative suffixes *-TTA*, *-U*, *-ELE*, *-inen*, and *-O/-U* are all transparent. Still, transparency alone does not lead to derivative types being used early or frequently, because both *-jA* and *-in* are transparent but used only later in CS; later their productivity is also obvious in the diary data. Frequency seems to be more important for early emergence: the verb *auttaa* ‘help’ emerges very early, as the first *TTA*-derivative verb; it is frequent but not transparent because of the exceptional loss of the *p*-sound in the derivation process (*apu* + *TTA* → *auttaa*).

Because Mari had a strong trochaic stage, it was important for her that the word had a simple and short form especially when the production of full forms is concerned, because Mari used forms which either fit or – if they were longer – were truncated by her to fit the trochaic pattern.

6.3 Research question 3: What is the relation between derivation, compounding and inflection in terms of emergence and productivity?

It is interesting to note that when at the age of 1;8 Mari’s grammatical development proceeded to protomorphology with several mini-paradigms and analogical inflectional forms in both verb inflection (Laalo 2003: 337–338, 342–344) and nominal inflection (Laalo 2009: 81–82) she also had a spurt in the use of compounds (Laalo 2017: 195). Thus, the onset of her compound production was simultaneous with the new active phase in grammatical development, and it is essential in both to combine linguistic elements with each other and produce new combinations of elements.

Because of her strong trochaic stage Mari had only a few early derivatives, namely the bisyllabic verbs *auttaa*, *kääntyy*, and *kaatuu* already at 1;7 (also many shortened derivative forms) and *näyttää*, *keitto* ‘soup’, and *keitti* (← *keittiö* ‘kitchen’) at 1;8. In the recordings, Mari uses the causative verb *syö-ttä-ä* (eat-CAUS-INF) ‘feed’ first at 2;1, but in the diary data she uses *syöttää* in both the present and past already at 1;8 (*syö-ttä-ä* (eat-CAUS-PRS.3SG) ‘feeds’ : *syö-tt-i* (eat-CAUS-PST) ‘fed’), along with the mini-paradigm of the base verb *syö* ‘eat’. In the diary data, Mari has several *U*-verbs already at the age of 1;7, namely *kuivuu* ‘gets dry’, *kuuluu* ‘is heard’, and *maistuu* ‘tastes’.

In the recordings, Tomi has only one compound at 1;8 but already three compounds at 1;9 (Laalo 2017: 198). Tomi uses the first mini-paradigms of verbs at 1;8 (Laalo 2011: 90) and of nouns also at 1;8 (Laalo 2017: 164). In the recordings, Tomi

uses the first derivative (*kaatuu*) at 1;8 and several *TTA*-verbs and *inen*-adjectives at 1;9. In the diary data, he already has several *U*-verbs at 1;7: two normal (*kaatu* ‘fell down’, *mahtuu* ‘fits’) and one fossilized (*tippu* ‘fell down’), and he has four further *U*-verbs at 1;8.

6.4 Summing up

The first derivational patterns to emerge in CS belong to types that are frequently used in CDS and fit the trochaic pattern of two syllables. The early emergence thus depends first of all on frequency, especially type frequency in CDS. Productivity also plays a role, because derived verbs, nouns, and adjectives that emerge early belong to productive derivational types. On the other hand, some very productive types (for example agent and instrument nouns) are used only later in CS. Many of the first derivation types are also transparent.

The relation between derivation, compounding, and inflection is close: both children investigated use their first derived words at the same time when they start to inflect words and use compounds. Thus, morphology seems to emerge simultaneously in all its major components.

References

- Clark, E. V. 1993. *The Lexicon in Acquisition* [Cambridge Studies in Linguistics 65]. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511554377>
- ISK 2004. *Iso suomen kielioppi* (The large grammar of Finnish) [Suomalaisen Kirjallisuuden Seuran Toimituksia 950]. Hämeenlinna: Suomalaisen Kirjallisuuden Seura (The Finnish Literature Society).
- Laalo, K. 2003. Early verb development in Finnish: a preliminary approach to miniparadigms. In *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective* [Studies on Language Acquisition 21], D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 323–350. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110899832.323>
- Laalo, K. 2007. Diminutives in Finnish child-directed and child speech. In *The Acquisition of Diminutives: A Cross-linguistic perspective* [Language Acquisition and Language Disorders 43], I. Savickienė & W. U. Dressler (eds), 263–278. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.11laa>
- Laalo, K. 2009. Acquisition of case and plural in Finnish. In *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective* [Studies on Language Acquisition 30], U. Stephany & M. Voeikova (eds), 49–89. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110217117.49>
- Laalo, K. 2011. *Lapsen varhaiskielioppi ja miniparadigmat* (The protogrammar and mini-paradigms in child language) [Suomalaisen Kirjallisuuden Seuran Toimituksia 1309]. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Laalo, K. 2017. Acquisition of compound nouns in Finnish. In *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61], W. U. Dressler, F. N. Ketrez & M. Kilani-Schoch (eds), 191–207. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61.09.laa>
- Laalo, K. 2021. Directives in Finnish language acquisition. In *Development of Modality in First Language Acquisition: A Cross-linguistic Perspective* [Studies on Language Acquisition 54], U. Stephany & A. Aksu-Koç (eds), 347–378. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9781501504457-010>
- Laitsaari, L. & Laitsaari, M. 1996. *Täivaan kynä, linnun lehti – sanoja lasten suusta* (Collection of neologisms from Finnish-speaking children). Helsinki: WSOY.
- Lieko, A. 1998. Lapsen kielen uudissanat (Neologisms in child language). *Virittäjä* 102: 550–570.
- Nygrén, S. 2019. Litikuivan lintumyräkän sydänlaulu: Lasten uudissanoista (About lexical innovations in the speech of two girls). MA thesis, Tampere University.
- Riionheimo, H. 2002. Havaintoja kahden itäsuomalaisen lapsen morfologisista poikkeamista: Uudisjohdokset (Observations from the morphology of two children from Eastern Finland: Innovative derivations). *Virittäjä* 106: 432–443.
- Savonen, T. 2019. Täältä tulee välipalaja! Tutkimus lapsenkielen johdoksista (About derivatives in child language). MA thesis, Tampere University.
- Silvennoinen, H. 2014. Lentolukki ja epäterveellinen porkkana – lapsen luovat ilmaiset kognitiivisen kielitieteen näkökulmasta (The child's creative expressions from the perspective of cognitive linguistics). MA thesis, Tampere University.
- Vänttilä, S. 1998. Suomenkielisten lasten uudisjohdokset ja Clarkin sananmuodostuksen periaatteet (The innovative derivations of Finnish-speaking children and the word-formation principles of Clark). In *Kielen ituja. Ajankohtaista lapsenkielen tutkimuksesta* (Actual topics about child language research) [Suomen ja saamen kielen ja logopedian laitoksen julkaisuja 10], M. Karjalainen (ed.), 99–127. Oulu: Oulun yliopisto.

Noun and verb derivations in early Turkish child and child-directed speech

F. Nihan Ketrez and Ayhan Aksu-Koç
İstanbul Bilgi University / Boğaziçi University

This chapter presents the emergence of derivational morphology in nouns and verbs in the speech of two Turkish-speaking monolingual girls between the ages 1;3 and 3;0, taking into account the patterns in their child-directed speech. Derivational morphology emerges early, before age 3;0, although compounding is an option in addition to derivation in the language. In both children's speech the causative is the first derivational morpheme and it emerges simultaneously with inflections followed by other types of derivations. Instrument nouns emerge earlier than agent nouns in both children's speech. The study further shows that children with different developmental paces follow similar paths in their acquisition of derivational morphology and their speech reflect the derivational patterns in their child-directed speech.

Keywords: Turkish, noun and verb derivation, compounding, child-directed speech, voice suffixes

1. Derivation in Turkish

Turkish is an agglutinating language and a great majority of derivations is realized through suffixation. Suffixes are categorized according to the lexical class of words that they are attached to (nominal versus verbal as two major categories). Göksel and Kerslake (2005) list approximately 150 distinct derivational suffixes that can be attached to noun, verb, adjective and adverb bases as well as numerals, pronouns, interjections and onomatopoeic words.¹ Some of these suffixes are attached to a

1. An even longer list can be found in Zülfişar (1991). See also Wilkens (2016) for a list of derivation suffixes with examples along with a review of studies on derivation in Modern Turkish and Aksan (2001) for an extensive discussion of agent derivations. Uzun et al. (1992) documents the derivation suffixes with their database frequencies. See Erdem (2011) for a discussion of inflection versus derivation in Turkish.

number of different types of bases. Derivational suffixes are distinct and salient morphemes. They usually bear the word stress, which falls on the final syllable in a typical Turkish word.

In noun derivations, the suffixes *-CI* (e.g. *kitap* ‘book’ *kitap-çı* ‘book seller’ or ‘bookstore’) and *-LIK* (e.g. *kral* ‘king’ *kral-lık* ‘kingdom’ or *göz* ‘eye’ *göz-lük* ‘glasses’) are listed as the most productive and frequent ones.² The diminutive suffix *-CIK* (e.g. *ayı* ‘bear’ *ayı-cık* (bear-DIM) ‘teddy bear’) and its various forms are also among the frequent derivational suffixes in the nominal domain.³ It is possible to derive nouns with adjective and verb bases as well (e.g. *temiz* ‘clean’ *temiz-lik* ‘cleaning’, *öğret-* ‘teach’ *öğret-men* ‘teacher’). In addition to these, some verbal inflections (e.g. the infinitive *-mA/-MAK* and the future marker *-(y)AcAK*) can have a derivation function, deriving nouns from verb bases (e.g. *dol-* ‘fill’ *dol-ma* ‘stuffed vegetable’, *ye-* ‘eat’ *ye-mek* ‘food/meal’, *iç-* ‘drink (V)’ *iç-ecek* ‘drink (N)’).⁴

In the verbal domain, voice suffixes, especially the causative (e.g. *düş-* ‘fall’ *düş-ür-* ‘drop’) and the passive (e.g. *el-le-* (hand-*LA*) ‘touch’ *el-le-n-* (hand-*LA*-PASS) ‘to be touched’), are the most frequent derivational morphemes. Reflexive (e.g. *giy-* ‘put on’ *giy-in-* ‘get dressed’) and reciprocal (e.g. *öp-* ‘kiss’ *öp-üş-* ‘kiss one another’), as well as the multiple applications of voice suffixes (double causatives, causative-passive combinations) are used productively to derive new verbs from verb bases. Although voice-alternating suffixes are listed under “inflection” sections in some reference grammars (e.g. Göksel & Kerslake 2005), it is not unconventional to study them under derivation. They are included in derivation counts in most studies on early Turkish lexicon (e.g. Sofu 1995). They are listed under derivations in traditional Turkish grammars and derivational suffix documentations (e.g. Banguoğlu 1990; Çotuksöken 1980; Ergin 1985; Korkmaz 2003; Sebüktekin 1974; Uzun et al. 1992, among others). Voice suffixes can be analysed as derivational

2. Suffixes alternate according to the rules of vowel harmony, consonant assimilation and other regular morphophonological processes. Those alternating vowels and consonants are represented by uppercase characters.

3. While it is possible to name and categorize some of the suffixes (e.g. the diminutive, the causative), in most cases, derivational suffixes have variable functions or interpretations and their glosses correspond to different English suffixes in individual examples. The suffix *-LIK*, for example, can derive an instrument noun (e.g. *göz* ‘eye’ *göz-lük* ‘glasses’) or a state (*özgür* ‘free’ *özgür-lük* ‘free-dom’ *çocuk* ‘child’ *çocuk-luk* ‘child-hood’, *aç* ‘hungry’ *aç-lık* ‘hunger’). That is why the suffixes are not always named or translated when they appear in isolation throughout the chapter.

4. *-mA(K)*, which is an infinitive marker as well as a derivation suffix in modern Turkish and other Turkic languages, is observed as a derivation marker in old Turkic scripts, evolving into an infinitive marker later (Ahmodava 2016). The other inflections listed here have similar historical developments (Korkmaz 2003).

because they result in a significant modification in the argument and event structure of the verb (e.g. the verb *öl-* 'die' becomes the verb *öl-dür-* 'kill' with the attachment of the causative suffix) resulting in a new lexeme, listed separately in dictionaries. Another property that supports their categorization as derivational is that verbs with voice morphology can be substituted, in some contexts, by mono-morphemic verbs that do not bear voice morphology (e.g. the causative verb *çık-ar-* 'take-out' can be replaced by the monomorphemic verb *al-* 'take'). Inflectional morphology (e.g. tense or case) is determined by the rules of syntax and cannot be substituted by forms that do not carry appropriate syntactic information. On the other hand, it is important to acknowledge that voice suffixes pattern like inflections due to their productivity and frequency and, as we discuss in this chapter, their hybrid nature affects the way they emerge, providing evidence that they are somewhere between inflectional and derivational ends of the continuum of word formation, as proposed in studies by Bybee (1985), Dressler (1989) and Stephany (1982).

Apart from the derivations with voice suffixes, verb derivations with noun (e.g. *el-* 'hand' *el-le-* 'to touch'), adjective (e.g. *iyi* 'good' *iyi-leş-* 'get better'), or onomatopoeic word bases (e.g. *hav-la-* 'bark') are also possible but relatively less frequent. Among such derivational suffixes, *-la* (e.g. *hazır* 'ready' *hazır-la-* 'to prepare') is the most frequent one, not only in terms of the lexemes it is attached to but also because it can be attached to different types of lexical classes (İleri 2007).

Derivation through prefixation is restricted to a number of borrowed prefixes and it is not productive. Some examples are Persian negative prefixes *na-* and *bi-* (e.g. *na-mert* 'coward', *bi-haber* 'uninformed') and English *a-*, *anti-*, (e.g. *a-tipik* 'atypical', *anti-biyotik* 'antibiotics'). The general understanding is that the prefixes are not borrowed as bound morphological forms, but those words that bear prefixes are borrowed as unanalysed words. This explains why these prefixes are not productive in the language.

Conversion, reduplication and compounding are alternative word-formation strategies in Turkish. Conversion is relatively rare. Some examples are *kuru* 'dry' and *kuru-* 'to dry', *aç* 'hungry' and *aç-* 'to open', *boya* 'color pencil, paint' and *boya-* 'to paint, color'. Reduplication of onomatopoeic words or interjections are observed in the derivations of nouns (e.g. *havhav*, corresponding to English *bowwow* 'dog' (*köpek* in adult speech) *vakvak* (quack-quack) 'duck' (*ördek* in adult speech)).⁵ Another form of reduplication is echo-word formation with /m/ in the word initial position

5. The sounds the animals make (*vak* and *hav*) are reduplicated and become the name of the animal (*vakvak* and *havhav*). In their derived noun forms, these words have the word stress on their final syllables. When they are produced as the sound the animals make, they are pronounced as two separate words, with separate word stress (*vak, vak* or *hav, hav*). So, these forms cannot be analysed as conversion.

before a vowel (e.g. *ördek mördek* ‘duck-mduck’ meaning duck etc. or *kitap-mitap* (book-mook) ‘book etc.’).

Compounding is observed as a productive, but a relatively less frequent word-formation process. Nominal compounds include noun-noun (e.g. *babaanne* (father+mother) ‘grandmother’), adjective-noun (e.g. *cicianne* (pretty+mother) ‘godmother’), verb-verb (e.g. *yapboz* (do+undo) ‘puzzle’) compounds as well as more productive and frequent possessive compounds (e.g. *portakal su-yu* (orange+water-POSS/CM) ‘orange juice’) which bear a third person singular possessive marker attached after the rightmost element of the compound. Noun-case-Noun kind of compounds (e.g. *kar-dan adam* (snow-ABL+man) ‘snow man’) are relatively rare. Turkish has synthetic compounds as well (e.g. *bilgisayar* (information+count-er) ‘computer’) but they are also relatively rare, especially in early CS (child speech) and CDS (child-directed speech). In the case of verbs, *et-*, *yap-*, *ol-* ‘make, do, become’ combine to form compound verbs with nouns and adjectives (e.g. *telefon et-* (telephone+make) ‘to phone’, *hasta ol-* (sick+become) ‘get sick’, *yemek yap-* (food+make) ‘cook’). Compound verbs can be formed with other types of verbs as well (e.g. *şarkı söyle-* (song+say) ‘sing’). The language also allows alternative means of word formation (compounding and derivation) for the same meaning. For example, ‘to get sick’ can be expressed with a compound verb *hasta ol-* (sick+become) ‘to get sick’ and also a derivation *hasta-lan-* (sick-lAn) ‘to get sick’.

This chapter focuses on morphological derivation and discusses other word-formation examples such as compounding only marginally and in relation to derivation.⁶ Among the morphological derivations, voice derivations (causative, passive, reflexive and reciprocal) in the verbal domain, and diminutives in the nominal domain are presented in the tables to display the complete picture of word formations, but they are not included in the discussions as they are handled in detail elsewhere (Ketrez 1999; Ketrez & Aksu-Koç 2007). Conversion examples will not be included in the derivation counts and the analyses because the direction of the conversion, and therefore whether a word is derived or not, is not always clear. Only those words that have transparent and distinct derivational suffixes are considered morphologically complex and were included as derived in the analyses in the present chapter.

6. See Ketrez (2017) for a discussion of the emergence of nominal compounds in early Turkish.

2. Acquisition of derivation

Previous work on the acquisition of derivation in Turkish shows that it emerges early, around 2;0, along with compounding and inflection (Aksu-Koç & Slobin 1985; Ekmekçi 1987; Ketrez & Aksu-Koç 2007; Ketrez 2017; Sofu 1995). Sofu (1995, 2016), based on data collected between 2;0–3;6, reports that at age 3;6 both compounds and derivations are observed in the children's speech but the relative proportions of derived nouns and verbs versus compounds indicate that derivation is the more dominant process. Turkish contrasts with so-called compounding languages in that sense and these findings are discussed by Sofu (1995, 2016) as a support for Clark's (1993) observation that children's language displays the word-formation properties of their languages at a very early age.

Dominance of derivation versus compounding is observed more clearly in bilingual children's speech. Verhoeven (1988), for example, shows that six- and eight-year-old bilingual children acquiring Dutch and Turkish prefer compounding in Dutch and derivation in Turkish as the most common word-formation strategy in their speech. In innovative word formations, however, compounding and derivation proportions are very close in six-year-olds' speech and compounds outnumber the derivations in eight-year-olds' speech, which is attributed to the overuse of *şey* 'thing' as in *çay şeyi* 'tea thing' or *şey motoru* 'thing motor'. Although the study does not discuss the reason, one may attribute the increase in compounding in the innovative words to the increasing exposure to a compounding language, or the attrition in the first language. On the other hand, preference for compounding for innovative words is not a property of bilingual language acquisition alone. Akyol-Bal and Sofu (2014) report a similar trend in the lexical innovations of monolingual Turkish children. Children prefer compounding in their creative lexical productions especially during the early phases of language development (ages 2;0–4;0) which is attributed to the transparency and relative simplicity of compounds when compared to derivations. This pattern is noted to be similar to the one observed in the acquisition of English reported by Clark (1993) where derivation is discussed as a property of a later developmental stage.

Among the derivational suffixes, the agentive marker *-CI* and *-LIK* are reported to be the earliest and the most frequent derivation markers, deriving nouns from noun bases, in both CS and CDS in Sofu (2016) and in a much earlier study by Cüceloğlu (1973). In bilingual data, as well, *-CI* and *-CIİK* (*-CI+LIK*), which is treated as a distinct derivation marker, are the most frequent ones along with *-LEŞ*, *-LI* (the attributive marker that expresses the resident of a place when attached to a place name, e.g. *Ankara-lı* 'someone from Ankara'), *-CA* (expresses the name of a language when attached to a nation name *Alman-ca* (German-*CA*) 'German language') and *-gİL* (expressing family relations, e.g. *amcam-gil* 'my uncle and his

family') in established words. *-CI* and *-CIIK* are reported in children's innovations, and over-marking of agency on nouns that are already agentive such as *öğretmen-ci* 'teacher-**CI*'. Bilingual children attach productive Turkish derivational suffixes to Dutch words and create examples such as *tikker-ci* 'touch-*CI*' or *tikken-le* (touch-*IA*) 'to touch' or *ijs-çi* (ice cream-*CI*), 'ice cream seller' as well, providing further evidence for productivity (van der Heijden 2011; van der Heijden & Verhoeven 1994; Verhoeven 1988). Such overgeneralization errors of derivational morphemes and creative examples of derivation come from children beyond age 3;0 or 4;0 and therefore they are properties of a late morphological period.⁷

In verb derivations, *-IA* (e.g. *el* 'hand' *el-le-* 'to touch') is the most frequent derivational suffix that derives verbs from nouns in both Sofu's (2016) and Verhoeven's (1988) study, while other derivational suffixes are recorded only marginally in a few instances and fail to meet any productivity criterion. Apart from *-IA*, the most frequent verbal derivations are realized with voice morphology (Sofu 1995). Ketrez (1999) reports that voice morphology emerges before 2;0, along with inflections, and evidence for productivity, the use of the underived base forms interchangeably with the forms with voice morphology, sometimes within the same session, is observed quite early.

Although this chapter is not the first study on the acquisition of derivational morphology in early Turkish, it is the first one that discusses the derivation patterns in CS in relation to CDS. Another unique contribution of the present study is that the two children, with different paces of development, are studied individually, without merging their data. This has enabled us to document that children follow the same patterns of morphological development although they proceed with different paces.

3. Method

Spontaneous interaction of two monolingual Turkish speaking girls, Deniz and Zeynep, were recorded with their family members, twice a month for 20–30-minute long sessions at their own home. Both girls come from families with middle socio-economic status, living in Istanbul and have university educated parents. Zeynep's mother works full time and her grandparents take care of her with the assistance of a live-in nanny. Her CDS included the grandparents, other family members and the nanny in addition to the mother. Deniz's speech was recorded

7. Similar innovative examples are reported by Ekmekçi (1990) and Kabadayı (2012) for monolingual children's speech.

by her stay-at-home mother on an audiotape. Her father and grandmother participated in some sessions.

Table 1 presents the number of sessions, children's age range, MLU (Mean Length of Utterance) range (in morphemes) and total number of utterances and word tokens in each corpus. The MLU range displays the highest and the lowest MLUs recorded during the period analysed, they are not necessarily the earliest and the latest values. Although Zeynep has a higher number of sessions, which results in a longer total duration and two times more CDS, the amount of speech produced by the children is similar both in terms of the number of utterances and word tokens.

Table 1. Participants, their age and MLU range, total utterances and word tokens

	No. of sessions	Age range	MLU range	Utterances	Word tokens
CS(D)	21 (~420 minutes)	1;3–2;0	1.20–3.79	3,831	7,026
CDS(D)			4.00–6.22	7,217	20,542
CS(Z)	33 (~660 minutes)	1;4–3;0	1.0–3.81	3,979	6,623
CDS(Z)			3.54–4.93	14,403	41,335

All the participants' utterances were transcribed according to the CHAT conventions of CHILDES (MacWhinney 2000) and coded morphologically. All the nouns and verbs were targeted for analysis and their stems were further coded as morphologically complex (derived) or simplex. Potential productivity of derivation is evaluated based on the use of individual derivational suffixes. Those suffixes that appear on three or more lemmas are considered potentially productive.

4. Results

4.1 Emergence of morphology in different paces

Figure 1 shows the MLU (in morphemes) of Deniz and Zeynep. Both children start inflecting words before 2;0, Deniz around 1;5–1;7, Zeynep around 1;8–1;9. Deniz displays a faster development and a sharper increase in MLU, followed by Zeynep with approximately six months difference. Deniz's protomorphology stage is observed to be between 1;6–1;9 (Aksu-Koç & Ketrez 2003; Ketrez 2017) while Zeynep's words become morphologically complex only after age 1;9. Zeynep's protomorphology stage is observed after 2;0 and the morphological complexity of her speech reaches its peak level around 2;11. This is the complexity level Deniz reaches around age 2;0. As will be shown below, Deniz and Zeynep have similar developmental patterns in their acquisition of derivational morphology, yet with different developmental paces.

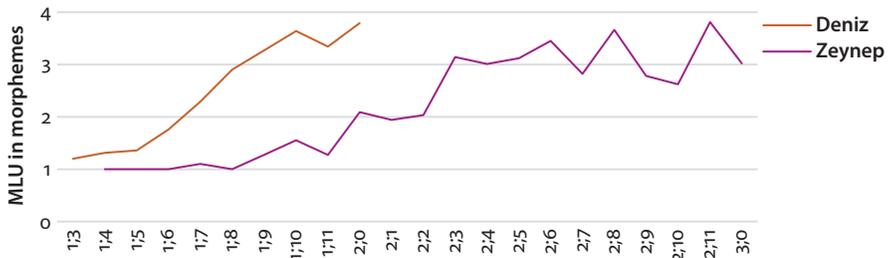


Figure 1. MLU in morphemes between 1;3–3;0

In their first sessions, the children, Deniz at 1;3 and Zeynep at 1;4, have more N than V lemmas in their speech. In terms of token frequencies Ns outnumber Vs during the first months of their development. This pattern continues until 1;8 in Deniz's speech and 2;3 in Zeynep's speech. After that, we see a V dominance in terms of token frequencies and this increase in V tokens results in an increase in the MLUs of both children, as Vs are morphologically more complex than Ns. A similar shift from N dominance to V dominance, which is observed at different ages but around the same MLUs in both children's speech, is not observed in lemma frequencies. In terms of lemmas, Ns are either more frequent, or they are equal to Vs throughout the period analysed. In CDS of both children we see a similar trend.

4.2 Derivation versus other word-formation options in CDS and CS

Table 2 shows the numbers and proportions of different types of word-formation options including compounding and reduplication, in addition to derivation through suffixation. As seen in this table, reduplication is observed in a few examples in only noun derivations.

Table 2. Types of word formation in Ns and Vs (lemmas/tokens)

	CDS(D)	CS(D)	CDS(Z)	CS(Z)
V total	333/5480	141/2367	419/12074	218/1693
Compound (V)	37/100	11/20	54/142	24/41
% Compound (V)	11.11/1.82	7.80/0.84	12.8/1.17	11.0/2.42
Derivation (V)	139/533	51/146	213/1019	47/118
% Derivation (V)	41.74/9.72	36.17/6.16	50.83/8.43	21.55/6.96
N total	749/5183	354/2099	848/8061	358/1526
Compound (N)	129/324	31/160	98/554	41/129
% Compound (N)	17.2/6.25	8.75/7.62	11.5/6.87	11.45/8.45
Reduplication (N)	7/22	2/16	5/21	7/23
% Reduplication (N)	.93/4.2	.56/.76	.47/0.25	1.95/1.50
Derivation (N)	87/436	38/109	112/518	40/114
% Derivation (N)	11.61/8.41	10.07/5.19	13/6.42	11.17/7.47

In both children's corpora, derivation is the most frequent word-formation strategy in verbs in both CDS and CS and in terms of both lemma and token frequencies. 42% of the verb lemmas (9.72% of verb tokens) in CDS and 36% of the verb lemmas (6.16% of verb tokens) in CS are formed through derivation in Deniz's corpora. In Zeynep's CDS, we see a similar picture: 51% of the verb lemmas (8.43% of verb tokens) are derivations. In her CS, 21% of lemmas (7% of verb tokens) are derived through suffixation. The proportions are lower in token frequencies, but they are still higher than the compounding frequencies. About 10% of the verb lemmas and 0.8–2% of verb tokens are compounds in CDS and CS.

In noun derivations, we have a different picture. When compared to the verbs, the nouns are less likely to have derivational suffixes. In Deniz's CDS about 12% of lemmas (vs. 42% in verbs) bear derivational suffixes. In Deniz's speech 10% of the noun lemmas are derived. A similar proportion pattern is observed in Zeynep's speech. 13% of lemmas (vs. 51% of lemmas in verbs) are derivations in her CDS. Zeynep's CS is similar to Deniz's CS with 11% of nouns bearing derivational morphology. Compounds outnumber the derivations in Deniz's CDS, but this is not the pattern observed in the CS where we see similar proportions of derivations and compounds. In Zeynep's CDS and CS, compounds and derivations are close to each other in terms of both lemmas and tokens.

4.3 Verb derivations: An overview

Among the verb derivations, those that are realized through voice suffixes (causative, passive, reflexive, reciprocal) are the majority (70–76% of all derivations in terms of lemmas and 62–70% of all derived tokens), as seen in Table 3. So we can attribute the big proportion of derivations (versus compounding) in verbs to voice-alternating suffixes.

Table 3. Causative, passive, reflexive and reciprocal verbs

	CDS(D)	CS(D)	CDS(Z)	CS(Z)
Causative <i>-Dir, -Ar/-Ir, -t</i>	41/174	17/20	82/454	14/31
Passive <i>-(I)n/-(I)l</i>	46/140	14/53	40/134	10/19
Reflexive <i>-(I)n</i>	13/38	6/32	26/117	9/24
Reciprocal <i>-(I)ş</i>	5/16	2/16	9/13	–
Total Voice	105/368	39/103	157/718	33/74
Total Derivation	139/533	51/146	213/1019	47/118
% Voice	75.5/69.0	76.47/70.54	73.70/70.46	70.21/62.71

In addition to the voice suffixes that derive verbs from V bases, verbs can be derived by means of suffixation of a variety of derivational suffixes to N, ADJ, Number and ONOM (onomatopoeia) bases (Table 4). In both corpora, 11 such distinct suffixes

are observed (in addition to the six voice suffixes). Among these suffixes, *-IA* has a special status because it can be attached to a variety of different lexemes and different lexical classes as bases (e.g. *bez-le-* (diaper-*IA*) ‘to diaper’, *temiz-le-* (clean-*IA*) ‘to clean’, *hav-la-* (bark-*IA*) ‘to bark’). As seen in the top rows of Table 4, in both Zeynep’s and Deniz’s CDS and CS, it is the most frequent derivational suffix. Of the verbs not derived by voice suffixes, 87% of verb tokens, and 62–65% of verb lemmas bear this suffix in CDS. In Deniz’s CS 83% of V lemmas and 91% of V tokens, in Zeynep’s CS about 57% of V lemmas and 66% of V tokens are derived with *-IA*.

Table 4. Verb derivational suffixes, their bases, examples and lemma/token frequencies in CS and CDS

Base-suffix	Examples	CDS(D)	CS(D)	CDS(Z)	CS(Z)
N- <i>IA</i>	<i>başla-</i> (head- <i>IA</i>) ‘start’	13/124	6/29	22/186	7/28
ADJ- <i>IA</i>	<i>temizle-</i> (clean- <i>IA</i>) ‘clean’	7/16	4/10	6/22	1/1
ONOM- <i>IA</i>	<i>havla-</i> (bark- <i>IA</i>) ‘bark’	2/4	–	7/53	–
N- <i>IA</i> n	<i>kirlen-</i> (dirt- <i>IA</i> n) ‘get dirty’	3/4	–	7/10	1/6
N- <i>IA</i> ş	<i>yerleş-</i> (place- <i>IA</i> ş) ‘be placed’	2/2	–	2/5	1/1
ADJ- <i>IA</i> ş	<i>iyileş-</i> (good- <i>IA</i> ş) ‘get better’	2/2	–	2/4	1/1
Num- <i>IA</i> ş	<i>birleş-</i> (one- <i>IA</i> ş) ‘merge’	–	–	2/6	–
N- <i>Al</i>	<i>yönel-</i> (direction- <i>Al</i>) ‘go towards’	–	–	1/1	–
ADJ- <i>Al</i>	<i>düzel-</i> (straight- <i>Al</i>) ‘get straight’	2/5	1/2	1/1	1/1
N- <i>A</i>	<i>kana-</i> (blood- <i>A</i>) ‘bleed’	–	–	1/1	–
ONOM- <i>DA</i>	<i>kıvrıda-</i> (wriggle- <i>DA</i>) ‘wriggle’	2/4	–	1/1	–
ADJ- <i>Ar</i>	<i>kı(rmı)zar-</i> (red- <i>Ar</i>) ‘get red’	–	–	1/2	1/1
ADJ- <i>Il</i>	<i>küç(ük)ül-</i> (small- <i>Il</i>) ‘get smaller’	–	–	1/1	–
V- <i>Ik</i>	<i>gör(z)ük-</i> (see- <i>Ik</i>) ‘appear’	–	–	1/4	–
ADJ- <i>Ik</i>	<i>acık-</i> (hungry- <i>Ik</i>) ‘get hungry’	1/4	1/2	–/–	1/3
N- <i>sA</i>	<i>susa-</i> (water- <i>sA</i>) ‘get thirsty’	–	–	1/3	1/2
V- <i>Im</i> sA	<i>gülümse-</i> (laugh- <i>Im</i> sA) ‘smile’	–	–	1/1	–
Total		34/165	12/43	56/301	14/44

The suffix *-IA* is followed by the suffixes *-IA*ş (e.g. *yer-leş-* (place-*IA*ş) ‘be placed’) and *-IA*n (e.g. *hasta-lan-* (sick-*IA*n) ‘get sick’) in CDS but these are not observed in Deniz’s speech, and appear in one lemma each in Zeynep’s speech. *-IA*, *-IA*ş and *-IA*n are the only suffixes that are attached to N bases in CDS (18 and 32 N base lemmas in total in Deniz’s and Zeynep’s CDS, respectively).

Six suffixes are attached to ADJ bases (two of them are *-IA*ş and *-IA* (*iyi-leş-* (good-*IA*ş) ‘get better’ and *hazır-la-* (ready-*IA*) ‘prepare’) which are attached to Ns as well and they are observed in 12 lemmas in Deniz’s CDS and 11 lemmas in Zeynep’s CDS. Two suffixes, *-IA* and *-DA*, are attached to the ONOM category (*hav-la-* (bark-la) ‘bark’ and *kıvrı-da-* ‘wriggle’). They are produced in four/six lemmas in CDS but they are not observed in CS.

Table 4 does not include the voice suffixes as their frequencies are presented in Table 3. Three different types of causative morphemes (*-DIr*, *-Ar/-Ir* and *-t*) are counted as three distinct derivations because their distribution is not phonologically conditioned. Passive, reciprocal and reflexive suffixes have phonologically conditioned allomorphs, therefore they are listed as single derivational suffixes. V bases (Table 3) outnumber the other bases (Table 4) in verb derivations in CDS and this is mostly due to the frequency of voice suffixes. It is also important to note that voice suffixes are the only suffixes that are attached to more than one V base to derive verbs in CDS and CS during the period analysed. So derivation gains productivity with the voice suffixes at the beginning. Other than these, there are two derivational suffixes (*-ImSA* and *-Ik*) that are attached to a V base to derive verbs (see Table 4 for examples) but they appear on one verb each, only in Zeynep's CDS.

4.4 Emergence of verb derivations

Table 5 shows the chronological order of emergence of derivation versus other word formations together with inflection in CS. In Deniz's recordings, we already have a causative verb at the very first session (*kapa-t-* 'turn-off') at 1;3, together with four other non-derived verbs. In the following sessions, she does not produce any derived verbs. From 1;7 onwards, the causative emerges in two additional lemmas (*benze-t-* 'make resemble' and *bit-ir-* 'finish') and from this age onwards, derivations increase in her speech.

The first derivation with another type of suffixation is *bez-le-* (*diaper-IA*) 'to diaper', derived from a N base, recorded at 1;8. Another example with the same suffix is observed at 1;9, *el-le-* (*hand-IA*) 'to touch'. Deniz attaches this suffix to ADJ bases as well in the following sessions (e.g. *temiz-le-* (*clean-IA*) 'clean' at 1;10). The second derivation with an ADJ base is recorded at 1;9: *düz-el-t-* (*düz-AL-CAUS*) 'make straight' and it has the base *düz* 'straight', it is derived with *-Al*, and then made causative.

Table 5. Age of emergence of derivation, compounding and inflection

	Derivation (Voice)	Compound	Derivation (other)	Inflection
CS(D)	1;3 (<i>kapa-t-</i> 'turn off') 1;7 (<i>bit-ir-</i> 'finish')	1;7 (<i>ütü+yap-</i> 'iron')	1;8 (<i>bez-le-</i> 'diaper') 1;9 (<i>el-le-</i> 'touch')	1;3 (<i>düş-PST</i> 'fell') 1;7 (the first 3-member paradigm)
CS(Z)	1;9 (<i>kapa-t-</i> 'turn off') 1;11 (<i>bit-ir-</i> 'finish')	2;1 (<i>kayıp+ol-</i> 'disappear')	2;4 (<i>baş-la-</i> 'begin') 2;8 (<i>ebe-le-</i> 'tag')	1;9 (<i>kapa-t-PST</i> 'closed') 1;11 (the first 3-member paradigm)

In Zeynep's speech we see a similar pattern. As observed in Deniz's recordings, the first derived verb is the causative verb *kapa-t-* 'turn off', recorded twice at 1;9. Starting at 1;11, we observe an increasing number of voice-alternating suffixes attached to various bases. The second causative verb, which is the second one in Deniz as well, is *bit-ir-* (finish-CAUS) 'finish'. At 1;11, she fails to produce the causative suffix in the verb *uy-dur-* (fit/match-CAUS) 'make up' that appears in a song. Similarly, she fails to produce the derivational suffix *-lAn* in the word *can-lan-* (heart/life-*lAn*) 'come to life' that again occurs in a song that she is singing. This verb is recorded with the derivational suffix only at 2;11. The first derivation (with one of the other types of suffixes) is recorded only at 2;4: *baş-la-* (head-*lA*) 'start'. This is a derivation from a N base with the *-lA* suffix, which is observed once more attached to another base at 2;8: *ebe-le-* (it-*lA*) 'tag'. Another suffix that appears in more than one lemma is *-lAş* that is observed in the word *iyi-leş* (good-*lAş*) 'to be/become well' and *yer-leş-* (place-*lAş*). These two suffixes, *-lA* and *-lAş*, are the only ones that can be considered potentially productive in addition to the voice-alternating suffixes in her speech.

In Zeynep's speech the first compound (*kayıp+ol-* (lost+be) 'get lost, disappear') is recorded in 2;1, this is after voice-alternating suffixes that emerge at 1;9, but before the other derivations, which are recorded at 2;4 for the first time. From that age onwards, compound verbs (e.g. *boya+yap-* (paint+make) 'color', *şaka+yap-* (joke+make) 'joke', *banyo+yap-* (bath+make) 'bathe') increase in number but they never exceed the number of derivations. In Deniz's speech, too, causative verbs are the first derivations, and compounds emerge only after, or simultaneously with the other types of derivational suffixes. The first compound verb *ütü+yap-* (iron+do) 'to iron' is recorded at 1;7 and the next one is *çiş+yap-* (pee+do) 'to pee' recorded at 1;8.

In both children's speech inflections emerge around the time the causative verbs emerge, 1;3 in Deniz, and 1;7 in Zeynep, which is before other types of derivations. The first 3-member inflectional "mini-paradigm" (Bittner, Dressler & Kilani-Schoch 2000; Kilani-Schoch & Dressler 2002) appears in 1;7 and 1;11 in Deniz and Zeynep's speech, respectively, and these ages coincide with the second causative verbs in their speech. Children go through the same emergence phases with the same suffixes (even lemmas, in the case of the causatives) but at different ages.

Complex derivations, that is, multiple suffixations are observed as rare instances in verb derivations. Deniz, for example, produces the passive verb *el-le-n-* (hand-*lA*-PASS) 'be touched' that is derived from the N base *el* 'hand' with *-lA* and then it is passivized with *-n* (Deniz 2;0). Another such derivation is observed in both Deniz's and Zeynep's speech: *düz* 'straight' is used to make a verb with *-Al*, *düz-el-* 'get straight' and then it is causativized: *düz-el-t-* 'make straight' (Deniz, 1;10, Zeynep 2;7).

In summary, in both children's speech, emergence of word formation starts with derivation in the verbal domain, along with inflections, and the first derivation examples are causative verbs *kapa-t-* 'close' and *bit-ir-* 'finish'. Compound verbs emerge later, but before the other derivational suffixes. In both children's speech, voice alternations outnumber the other types of derivations, and due to this pattern, verb bases are more frequent than other types of bases. Among the other types of suffixes *-la* appears as the most frequent one in terms of the variety of the bases it is attached to and the only one that meets the productivity criterion in CS (attached to five N and three ADJ bases in Deniz's CS and attached to seven N and one ADJ base in Zeynep's CS). In Zeynep's CS *-laş* is the other suffix that is attached to more than one base (one N and one ADJ). All the other derivational suffixes are observed only on one lemma each so their productivity cannot be assumed.

4.5 Noun derivations: An overview

As shown in Table 2, derived nouns constitute about 12% of the noun lemmas and less than 10% of noun tokens in CS and CDS. Of these derivational suffixes, diminutives that are formed by *-CIK/-CIğIm*, *-o*, *-(I)ş* or *-(I)şko* are the most frequent ones as a group, observed in 18% of the derived lemmas, and about 20–26% of the derived tokens in both CDS and Zeynep's CS (see Table 6). There are relatively fewer diminutives in Deniz's CS.

Table 6. Diminutive and derivation frequencies in CS ad CDS (lemmas/tokens)

	CDS(D)	CS(D)	CDS(Z)	CS(Z)
Diminutive	16/88	4/6	20/114	12/30
Derivation	87/436	38/109	112/518	40/114
% of Diminutives	18.39/20.18	10.52/5.50	17.85/22	19.67/26.31

As can be seen in Table 7, in Deniz's CDS, there are 10 distinct suffixes that are attached to N bases (44 lemmas) and two of these suffixes (*-CI* and *-IİK*) can be attached to ADJ bases as well. There are 15 suffixes that are attached to V bases (54 lemmas) in total, and two of these (*-IntI* and *-Ak*) can be attached to an ONOM base as well. In Deniz's speech a parallel distribution is observed. Seven suffixes are attached to N bases, one of which is attached to an ADJ base as well (in total 13 lemmas), and 10 suffixes are attached to V bases, one of which is observed on a ONOM base as well (in total 24 lemmas). Similarly, in Zeynep's CDS there are 10 suffixes that are attached to N or ADJ bases (62 lemmas), and 15 suffixes that are attached to V or ONOM bases (50 lemmas). In Zeynep's speech eight suffixes are attached to V bases (20 lemmas) and eight suffixes are attached to N bases (21 lemmas).

Table 7. Noun derivational suffixes, their bases, examples and lemma/token frequencies in CDS and CS

Base-suffix	Example	CDS(D)	CS(D)	CDS(Z)	CS(Z)
N-DIM:CIK	<i>kedicik</i> (cat-DIM) 'kitty'	11/82	3/5	13/91	5/13
N-DIM:Iş	<i>bebiş</i> (baby-DIM) 'little baby'	5/6	–	6/22	5/13
N-DIM:Işko	<i>babışko</i> (father-DIM) 'daddy'	–	–	1/1	1/1
N-DIM:o	<i>Zeyno</i> (Zeynep-DIM) 'Zeyno'	–	–	–	1/3
N-Ilk	<i>gözlük</i> (eye-Ilk) 'glasses'	8/22	2/3	10/51	3/6
ADJ-Ilk	<i>temizlik</i> (clean-Ilk) 'cleaning'	3/4	–	9/18	–
N-cl	<i>yolcu</i> (road-cl) 'passenger'	6/26	2/6	12/31	–
ADJ-cl	<i>eskici</i> (old-cl) 'ragman'	1/1	1/1	–	–
N-dA	<i>gözde</i> (eye-DA) 'favorite'	1/1	–	–	–
N-II	<i>köylü</i> (village-II) 'peasant'	2/2	1/1	1/1	–
N-DAş	<i>arkadaş</i> (back-DAş) 'friend'	2/17	1/8	3/29	1/1
N-ArI	<i>dışarı</i> (out-ArI) 'outside'	2/11	1/1	5/75	4/7
N-InCA	<i>ortanca</i> (middle-InCA) 'middle'	–	–	1/2	–
N-cA	<i>Almanca</i> (German-CA) 'German'	1/1	–	1/4	–
N-(y)A	<i>Aliye</i> (Ali-yA) 'Aliye'	2/5	1/1	–	1/1
V-(y)AcAK	<i>tutacak</i> (hold-(y)AcAK) 'handle'	4/18	3/11	2/53	2/13
V-Ak	<i>uçak</i> (fly-Ak) 'plane'	5/57	4/16	5/53	5/12
ONOM-Ak	<i>çingirak</i> (tinkle-Ak) 'rattle (object)'	1/4	1/6	–	–
V-(G)I	<i>sargı</i> (wrap-GI) 'bandage'	4/22	2/3	3/13	–
V-gIn	<i>yangın</i> (burn-gIn) 'fire'	–	–	1/1	1/2
V-mAK	<i>yemek</i> (eat-mAK) 'food'	2/33	2/15	2/45	3/16
V-mA(cA)	<i>dondurma</i> (freeze-mA) 'ice cream'	12/33	4/11	7/33	4/13
V-IcI	<i>bakıcı</i> (care-take-cl) 'caretaker'	–	–	3/4	–
V-(A)mAç	<i>saklanbaç</i> (hide-(A)mAç) 'hide-and-peek'	–	–	–	1/4
V-KAn	<i>yapışkan</i> (stick-KAn) 'sticker'	1/1	–	–	–
V-(I)m	<i>doğum</i> (be born-Im) 'birth'	3/18	2/2	8/12	–
V-Inç	<i>bilinç</i> (know-Inç) 'consciousness'	–	–	1/1	–
V-In	<i>yayın</i> (spread-In) 'publication'	1/1	–	1/4	–
V-I	<i>örtü</i> (cover-I) 'cloth'	7/40	3/11	5/59	2/7
V-IntI	<i>kaşıntı</i> (itch-IntI) 'itching'	1/1	–	4/4	–
ONOM-IntI	<i>gürültü</i> (gurgling sound-IntI) 'noise'	1/1	–	–	–
V-(I)k	<i>delik</i> (drill-(I)k) 'hole'	7/22	2/4	5/11	–
V-mIş	<i>dolmuş</i> (fill-mIş) 'shared taxi'	–	–	1/3	–
V-Iş	<i>bakış</i> (look-Iş) 'look'	2/2	1/1	–	–
V-A	<i>yara</i> (split-A) 'scar'	1/2	1/2	–	–
V-mAz	<i>yaramaz</i> (do good-mAz) 'naughty'	1/1	–	–	–
V-mAn	<i>öğretmen</i> (teach-mAn) 'teacher'	1/2	–	2/11	1/2
Total		87/436	38/109	112/518	40/114

As presented above in Section 4.3, in the derivation of verbs, most of the derivational suffixes were attached to N or ADJ bases. In the N derivations, there is a more even distribution of N and V bases in both CS and CDS. Another difference in N and V derivations is the variety of the suffixes. In N derivations there are 33 different types of suffixes in both corpora in total. In V derivations, we observed 12 suffixes (and 6 of these were voice suffixes).

Among the suffixes, *-IİK*, which is reported to be one of the most productive suffixes in Turkish in Göksel and Kerslake (2005), is the most frequent one in N derivations with N bases (after the diminutive suffixes as a group). It can be attached to ADJ bases as well and this results in an even more frequent use. *-CI* is also observed in a number of different noun lemmas in Deniz's CDS. *-Ak* is the most common N derivational suffix that can be attached to V bases and it is observed in both children's CDS. The suffix *-mA* is observed in a number of different lemmas in Deniz's CDS, including compounds such as *oku-ma saat-i* (read-*mA* hour-POSS/CM) 'reading hour' but it is a relatively less frequent suffix in Zeynep's CDS, recorded in five lemmas. In addition to its function as a derivation marker it is one of the infinitive forms (Bayraktar-Erten 2017) so it is a frequent form in the language in general.⁸ This is important because, as will be discussed below in Section 4.7, this is the suffix that is used in children's creative examples.

Complex derivations with multiple suffixes are observed as rare instances. In one such example, *don-dur-ma* 'ice-cream', the intransitive verb *don-* 'freeze' takes the causative marker *-Dir* and becomes a transitive verb *dondur-* 'freeze' and then it takes the *-mA* derivational suffix to become a noun 'ice cream'. Another example is *tren-ci-lik* 'pretend train game'. The noun agentive suffix *-CI* is attached to the word *tren* 'train' and then *-IİK* is attached to name the game: *tren-ci-lik* (N) 'a pretend train game' *-IİK* or *-CIİİK* is very productive in CDS and CS and is one of the early complex derivations as pretend games are very typical in the child communication context. In the literature (e.g. Verhoeven 1988) *-CIİİK* is analysed as one single suffix, most probably because the intermediary derivation with *-CI* only is not always a legitimate word in the language (e.g. the derivation *doktor-cu* in *doktor-cu-luk* is not a possible word). Derived words can appear in compounds as well and pose another level of complexity. One such example is *su-la-ma kova-sı* (water-*lA*-*mA*+can-POSS/CM) 'watering can'.

8. The suffix that appears in deverbal nouns such as *don-dur-ma* 'ice cream' is clearly a derivational suffix as it derives a distinct lexeme referring to the substance ice cream and the causative verb *dondur-* 'freeze' no longer bears verbal features. It cannot have an object for example. In such examples the word no longer refers to the freezing event or the freezing action, although it is possible to have such a reading of the word *don-dur-ma* in appropriate contexts (e.g. *Ali-nin meyve-ler-i dondur-ma-sı-nu iste-di-m* 'I wanted Ali to freeze the fruit'). In the present study, we have not included the *-mA*-marked verbs in the subordinate clauses as they retain their verbal features.

4.6 Emergence of noun derivations

In Deniz's speech, the first attempted derivation is observed at 1;4. She produces *cice*, a reduced form of the word *eski-ci* (old-*CI*) 'ragman', maintaining the derivation morpheme. Because she produces the *-CI* part, we considered this the first production of a derivational morpheme. In the same session, she also produces the word *del-ik* 'hole' that is derived with the attachment of *-Ik* to the verb base *del-* 'pierce/make holes'.⁹ The next derivation is the proper noun *Ali-ye* (Ali-*yA*) that has the suffix *-yA* that derives female names from male names. Starting at 1;7, Deniz's derivations increase in number. In this session she has the word *arka-daş* 'friend', derived by the attachment of *-dAş* to the noun base *arka* 'back', and *ört-ü* 'cloth', derived with *-(g)I* based on the *V ört-* 'cover'. The other derivations that follow these in the following sessions are all V-based derivations: *kapı* (*kapa* 'close'-*I*) 'door', *ye-mek* (*ye-* 'eat'-*mAk*) 'food/meal', *yat-ak* (*yat-* 'lie'-*Ak*) 'bed', for example. There are other N-based derivations as well, but they emerge later and are fewer in number. Diminutive derivation is observed only in two words recorded at 1;9 (*ayı-cik* (bear-DIM) 'teddy') and 2;0 (*Kaan-cığım* (Kaan-DIM) 'dear Kaan').

There are more verb bases and more suffixes that are attached to verbs in Deniz's derivations. Among N suffixes, *-CI*, *-CIK* and *-IİK* are the only ones that are observed in multiple bases, and *-CI* and *-CIK* are the only suffixes that are attached to three lemmas and thus meet the productivity criterion. In the case of the V suffixes, *-Ak* ($n = 5$), *-I* ($n = 3$), *-mA(cA)* ($n = 4$), and *-(y)AcAk* ($n = 3$) are attached to three or more lemmas. So, in Deniz's CS, derivations with V bases and V suffixes are not only more frequent, they are also more productive. This is similar to the pattern observed in her CDS.

In Zeynep's speech, the first derivation is a proper noun with a diminutive marker, *Ayşe-cik* (Ayşe-DIM) 'dear/little Ayşe', that appears in a song, and the second one, recorded in the same session, is a failed attempt, *çift-lik* 'farm', which is derived with the attachment of the suffix *-IİK* to the word *çift* 'pair'.¹⁰ Zeynep produces only the first syllable of the word, without the derivational suffix. This sounds more like a phonological omission rather than an omission of the derivational morpheme as the syllable. If we do not count these two early occurrences

9. She actually points to her slippers (*ter-lik*) and produces the word *de:lik* and the mother interprets this word as *del-ik* 'hole' because there is a hole in her slippers. Her intended utterance could also be *terlik*, another derived word.

10. Here the word *çift* 'pair' refers to a pair of oxen that are used in the farm. Although it is a historically derived word, its derivation status is disputable as far as the present day Turkish, especially the child/child-directed speech is concerned. On the other hand, the word *çift* meaning 'pair' or 'couple' is a common word in the language.

in the corpora, we can say that the actual emergence of derivational morphology in nouns is only at 2;2 with the word *oyuncak* ‘toy’ derived from the V base ‘oyna’ ‘game/play’ with the attachment of the suffix $-(y)AcAK$ that derives instruments.¹¹ This word is produced five times in the same session, and Zeynep continues using it in the following sessions as well. The second derived form is another instrument produced at 2;4, *uç-ak* ‘plane’, and is derived from the verb *uç-* ‘fly’ with the attachment of the suffix $-AK$. The following derived forms produced in the same session are *salın-cak* ‘swing’, (*sallan-/salın-* ‘to swing’- $(y)AcAK$), *ekmek* ‘bread’ (*ek-* ‘plant’- mAK) and *don-dur-ma* (freeze-CAUS- mA) ‘ice cream’. Each of these derivations are realized by different derivational suffixes but they all have V bases. The earliest derivations that we have not counted (*çift-lik* and *Ayşe-cik*) have N bases but their productivity is questionable (in the case of *Ayşe-cik*, because it appears in a song) or they reemerge only after 2;5 (in the case of *çift-lik*). Apart from them, the other N-based derivation is *göz-lük* ‘glasses’ derived by the same suffix, $-IİK$, we have in the word *çift-lik* ‘farm’.

As seen in Table 7, in Zeynep’s CS, there are eight distinct suffixes that are attached to 19 verb bases (69 tokens). Those that derive Ns from Ns are eight distinct suffixes (three of them are diminutives) and they are attached to 21 distinct noun lexemes (45 tokens). Therefore, in Zeynep’s speech, V bases and suffixes that derive Ns from V bases are not more frequent, but in terms of token frequencies, V bases outnumber the N bases.

In terms of productivity, V-N derivations and N-N derivations look similar. There are three suffixes that meet the productivity criterion with three or more lemma examples: $-Ak$ ($n = 5$), $-mAK$ ($n = 3$), $-mA(cA)$ ($n = 4$). The first two of these are observed as productive suffixes in Deniz’s speech as well. In N-N derivations, there are four suffixes, the diminutives $-CIK$ ($n = 5$), and $-Iş$ ($n = 5$), as well as $-IİK$ ($n = 3$) and $-ArI$ ($n = 4$). Only $-CIK$ meets the productivity criterion in Deniz’s speech. In Zeynep’s speech, V-N derivations do not outnumber the N-V derivations (unlike Deniz’s speech), but the types of suffixes that meet the productivity criterion (with three or more lemmas) or the ones that are potentially productive (occurring on two lemmas) mostly overlap (Table 7).

The first examples of compounds are recorded at 1;7 in Zeynep’s speech but then they are not observed again until 1;10. After 1;10 they reemerge and increase in number until 2;9 but their proportion does not exceed 25–27% of all noun lemmas throughout the period analysed, ending up with a mean compound proportion of about 11% of all nouns (lemmas). This is the same mean proportion of compounds

11. The word root of the derived word *oyuncak* is analysed as the V *oyna-* rather than the noun *oyun* because $-(y)AcAK$ is a suffix that can be attached to V bases. So the word is analysed as *oyna-yacak*.

in CDS. In Zeynep's speech the earliest complex forms are two different compounds for 'grandmother' (*baba+anne*, *anne+anne* 'grandmother'). She has a failed attempt to produce another compound *ayak+kab-ı* (shoe+container-POSS/CM) 'shoe'. She produces the word as /pakka/ in a reduced form, so we have not included it in our counts. The two grandmother compounds are bare NN compounds. The first possessive compound attempt is with *uğur+böceğ-i* (good luck+bug-POSS/CM) 'ladybug' but it is a failed attempt at 2;1 as she produces only the first part of the compound. The first possessive compound in full form, *deniz+kız-ı* (sea+girl-POSS/CM) 'mermaid', is recorded only at 2;4. This is around the time she produces the first derivations. In the meanwhile, she produces other NN compounds such as, *Ali baba* 'father Ali' *bakkal+amca* 'shopkeeper uncle' some of which appear only in songs or riddles she recites. Of 36 compound lemmas produced by Zeynep, 16 are possessive compounds (44%), only one of them is a synthetic compound (*bilgi+say-ar* (information-count-Ar) 'computer') and the rest are bare NN (*baba+anne* 'grandmother'), ADJN (*cici+anne* 'god mother') or VV (*yap+boz* (do-undo) 'puzzle') compounds. She omits the compound marker in one example and produces a NN compound instead of a possessive compound. In another instance, she produces *tansiyon+ölç* 'bloodpressure+measure' (N+V) to refer to a blood pressure gauge (*tansiyon+alet-i*, NN-POSS/CM compound or a synthetic compound *tansiyon+ölç-er*). This example is important to note because she uses a V instead of a N in the compound, and this is consistent with her tendency to use Vs as Ns. She uses the V *iç* 'drink' to refer to 'water' in a glass, as well. Such N-V conversions are not grammatical in the language. In Deniz's speech, the first compounds are *ay+dede* (moon+grandpa) 'moon', recorded at 1;3 and *baba+anne* 'grandmother', recorded at 1;5. The first possessive compound *ayak+kab-ı* 'shoe' pronounced as /a:baɣɪ/, with a metathesis, is produced at 1;5 but it is not considered the first example of a NN-POSS/CM compound due to its reduced pronunciation (Ketrez 2017). The second NN-POSS/CM compound, *Kurabiye+Canavar-ı* 'Cookie Monster' is recorded at 1;8 followed by *portakal+su-yu* (orange water-POSS/CM) 'orange juice' at 1;9. This is later than the derivations that emerge around 1;3–1;4, but around the time she produces her first diminutives.

Derivation by suffixation is recorded six months after the production of the first compounds in Zeynep's CS. These first derivations all appear in songs and it is likely that they are not analysed forms. It takes even more, about three more months, for her to produce the first real derivation (*oyun-cak* 'toy'). Although the number of derivations increase along with different types of suffixes, a similar increase is observed in compounds as well and derivations and compounds have similar proportions in her speech until age 3;0, 11% of lemmas and 8% of tokens. We observe a similar trend in Deniz's speech in lemmas 9% (compounds) versus 10% (derivations). In terms of token frequencies, however, she has slightly more

compounds (8%) than derivations (5%) (Table 2). All these show that derivation is not a (more) dominant word-formation strategy in children's speech. This result is confirmed by the ages of emergence of compounds and derivations, which are shown in Table 8 along with their first examples.

Table 8. Age of emergence of compounds, derivations and inflection

	Compound	Derivation (other)	Derivation (diminutive)	Inflection
CS(D)	1;3 (<i>ay+dede</i> 'moon+grandpa')	1;4 (<i>eski-ci</i> 'ragman', <i>del-ik</i> 'hole')	1;9 (<i>ayı-cık</i> 'teddy <i>bear</i>)	1;3 <i>abi-ye</i> 'brother-DAT'
CS(Z)	1;7 (<i>anne+anne, baba+</i> <i>anne</i> 'grandmother')	2;2 (<i>oyun-cak</i> 'toy')	2;4 (<i>Ayşe-cik</i> 'dear/ <i>little Ayşe</i>)	1;7 (<i>o-nu</i> 'it-ACC' <i>ora-da</i> 'there-LOC')

In both children's speech compounds emerge earlier than derivations. Diminutive is not the earliest derivation in Deniz's speech. In Zeynep's speech the first diminutives that are recorded at 2;1 appear only in songs, so their actual emergence date is recorded as 2;4, which is after the emergence of the word *oyun-cak* 'toy' at age 2;2. Derivation emerges after the first nominal inflections (case markers) in both children's speech.

4.7 Neologisms

There are two neologisms in Deniz's speech recorded at 1;11 and 2;0, and they are both V-based, produced with the suffix *-mA*. In one example, she says *goyma* (put-*mA*) referring to either a 'cup' or a 'pot' while she is serving milk to her mother during a pretend play. Another derivation with *-mA* is *kesme* (cut-*mA*), referring to 'scissors' (*makas* in the adult language). These examples show that at least at age 1;11–2;0, Deniz is actively creating words with a derivational suffix that is productive in the language.

Zeynep has two suffixation neologisms in her speech, too. She refers to the rolling pin (*merdane* in adult speech) with her own creation *aç-mak* (open-*mAK*). *-mAK* is another infinitive marker just like *-mA*, and it is one of the productive derivational suffixes that we observe in CS and CDS in words such as *ye-mek* 'food' (eat-*mAk*) and *ek-mek* 'bread' (plant-*mAk*). She also refers to a 'fire fighter' (*itfaiye-ci* in Turkish, N-*CI* derivation), with the word *kurtarma* (rescue (V)-*mA*). These examples are similar to the ones observed in Deniz's speech.

Other neologism examples are observed in the form of bare Vs used as a N, or reduplication of bare Vs, or the use of a V in a N position in Zeynep's speech. These are not legitimate neologism examples because such N-V conversions are not grammatical in Turkish. She produces the V *iç-* 'drink' to refer to *su* 'water'. She produces

VV *çık+çık* ‘come-out+come-out’ at 2;10 referring to the toy that comes out of Kinder Surprise chocolate. As stated above, she also had one synthetic compound *tansiyon+ölç* (blood pressure+measure) referring to a blood pressure gauge using a bare V ungrammatically in a NN-POSS/CM compound (*tansiyon+alet-i* ‘blood pressure+gauge’). Another instance of a bare V used in a N position is observed in the compound verb *kok+yap-* (smell (V)+make) ‘pass gas’. A grammatical production could be the noun form *kok-u* ‘smell (N)’ with the verb *yap-*. All these neologism examples are ungrammatical conversions because they involve verb bases, used in a bare form as nouns, which is not observed in the adult grammar.¹²

Overproduction of the other derivational suffixes are not observed, although there are a lot of such examples in the literature (Ekmekçi 1987, 1990; Kabadayı 2012). This is most probably because such derivations are characteristics of a later developmental period (after 3;0).

4.8 Agent vs. instrument nouns

Table 9 shows the emergence age of derived agent and instrument nouns, and the total number of agent and instrument nouns in CS and CDS of each child. As seen in the table, instrument nouns emerge earlier than agent nouns in both children’s speech. In Deniz’s CDS instrument nouns outnumber agent nouns, while in Zeynep’s CDS there are more agents, but the numbers are also closer. In Deniz’s speech, out of 34 derived lemmas, 21 lemmas refer to instruments or objects (e.g. *oyun-cak* ‘toy’, *tara-k* ‘comb’), six of them refer to agents (e.g. *arka-daş* ‘friend’, *balon-cu* ‘balloonman’). The others cannot be categorized as instrument or agent (e.g. *göster-i* ‘show’). In Zeynep’s speech out of 40 derivations 18 refer to objects (e.g. *göz-lük* ‘glasses’), 15 refer to agents (e.g. *öğret-men* ‘teacher’) and the rest refer to other types (e.g. *yan-gın* ‘fire’).

Table 9. Agent vs. Instrument nouns

	D	Z
Agent Ns	1;7 (<i>arka-daş</i> ‘friend’)	2;4 (<i>Ayşe-cik</i> ‘dear/little Ayşe’)
Agent N lemmas CS/CDS	7/27	15/39
Instrument Ns	1;4 (<i>ört-ü</i> ‘cloth’)	2;2 (<i>oyun-cak</i> ‘toy’)
Instrument N lemmas CS/CDS	21/46	18/31
Other	6/18	7/40

12. Nominal compounds such as *gel+git* (come-go) ‘ebb and flow’ or *çek+yat* (pull-lie) ‘sofa bed’ have bare forms of the verbs in nominal compounds. Such bare verbs do not form verbal compounds with *yap-*.

As presented in Table 10, in both CDS and CS, most suffixes that are attached to N bases derive agent nouns. Among these, the most frequent ones are the diminutive suffixes and the agentive *-CI*. (e.g. *ayı-cık* ‘teddy’, *balon-cu* ‘balloon man’). The other agent derivations with other suffixes (e.g. *arka-daş* ‘friend’, *can-baz* ‘acrobat’, *göz-de* ‘favorite’) are recorded once or twice. The only suffix that derives only instruments (along with other types of nouns such as *sağ-lık* ‘health’) is *-lik* (e.g. *göz-lük* ‘glasses’). In contrast, the derivations that have V bases are mostly instruments and in terms of token frequencies and lemmas they outnumber the ones derived with N bases. The only agent derivations with V bases are the words *öğret-men* ‘teacher’ and *yara-maz* ‘naughty’, which is actually an ADJ used as a noun.

Table 10. Agent vs. Instrument nouns with N/ADJ and V/ONOM bases (CS/CDS)

	Agent (D)	Instr. (D)	Other (D)	Agent (Z)	Instr. (Z)	Other (Z)
N-DIM: <i>CIK</i> , -DIM: <i>Iş</i> , -DIM: <i>o</i>	2/12	-/2	1/1	11/18	-/1	1/1
ADJ/N- <i>CI</i> , - <i>dA</i> , - <i>bAz</i> , - <i>dAş</i> , - <i>yA</i>	5/12	-	-	2/15	-	-
N- <i>II</i>	-/1	-	-	-/1	-	-
ADJ/N- <i>LIK</i>	-	1/4	1/3	-	3/5	-/13
N- <i>cA</i>	-	-/1	-	-	-/2	-
N- <i>ArU</i>	-	-	2/1	-	-	4/5
Total N/ADJ bases	7/25	1/7	3/6	13/34	3/8	5/19
ONOM/V- <i>(y)AcAK</i> , - <i>Ak</i> , - <i>(G)I</i> , - <i>mAk</i> , - <i>kAn</i> , - <i>In</i> , - <i>I</i> , - <i>(I)k</i> , - <i>A</i> , - <i>kAn</i>	-	17/35	-	-	14/20	-/2
ONOM/V- <i>mA(cA)</i>	-	3/4	1/8	1*/-	2/2	1/5
V- <i>IcI</i>	-	-	-	-/3	-	-
V- <i>Im</i> , - <i>IntI</i> , - <i>Iş</i> , - <i>gIn</i> , - <i>In</i> , - <i>Inç</i>	-	-	2/6	-	-	1/14
V- <i>mAz</i> , - <i>mAn</i>	-/2	-	-	1/2	-	-
Total V/ONOM bases	-/2	20/39	3/12	2/5	16/22	2/21
Total	7/27	21/46	6/18	15/39	19/30	7/40

Different sets of suffixes derive agents and instruments in the language (with a few exceptions) and the children seem to be sensitive to this distinction. There is one error produced by Zeynep that violates the adult pattern. She produces the word *kurtar-ma* ‘rescue-*mA*’ to refer to a fire fighter, an agent. This is a violation because *-mA* derives either instruments/objects (e.g. *dondur-ma* ‘ice cream’ or *uçurt-ma* ‘kite’) or other type of nouns such as action nouns (e.g. *konuş-ma* ‘speech’ or *sula-ma* ‘watering’).

5. Discussion and conclusion

In this chapter, we presented the emergence of derivational morphology in the speech of two Turkish speaking monolingual girls between the ages 1;3 and 3;0. We have shown that although they have different paces of development, they display similar patterns in their acquisition of derivational morphology.

In both children's CDS, less than 15% of noun lemmas and an even smaller proportion of tokens are derived forms. Around 30 different types of suffixes are used to derive Ns from mainly N, V and ADJ bases. Instrument nouns emerge earlier than agent nouns in both children's speech, and in terms of their proportions, they outnumber agent nouns in one child's speech, while in the other one, a more even distribution is observed. Verb derivations are realized with a smaller number and variety of derivational suffixes, 12 in total, 6 of which are voice suffixes. Despite that, verbs are richer in terms of derivational morphology, just as they are richer in terms of inflectional morphology, as 40–50% of verb lemmas bear at least one derivational morpheme. About 70% of these derivations are realized with voice suffixes. In the CDS of both children, verbs are derived more often from N bases than V bases if the variety of suffixes are considered. In terms of token frequencies, due to the frequency of causative and passive verbs, V bases outnumber the N bases.

In this study, following the earlier literature on language acquisition and traditional grammars we cited, we analysed voice suffixes as derivations, but we also acknowledged that they have a hybrid behavior that should not be overlooked. We observed that voice morphology emerged along with inflectional morphology, before other types of derivation. The timing of emergence cannot be an evidence that they are in the inflection category though, because the simultaneous emergence pattern could be attributed to the properties that they share with inflectional morphology (transparency, regularity, frequency etc.), not necessarily to the fact that they are inflections. It is still important to point out that if they were excluded in the derivation counts, nominal and verbal domains would have a much similar pattern in terms of the proportions of derivations versus compounding. Moreover, compounding would be an earlier acquisition than derivations in both children's speech and in both verbal and nominal domains.

Early emergence of NN compounds seems to be the pattern observed cross-linguistically in those languages where compounding is a word-formation option alternative to derivation (Clark 1993; Dressler, Ketrez & Kilani-Schoch 2017, among others). Compounds are preferred due to their semantic transparency and ease of formation especially during the earlier phases of language development and especially in neologisms. In our study, however, it is difficult to attribute the preference for compounding to a developmental phase because the proportions of

compounds versus derivations in CS and CDS are very close to each other. In general, children's speech reflects the derivational morphology patterns in their CDS. Further evidence for this parallel pattern is observed in the types and frequency of derivational morphemes. Those suffixes that are listed as the most frequent and the most productive suffixes in the target adult language are observed as such in the CDS and they are the first ones to emerge and meet the productivity criterion in CS. Turkish derivational morphemes are transparent and productive which most probably contribute to an early sensitivity facilitating an early emergence of such morphemes in CS.

We have shown that derivational morphology does not only emerge early, children also have a number of neologisms providing evidence for productive use of derivation as a word-formation strategy before age 3;0. Clark (1993) suggests that children acquiring derivational languages wait until a much later age to create novel words unlike children acquiring compounding languages where first innovative word formations in the form of NN compounds emerge quite early. What we observe in the acquisition of Turkish is that children can be creative with derivation at an early age even when compounding is an option in addition to derivation. In this study, we documented that children can have creative derivation examples before age 2;0. It is likely that the transparency, salience and regularity of derivational morphology contribute to the early emergence of derivational morphology in Turkish.

The results we presented in this chapter are similar to the results reported in earlier studies in the sense that both derivations and compounding emerge early, before three years of age. Derivation appears as the dominant word-formation strategy due to the frequent use of the voice morphology on verbs. The present study crucially shows that patterns observed in the CS mimic the patterns observed in the CDS and children with different developmental paces follow similar paths in their acquisition of derivational morphology.

Funding

The data analysed in this study were collected with the supports of Boğaziçi University Research Fund (Project numbers: 94 SO 064 and 96 S 0017, P.I. Ayhan Aksu-Koç) and TÜBİTAK (Turkish Scientific and Technical Research Council, Grant number: 111K270, P.I.: F. Nihan Ketz Sözmen).

References

- Ahmodova, A. 2016. From infinitives to deverbal nominals in the language of Orkhon inscriptions. *Journal of Humanities and Social Science* 21(8): 17–21.
<https://doi.org/10.9790/0837-2108011721>
- Aksan, M. 2001. *Agents in Turkish and Their Morphology* (Dilbilim İncelemeleri Dizisi) [Linguistics Research Series 3]. Mersin: Mersin University Press.
- Aksu-Koç, A. & Ketrez, F. N. 2003. Early verbal morphology in Turkish: Emergence of inflections. In *Mini-paradigms and the Emergence of Verb Morphology*, D. Bittner, W. U. Dressler & M. Kilani-Schoch (eds), 27–52. Berlin: Mouton de Gruyter.
- Aksu-Koç, A. & Slobin, D. I. 1985. The acquisition of Turkish. In *The Crosslinguistic Study of Language Acquisition, Vol. 1: The Data*, D. I. Slobin (ed.), 839–878. Hillsdale NJ: Lawrence Erlbaum Associates.
- Akyol-Bal, Ö. & Sofu, H. 2014. Türkçede yeni sözcük oluşturma (Lexical innovations in Turkish). *Turkish Studies* 9(3): 79–89. <https://doi.org/10.7827/TurkishStudies.6429>
- Banguoğlu, T. 1990. *Türkçenin Grameri* (Grammar of Turkish). Ankara: TDK Yay.
- Bittner, D., Dressler, W. U. & Kilani-Schoch, M. 2000. *First Verbs: On the way to Mini-paradigms* [ZAS Papers in Linguistics 18]. Berlin: ZAS.
- Bybee, J. L. 1985. *Morphology: A study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.9>
- Clark, E. V. 1993. *The Lexicon in Acquisition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511554377>
- Çotuksöken, Y. 1980. *Türkçe'de Ekler-Kökler-Gövdeler* (Affixes-roots-stems in Turkish). Istanbul: Alaz Yayınları.
- Cüceloğlu, D. 1973. Türkçe türetme ekleri üzerine bir çalışma (A study on the Turkish derivation affixes). *Hacettepe Üniversitesi Sosyal ve Beşeri Bilimler Dergisi* (Hacettepe University Journal of Social Sciences and Humanities) 5(1): 47–56.
- Dressler, W. U. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42: 3–10.
<https://doi.org/10.1515/stuf-1989-0102>
- Dressler, W. U., Ketrez, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins.
<https://doi.org/10.1075/lald.61>
- Ekmekçi, Ö. 1987. Creativity in the language acquisition process. In *Studies on Modern Turkish*, H. E. Boeschoten & L. T. Verhoeven (eds), 203–210. Tilburg: Tilburg University.
- Ekmekçi, Ö. 1990. Coinage of words in child's lexicon in acquiring Turkish. In *Proceedings of the Fourth International Conference on Turkish Linguistics*, S. Koç (ed.), 27–39. Ankara: Middle East Technical University Publications.
- Erdem, M. 2011. Türkçede yapım ve çekim eklerinin özellikleri ve sınırları (On the properties and boundaries of derivational and inflectional suffixes in Turkish). *Bilig* 58: 71–90.
- Ergin, M. 1985. *Türk Dil Bilgisi* (Turkish grammar). Istanbul: Boğaziçi Yayınları.
- Göksel, A. & Kerslake, C. 2005. *Turkish: A Comprehensive Grammar*. New York NY: Routledge.
- İleri, E. 2007. Überlegungen zur Untersuchung der Verben auf +IA+mAk. (Reflections on the examination of +IA+mAk verbs). In *Einheit und Vielfalt in der türkischen Welt: Materialien der 5. Deutschen Turkologenkonferenz Universität Mainz, 4–7 Oktober, 2002* [Turcologica 69], H. Boeschoten & H. Stein (eds), 116–130. Wiesbaden: Harrassowitz.

- Kabadayı, A. 2012. Okulöncesi çocukların Türkçe ediniminde yaptıkları kurallaştırma hatalarının incelenmesi (An investigation of preschool children's overregularization errors in the acquisition of Turkish). *Turkish Studies* 7(3): 1561–1573.
- Ketrez, F. N. 1999. Early verbs and the acquisition of Turkish argument structure. MA thesis, Boğaziçi University İstanbul.
- Ketrez, F. N. 2017. The emergence of nominal compounds in Turkish: A case-study on structural simplicity vs. input frequency. In Dressler, Ketrez & Kilani-Schoch (eds), 231–249.
- Ketrez, F. N. & Aksu-Koç, A. 2007. The (scarcity of) diminutives in Turkish child language. In *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43], I. Savickiene & W. U. Dressler (eds), 279–293. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.12ket>
- Kilani-Schoch, M. & Dressler, W. U. 2002. The emergence of inflectional paradigms in two French corpora: An illustration of general problems of pre- and protomorphology. In *Pre-and Protomorphology*, M. D. Voieikova & W. U. Dressler (eds), 45–59. Munich: Lincom.
- Korkmaz, Z. 2003. *Türkiye Türkçesi Grameri (Şekil Bilgisi)* (Turkey Turkish (The form)). Ankara: Türk Dil Kurumu Yayınları.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk, Vol. 1: Transcription, Format and Programs*. Mahwah NJ: Lawrence Erlbaum Associates.
- Sebüktekin, H. 1974. Morphotactics of Turkish verb suffixations. *Boğaziçi Üniversitesi Dergisi: Humaniter Bilimler-Humanities* 2: 87–116.
- Sofu, H. 1995. Acquisition of Lexicon in Turkish. PhD dissertation, Çukurova University Adana.
- Sofu, H. 2016. Adların ve eylemlerin edinimi (Acquisition of nouns and verbs). *Çukurova Üniversitesi Türkoloji Araştırmaları Dergisi* (Çukurova University Journal of Turcology Research) 1(1): 90–97.
- Stephany, U. 1982. Inflectional and lexical morphology – A continuum. *Glossologia* 1: 27–55.
- Uzun, N. E., Uzun, L. S., Aksan, Y. K. & Aksan, M. 1992. *Türkiye Türkçesinin Türetim Ekleri: Bir Döküm Denemesi* (Derivation affixes of Turkey Turkish: A documentation attempt). Ankara: Şirin.
- van der Heijden, H. 2011. Word formation processes in young bilingual children. In *Bilingualism and Migration*. G. Extra & L. Verhoeven (eds), 123–142. Berlin: Walter de Gruyter.
- van der Heijden, H. & Verhoeven, L. T. 1994. Early bilingual development of Turkish children in the Netherlands. In *The Cross-linguistic Study of Bilingual Development*, G. Extra & L. T. Verhoeven (eds), 51–73. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.
- Verhoeven, L. T. 1988. The bilingual child as a word maker: Word formation processes in Turkish and Dutch. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistic Society*, S. Axmayer, A. Jaisser & H. Singmaster (eds), 232–343. Berkeley CA: Berkeley Linguistics Society. <https://doi.org/10.3765/bls.v14i0.1790>
- Wilkins, J. 2016. Turkish. In *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 5, P. O. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds), 3367–3385. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110424942-016>
- Zülfikar, H. 1991. *Terim Sorunları ve Terim Yapma Yolları* (Terminology problems and ways of terminology derivation). Ankara: Atatürk Kültür Dil ve Tarih Yüksek Kurumu.

Conclusions

Veronika Mattes¹ and Wolfgang U. Dressler^{2,3}

¹University of Graz / ²University of Vienna / ³Austrian Academy of Sciences

1. Introduction

The present volume provides new insights into the early development of derivational morphology by means of longitudinal studies on ten languages. It is noteworthy that the domain of word formation has been understudied in comparison to other domains of language acquisition such as the lexicon, syntax or inflectional morphology.

Our results allow for some generalisations about the development of derivation, i.e. developmental processes that can be observed in all children independent of the language they acquire. But they also show important typological differences, i.e. an obvious impact of the structural properties of the specific language on emergence, use and the early course of development of derivational patterns in early child language.

Our language sample comprises languages which differ in their derivational richness, in their morphological systems and consequently in the characteristics of derivational patterns. In terms of classical morphological typology, the languages can be assigned to two synthetic categories: The agglutinating languages Turkish, Finnish, and (to a lesser extent) Estonian, have a very large inventory of derivational affixes, most of which are morphotactically transparent. The more fusional languages (albeit less so in derivation than in inflection) such as the Slavic languages Russian and Croatian, but also Lithuanian, French and Greek, are characterised by a higher percentage of non-transparent derivations. Although the Germanic languages Danish and German possess a broad inventory of derivational patterns, they are predominantly compounding. In the agglutinating languages, we observe an early emergence and productivity of derivational patterns, which we expected, because of the high transparency, salience and regularity of derivational morphology.

In many languages, word formation has more of an agglutinating character than inflection, which in turn tends to be more fusional (Skalička 1979). Compared to agglutinating languages, certain derivational patterns in fusional languages, with their rich polysemy or allomorphy, are morphotactically less transparent and

further removed from the ideal of a biunique relation between meaning and form. The acquisition of derivation does not appear to be facilitated by agglutinating languages, however, when compared with more fusional ones. Children tend to detect and use derivations roughly at the same age in all of the ten languages studied. The higher proportion of transparent patterns in agglutinating languages is nonetheless reflected by a larger number of patterns in early use by children.

The present volume offers a systematic description of the first phases of the acquisition of derivation. As mentioned above, this domain has not been studied in much detail until now. Previous studies on word formation have shown that diminutives and compounds are the first nominal patterns to emerge (Clark 2014: 426; Dressler, Ketrez & Kilani-Schoch 2017; Dressler, Lettner & Korecky-Kröll 2012; Savickienė & Dressler 2007). There are several language specific studies on the acquisition of derivation (see Introduction chapter), but our volume takes the first cross-linguistic perspective on the early development of derivation.

Children clearly focus from the very beginning on those word-formation patterns that are favoured in their language (Clark & Berman 1984: 578). However, the claim has been made that derivation emerges either later than compounding or at least simultaneously with it due to its lower semantic transparency (Dressler, Ketrez & Kilani-Schoch 2017: 288). This view is challenged by our findings, since derived lexemes (in addition to diminutive nouns) emerge earlier than compounding in half of the languages in our volume (see Section 2).

With respect to productivity, the common view is that children start to make productive use of derivational affixes at the same age across languages and that the productive use of derivation lags behind that of compounding due to its greater complexity and lower morphosemantic transparency (Berko 1958; Clark & Hecht 1982; Clark 1993: 176; Swan 2000). Furthermore, Clark (2016: 308) states that “[i]n languages that make little use of zero-derivation or of compounding, children produce fewer lexical innovations before age three or so”, compared to children acquiring languages that use these less complex and highly morphosemantic transparent morphological pattern. Her statement seems indeed to be confirmed when we look at the Estonian or Lithuanian data, however, the analysis of the French, Finnish or Turkish data clearly shows that this observation may not be generalised. Nevertheless, in spite of the fact that derivation is the dominating word-formation process in Turkish, noun-noun compounding is more productive in the early stages, due to its higher semantic transparency and simplicity (see Ketrez & Aksu-Koç this volume).

Although a relatively large repertoire of derivational patterns emerges in the first three years and is in part used productively by children, it only represents a small portion of the patterns found in adult-directed speech (ADS) and even more so in written language. Thus, notwithstanding the large achievements in word formation in the first three years that we observed, “[d]erivational affixes appear to

enter in large numbers only after age three” (Clark & Berman 2004: 1803). An even greater development in this domain is to be expected in later preschool age, with major progress taking place during school age up to adolescence (Anglin 1993; Nippold 2007; Ravid 2004).

2. Emergence of derived lexemes and derivational patterns

Looking at the emergence of derived lexemes in the speech of the children studied we find age differences among the ten languages examined and also, of course, between individual children. The earliest emergence of derived lexemes that we observed in the data lies between the ages 1;3 (Estonian) and 2;6 (French) in all languages. Recordings of derivations as early as 1;3 or 1;4 are single instances and cannot be considered as transparent to the child. Most children start to use derived words before the age of 2. The appearance of derived words in child speech (CS) on average, can be expected around 1;7 to 1;8 in any language and the number of derivations (tokens and lemmas) usually starts to increase from the age 1;7 onwards. Derived verbs appear shortly before derived nouns (except for diminutives) in all languages.

In the majority of languages, derivation emerges in the same developmental stage as compounding (with the exception of diminutives, which often emerge earlier, see Introduction chapter); the first derived words are found simultaneously with (Greek, German, Estonian, Finnish), or earlier (Danish, Croatian, Russian, Lithuanian, French) than compounds. In fact, there is only one language in our sample, Turkish, in which derivation emerges later than compounding. This is only true, however, if voice morphology is classified as inflectional and is thus not considered.

Since the age at which the first derived lexeme is uttered by the child (in the recordings) cannot be taken as the point at which the child starts to detect derivational morphology, we chose to apply the mini-paradigm criterion (see Introduction chapter) as starting point for the acquisition. The criterion is fulfilled when a pattern and a base have appeared in three different combinations in CS. This method does not of course prove that the child has established the respective pattern or even extracted a general rule, but it fixes a data point at which the child has enough material at her or his disposal to potentially detect the respective derivational pattern (“potential productivity”). The first mini-paradigms for derivation were fulfilled at very different ages (Croatian: 1;6, Finnish: 1;7/1;8, German: 1;9/2;1/2;3; Turkish: 1;7/1;11, Lithuanian: 1;7/1;10, French only one child: 1;9, Russian 1;8/2;5). In the majority of cases, verb derivations are the first to form mini-paradigms. In the Danish data, no mini-paradigm was identified up to the age of 3;0.

The respective various first detected patterns differ among the languages. In the case of verbs, causative derivations are among the earliest categories, if the language makes productive use of them (e.g. Turkish, Finnish, Estonian). With noun derivation, instruments and agents dominate. There is also a language in our sample, however, in which nominal derivations (other than diminutives) remain very infrequent at all in CS up until the age of 3 and this is Greek.

There is evidence that suffixes generally appear before prefixes (Clark 2016: 308) due to the early prevalence of the recency effect (see Introduction chapter). In the languages we studied, suffixes do indeed dominate in the nominal derivations produced by the children. In languages that regularly use prefixes for derivation, however, these also emerge very early in CS, in our sample mainly in Greek and Slavic verbal derivation. Furthermore, our data confirm that transparency plays a central role in productivity. Morphotactically and morphosemantically transparent affixes are productively employed at early stages of language development, whereas innovations involving potentially opaque affixes, that are more difficult to identify and interpret, only appear later, mainly after the acquisition of initial reading skills (see also Clark 1993; Gahl & Plag 2019).

3. Emergence of neologistic derivations and productivity

The appearance of neologisms, i.e. derivational patterns creatively applied by the child, is widely considered to be an important and reliable indicator of the productivity of these very derivational patterns themselves (Bowerman 1982; Clark 1982; Clark & Berman 1984, 2004; Meibauer, Guttropf & Scherer 2004; Swan 2000), since it reveals that a child is applying a morphological schema or rule creatively in order to produce a word form which has not been taken over from the input (Anglin 1993: 28).

The emergence of neologisms as a major proof for productivity, however, is unsuitable in longitudinal studies that are based on regular recordings that can only capture a very limited extract of the actual speech production by children. Furthermore, children differ in their propensity to create neologisms (depending among other things on their acquisition style and personality; see Anglin 1993: 139; Elsen & Schlipphak 2015: 2118). Even in the data of a very talkative and innovative child, lexical neologisms are relatively rare, as opposed to inflectional over-generalisations. One reason is that there is no obligatory syntactic context for lexical innovations (Swan 2000: 188). Although Clark (1993) calculated that her son produced 3 lexical neologisms per day on average up to 5;11, it is very improbable to record any of them in a 30 or 60-minute session once or twice a month.

Furthermore, some neologisms might also remain undetected in the recordings, as some of them accidentally result in a canonical word (see Hržica, this volume). We thus did not rely on neologisms as a chronological indicator of pattern productivity, but insofar as these were available from the recordings or from additional diary notes, they were included in the study as supplementary evidence to the “potential productivity criterion” (see Section 4.4).

Neologisms primarily serve to fill individual or language specific “lexical gaps” (Clark 1982: 390; Bushnell & Maratsos 1984; Elsen & Schlipphak 2015: 2119) that young children often experience because of their comparatively limited vocabulary. In later developmental stages, neologisms also have further functions (such as stylistic variation or intended creativity), but these do not play a role in the time span focused in the present volume.

As Clark (2016: 289) points out, “[children] attend to the relative productivity of different word-formation options for specific meanings [...] and favour more productive over less productive options”. Although the productivity of a word-formation pattern in the language acquired influences its productivity in CS (Dressler, Libben & Korecky-Kröll 2014; Laaha et al. 2006), it does not predict emergence and productivity in early CS; i.e. high productivity of an adult pattern does not necessarily lead to high productivity in early CS, and vice versa. Even low or zero productivity of a pattern in ADS can lead to innovative derivation in early CS, as long as it is transparent and occurs with sufficient lemma frequency in child-directed speech (CDS). Examples are nouns derived by the suffix *-e* in German (e.g. *Rutsch-e* ‘slide’ from the verb *rutsch-en* ‘slide’, for which the mini-paradigm criterion was fulfilled by Jan at 2;0) and nouns derived by the suffix *-mUs* in Finnish (used by Tomi in the neologism *ampu-mus* ‘gun; cannonball’ from the verb *ampua* ‘shoot’ at 2;8). Both patterns are frequent and transparent in the respective language, but not (very) productive in ADS, in the sense that they are not used synchronically to derive new lexemes. Neologisms of the kind suggest that morphotactic transparency and frequency of a pattern are the main factors for productivity in early CS, in which the child does not yet have enough language experience to consider the productivity of a pattern (see Clark 1993: 137; Goldberg 2006: 57). The knowledge of pattern productivity in the language is part of the stage of “core morphology” (see Introduction chapter) or, in Berman’s (2004) terms, the stage of conventionalisation, in which knowledge and use are integrated.

4. Other aspects of the development of derivational morphology

4.1 Preferences in derivational morphology

Summarizing the results for the role of the preferences in derivational morphology in early language acquisition (see Introduction chapter), in all languages studied, morphotactic and morphosemantic transparency enhanced morphological decomposition by the child. All derivational affixes that are acquired early by the children (in the sense of documented or potential productivity) are transparent in both senses.

In some languages we observed that prefixes or particles that are homophonous with prepositions and/or adverbs (with an obvious related meaning) are acquired earlier than others (e.g. German *an-*, *auf-*, *aus-*; Lithuanian *i-*, *su-*), obviously due to their higher morphotactic and morphosemantic transparency. What Libben (2014) states for compounding, would also appear to have an effect on such prefixes: The relationship of prefixed words with independent words adds additional word associations to the word family, which renders them more transparent.

The preference for suffixation over prefixation has been confirmed insofar as suffixed nouns dominate over prefixed nouns in CS derivations in languages that have derivational prefixes and suffixes. This observation must not, however, be reduced to the perceptive preference of children for suffixes over prefixes due to greater processing ease (recency effect), but might also be a mere consequence of the role prefixes play in our languages studied (see also Section 4.2). It would thus be very insightful to compare this finding to the course of acquisition in predominantly prefixing languages, such as Austronesian, Athabaskan or Bantu languages.

4.2 Word classes

One of the central questions in the present volume is whether there is a noun or a verb bias in the development of derivation.

The general result across our languages is that a bias to verb derivation – in terms of frequency and chronological order – does occur (in most languages only if we exclude nominal diminutives, that are on the border to inflection, see Introduction chapter). Derived verbs outnumber derived nouns in tokens and lemmas, they emerge earlier than (e.g. Croatian, Estonian, Greek, German) or at least simultaneously (e.g. Finnish, Russian) with derived nouns, and in our sample there are languages in which only verbal derivation becomes productive before age three (Estonian, Lithuanian, Greek). Only in French does noun derivation emerge earlier than verb derivation and it is more frequent. This finding might be unexpected, as for the early acquisition phase a general noun bias has been postulated (Gentner & Boroditsky 2001; see Introduction chapter), although discussed and challenged by

many authors at least for some languages, for example Korean, Mandarin Chinese, or Tzeltal (Choi 1998; Tardif 1996; Brown 1998). In German, a language for which the noun bias holds in principle, particle verbs appear in large numbers and are also used productively very early.

Adjectives, which typically emerge later than nouns and verbs (see Introduction chapter), appear as derived lexemes in CS and CDS of all languages. In Danish, German, Finnish, and French, at least one pattern that derives adjectives fulfils the mini-paradigm criterion in early CS (e.g. *-lig* in Danish, *-ig* in German, *-inen* and *-vA* in Finnish, and participles converted into adjectives as the only pattern in French). In others, e.g. Lithuanian or Croatian, adjectives remain unproductive up to age 3;0. In Greek, derivational suffixes of adjectives derived from nouns are semantically empty. Children derive adjectives from nouns rather than from verbs, as is the case in the target languages, since adjectives denote properties of objects rather than of actions or states.

One of the central functions of derivation is the change of word classes (in contrast to inflection and to compounding where the head of a compound most often maintains its original head status). For language acquisition, the ability to derive word classes from each other is – among other factors – associated with the development of word-class distinction, as pointed out by Berman (1988: 54):

[S]trong evidence for the ‘grammaticalization’ of word-class distinction is available once children are capable of forming denominal verbs and of constructing nouns out of verbs, or adjectives out of either nouns or verbs – in accordance with the structural constraints of their particular target language.

Comparing word-class changing and word-class maintaining derivation, we find that the majority of children’s noun derivations belong to the first group and verb derivation more often to the latter. For example, in Greek, word-class changing verbal derivation by suffixation is not transparent and not productive in CS, whereas word-class maintaining verbal prefixation is morphosemantically transparent and productive (e.g. the Greek intensifying prefixes *para-* expressing exaggeration, and *kse-* with a reversative meaning or *ksana-* ‘again’; see Stephany this volume). Higher frequency of a derivational relation may be connected with higher productivity and morphosemantic transparency.

In most of the languages studied in this volume that have suffixes and prefixes, there is a tendency in early CS for derivational prefixes and suffixes to be distributed according to word classes, in the sense that verbs are primarily derived by prefixes (and mainly word-class maintaining derivation), and nouns are mainly derived by suffixes, most frequently based on verbs and adjectives (Russian, German, Croatian, French). This might further support the categorization process of word classes in which derivational morphology is involved.

In languages with a rich morphological system (including inflection), the distribution of word-class-specific morphemes is more robust than in languages with poorer morphology (Kauschke 2007: 48). In the first type of languages, word classes are primarily established via inflectional and derivational morphology, whereas in the latter, word-class distinction is mainly a question of syntactic usage, i.e. under-specified lexemes can often be used in two or more syntactic functions (e.g. Vogel 2000). This leads us to conversion, a specific case of derivation in that it changes the word class of a lexeme without any formal marking (except inflection in morphological conversion), and as such is much more prevalent in languages with less elaborated morphology. It is on the one hand assumed to be the most simple way of deriving word classes, and as such to be preferred by children (Clark 1993, 2016: 301; Swan 2000: 190), although this assumption is based on acquisition of English, which is especially rich in conversion rules. On the other hand, it might not be preferred because it is non-iconic and (also) less transparent than affixation (see Introduction chapter). Conversion can be interpreted as a powerful means of changing word classes (e.g. Marchand 1963; Booij 2002), but conversion can also point towards a lack of word-class distinction (e.g. Bybee 1995; Haspelmath 1996) – either in the language in general (as argued for English), or in the child’s lexicon.

In some of the languages (i.e. Danish, German, French, Croatian and Russian), conversion is used to a considerable degree by children, but only in French and Danish it is preferred over word-class changing derivations by affixation. In 50% of the languages from our sample, conversion plays a minor role or does not appear at all in CS (Finnish, Lithuanian, Turkish, Greek, very rare in Estonian), reflecting its low occurrence in the respective target languages.

These contrasting results show that conversion is not a preferred derivational means per se due to its simplicity, but that its usage mainly depends on the structure of the language acquired. In any case, our cross-linguistic results confirm Clark’s (2016: 303) statement that “... where zero-derivation offers a productive option [...] children make use of it early on.”

4.3 Semantic categories of early derivations

The chronological order of the emergence and development of meanings expressed by children depends on general factors influencing acquisition, especially input frequency and conceptual complexity, but also the relevance for the communicative goals of children (Ambridge, Kidd, Rowland & Theakstone 2015). Derivational patterns and derived words are typically polysemous (Lieber 2004; Rainer 2014), but “[...] polysemy does not restrict productivity. It rather encourages an increase in productivity, especially as far as agent-instrument polysemy is concerned” (Scherer 2011: 50). A few studies, such as Clark & Hecht (1982) and Meibauer, Guttropf &

Scherer (2004), which focus the order of the acquisition of derivational semantic categories of nouns have shown that children begin early to use – rote learned – affixes for several nuances of meaning and that there would appear to be no “natural” chronology in the sense of a basic meaning and a gradual semantic extension.

Most derived nouns in CS across all the languages studied in the present volume, refer to objects (mainly instruments) and persons (mainly agents), to a lesser degree to locations, results and, in some languages (e.g. Estonian, Finnish, Russian, Turkish) also to processes/actions. In most languages, instrument nouns emerge earlier and are more frequent (in lemmas and tokens) than agent nouns, presumably because small children are more motivated in naming objects by referring to their functions rather than to persons (which they usually refer to with simple nouns) (see Salerni, Assanelli, D’Odorico & Rossi 2007). However, there are also languages in our sample, i.e. Greek and Estonian, in which instrument noun derivation is scarce or remains unproductive in CS. Not surprisingly, in this early phase, abstract nouns (and in some languages such as French, also action nouns) are rare or absent (e.g. Lithuanian) in CS and are not observed to be derived productively up to age 3.

For verbs, in our languages, aspectual or aktionsart derivations (mainly “telicity”, e.g. Russian, Croatian, German) or causatives (e.g. Turkish, Finnish, Estonian, Lithuanian) emerge early. As already discussed above (see Section 4.2), verbal derivation plays an important role in early CS in all languages, exhibiting high potential productivity.

When considering what is underrepresented in our CS data, we can conclude that the acquisition of derivation of action nouns, abstract nouns, categories of verbal derivation such as non-telic classes of aspect and aktionsart, and the majority of adjectival derivation, all occur in later development. This progress in derivational morphology goes hand in hand with cognitive and semantic development, and, later with the acquisition of written language (see Carlisle 1988; Nippold 2007; Ravid 2004; Tolchinsky 2004), including the different use of word formation in different genres (see Panagl 2019).

Our general observation is that in most of the 10 languages studied (e.g. Estonian, Lithuanian, Greek, Croatian, but not in French) verbs produced early are richer in derivational patterns than nouns (except for diminutives) (see Section 4.2). However, this result should be interpreted in a differentiated manner: In most cases it is valency-changing morphology that dominates in early verbs (e.g. Turkish, Estonian) or particle verb formation (German) and both morphological patterns are not prototypically derivational. Valency-change overlaps with inflection and particle verb formation with compounding. In any case, we see a tendency that derivations of prototypical nominal semantics, such as reference to persons or objects, emerge at an early stage, whereas the first derivations for verbs are often borderline cases of derivations, such as changes in voice or valency (mainly by causative marking).

4.4 (Potential) Productivity in CS

The productivity of patterns in CS is a key issue in the assessment of the underlying linguistic representations (Tomasello 2000: 67). As in our corpus studies based on longitudinal recordings of spontaneous speech a strict proof of productively applied derivational patterns is not possible, we looked for the potential productivity of a pattern, in applying the mini-paradigm criterion (see Introduction chapter).

In all languages we found a documented (by neologisms) or potential (by mini-paradigms) productivity of patterns of verbal derivations before the age of 3, but this is very largely not the case for derived nouns and even less so for adjectives. For example, in Estonian and Lithuanian, verbal derivation is potentially productive, but nominal derivation is not, except for diminutives. In Greek, potential productivity is higher for certain prefixed verbal derivatives, namely those which are morphotactically and morphosemantically transparent, than for nominal derivatives. But nominal derivation plays an important role in Danish, German, Finnish, Turkish, Croatian, Russian and French.

4.5 Complexity

A gradual rise of complexity can be predicted generally for acquisition (see Brown 1973; Dressler, Ketrez & Kilani-Schoch 2017; Swan 2000; also Clark 1993: 119–122 on simplicity). Complexity can be formal and/or conceptual (Slobin 1973, 1985). With respect to derivational morphology, formal complexity is caused by allomorphy, morphotactic opacity and combinations of derivational patterns in one lexeme. Differences regarding formal complexity in languages lead to different rates of acquisition (Clark 2016: 5), at least in terms of number of morphemes and morphotactic opacity. In the development of derivational morphology up to age 3, the complexity in the sense of combinations of affixes does usually not exceed two derivational affixes or patterns in one word (it is attested in Turkish and Greek though). However, derivation in combination with compounding appears soon after the first simple compounds and derivations (e.g. Estonian, German).

Conceptual complexity is caused by polysemy, morphosemantic opacity and abstract meanings. Abstract meanings are conceptually more complex than concrete meanings, predicting that concrete meanings of derivatives will emerge before abstract ones (Introduction chapter), and this is confirmed for derivational categories by our studies. Both aspects of complexity come into conflict in conversion, which does not lead to an increase of the number of derivational morphemes, but lacks iconicity (see Section 4.2).

5. Relations between CDS and CS

A main topic of the volume is the relation between CDS and CS in terms of frequency and emergence of derived words, i.e. to what extent does CS follow CDS?

Lemma or type frequency (i.e. lexical diversity) in CDS is the critical factor for the detection and generalisation of derivational patterns, as it is for word formation more generally (Bybee 1985, 2010; Clark 1993: 138–139; Cordes 2014; Elsen & Schlipphak 2015: 2131).

The relation between CS and CDS is always an indirect one. Although caretakers usually adapt to the stage of linguistic knowledge reached by the children (“fine tuning”, Snow 1995), CS does not mirror exactly the distributions of patterns in CDS, because children can only make use of input structures which they can both perceive well and integrate into their current language system (“uptake”, see Harris 1992; Clark 2010). With respect to derivation, such a mismatch between CS and CDS can be observed in several of our corpora, for example, where certain numbers of derived abstract nouns in CDS were found that are not reflected in CS. In CS almost exclusively derived lexemes with concrete meanings are found.

Overall, parallels in CS and CDS were found in the languages studied, but to varying degrees: for example, in Croatian the percentages of lemmas of derived verbs are parallel in CS and CDS, but that of derived nouns is higher in CDS than in CS. The noun derivations with abstract meanings in CDS do not occur in CS. In German, lemmas and tokens of prefixed verbs are much more frequent in CDS than in CS, as children obviously prefer to use the more salient and transparent particle verbs. In Russian and Estonian, no clear impact of CDS on CS could be detected.

Frequency and variety (repertoire) of derivational patterns in CDS are usually (much) higher than in CS. This is also true for the *late acquired* derivations, such as nominal prefixes, verbal suffixes, and affixes with abstract semantics.

As mentioned above (see Sections 3 and 4.3), chronological order of emergence and later mastery only partially depend on frequency in CDS (Ambridge et al. 2015; Ravid, Keuleers & Dressler 2020). The fact that frequency (i.e. lexical diversity) in CDS alone is not a predictor for productivity in CS but only works in combination with transparency, has been demonstrated for Greek (Stephany this volume): Despite their high lemma frequency in CDS, verbs derived by morphotactic and morphosemantic opaque suffixes do not promote the development of the derivational system while the much less frequently occurring transparent prefixed verbs do.

6. Summary and outlook

We conclude with the insight that the development of derivational morphology up to age 3 can be described as the stage of proto-morphology (see Introduction chapter), and that there is still a long way to go for the development of morphology proper (and the “proficient integration of knowledge and use”, Berman 2004: 14, i.e. full mastery of derivational morphology). As Clark (2016: 294) emphasizes, the age of 3 is only the *start* in word formation, which is, however, less the case for compounding in compounding languages and diminutive formation in most languages.

It is our belief that we have demonstrated that children acquire much knowledge on derivation in the course of their first three years. In each of the languages studied we see that children acquired a core inventory of derivational patterns as a basis to be elaborated rapidly and intensely with growing vocabulary, language experience, and cognitive capabilities.

Our investigations have shown as a generalisation for all languages studied that children detect frequent and transparent derivational patterns that are relevant for their communicative goals and that meet their cognitive abilities (Ambridge et al. 2015) at a relatively early stage, at about 2 years of age (e.g. valency changing verbal derivation, nominal suffixation for instruments and agents) and that they successfully derive verbs and nouns and, to a lesser extent, adjectives before their third birthdays. Although their repertoire is limited to a small subset of the large inventory of derivational patterns of the respective languages, the structural patterns acquired early cover a useful range of functions for the communicative needs of toddlers.

The aim of this volume has been to study the emergence and early development of derivational morphology in a cross-linguistic perspective. Its outcome can only be taken to be a first step towards a deeper understanding of the development of derivational morphology in children. As usual, open ends have remained. Our study based on spontaneous speech data should be supplemented by experimental studies for specific problems, for example when it comes to questions of children’s development of semantic knowledge or rule knowledge. Explanations for observations such as the preference of young children for derivations that are based on verbs (see Section 4.2) might be found when considering the acquisition of derivational morphology in languages with a verb bias, such as Korean or Mayan languages.

When seen from a general cross-linguistic perspective, studies of other language types and especially non-European ones (e.g. prefixing languages, languages with reduplication as a means of word formation, languages with much poorer derivational morphology) and a comparison with extragrammatical morphology would substantially enrich the insights into the acquisition of derivational morphology.

References

- Ambridge, B., Kidd, E., Rowland, C. F. & Theakston, A. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2): 239–253.
<https://doi.org/10.1017/S030500091400049X>
- Anglin, J. 1993. Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development* 58(10): 1–166. <https://doi.org/10.2307/1166112>
- Berko, J. 1958. The child's learning of English morphology. *Word* 14(2–3): 150–177.
<https://doi.org/10.1080/00437956.1958.11659661>
- Berman, R. A. 1988. Word class distinctions in developing grammars. In *Categories and Processes in Language Acquisition*, Y. Levy, I. M. Schlesinger & M. Braine (eds), 45–72. Hillsdale NJ: Lawrence Erlbaum Associates.
- Berman, R. A. 2004. Between emergence and mastery: The long developmental route of language acquisition. In *Language Development across Childhood and Adolescence* [Trends in Language Acquisition Research 3], R. A. Berman (ed.), 9–34. Amsterdam: John Benjamins.
<https://doi.org/10.1075/tilar.3.05ber>
- Booij, G. 2002. *The Morphology of Dutch*. Oxford: OUP.
- Bowerman, M. 1982. Reorganizational processes in lexical and syntactic development. In *Language Acquisition: The State of the Art*, E. Wanner & L. Gleitman (eds), 319–346. Cambridge: CUP.
- Brown, P. 1998. Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics* 36: 713–753. <https://doi.org/10.1515/ling.1998.36.4.713>
- Brown, R. 1973. *A First Language: The Early Stages*. Cambridge MA: Harvard University Press.
<https://doi.org/10.4159/harvard.9780674732469>
- Bushnell, E. & Maratsos, M. 1984. “Spoonling” and “basketing”: Children's dealing with accidental gaps in the lexicon. *Cognitive Development* 55: 893–902.
- Bybee, J. L. 1985. *Morphology: A Study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.9>
- Bybee, J. L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455. <https://doi.org/10.1080/01690969508407111>
- Bybee, J. L. 2010. *Language, Usage and Cognition*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511750526>
- Carlisle, J. 1988. Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics* 9: 247–266.
<https://doi.org/10.1017/S0142716400007839>
- Choi, S. 1998. Verbs in early lexical and syntactic development in Korean. *Linguistics* 36: 755–781.
<https://doi.org/10.1515/ling.1998.36.4.755>
- Clark, E. V. 1982. The young word maker: A case study of innovation in the child's lexicon. In *Language Acquisition. The State of the Art*, E. Wanner & L. Gleitman (eds), 390–425. Cambridge: CUP.
- Clark, E. V. 1993. *The Lexicon in Acquisition* [Cambridge Studies in Linguistics 65]. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511554377>
- Clark, E. V. 2010. Adult offer, word-class, and uptake in early lexical acquisition. *First Language* 30: 250–269. <https://doi.org/10.1177/0142723710370537>
- Clark, E. V. 2014. Acquisition of derivational morphology. In *The Oxford Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 424–439. Oxford: OUP.

- Clark, E. V. 2016. *The Lexicon in Acquisition*, 3rd edn. Cambridge: CUP.
- Clark, E. V. & Berman, R. A. 1984. Structure and use in the acquisition of word formation. *Language* 60(3): 542–590. <https://doi.org/10.2307/413991>
- Clark, E. V. & Berman, R. A. 2004. First language acquisition. In *Morphologie / Morphology: Ein internationales Handbuch zur Flexion und Wortbildung / An international Handbook on Inflection and Word Formation*, G. Booij, C. Lehmann, J. Mugdan & S. Skopeteas (eds), 1795–1805. Berlin: Walter de Gruyter. <https://doi.org/10.1515/9783110172782.2.19.1795>
- Clark, E. V. & Hecht, B. 1982. Learning to coin agent and instrument nouns. *Cognition* 12(1): 1–24. [https://doi.org/10.1016/0010-0277\(82\)90027-0](https://doi.org/10.1016/0010-0277(82)90027-0)
- Cordes, A.-K. 2014. *The Role of Frequency in Children's Learning of Morphological Constructions*. Tübingen: Narr.
- Dressler, W. U., Ketrez, F. N. & Kilani-Schoch, M. (eds) 2017. *Nominal Compound Acquisition* [Language Acquisition and Language Disorders 61]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.61>
- Dressler, W. U., Lettner, L. E. & Korecky-Kröll, K. 2012. Acquisition of German diminutive formation and compounding in a comparative perspective: Evidence for typology and the role of frequency. In *Current Issues in Morphological Theory: (Ir)regularity, Analogy and Frequency* [Current Issues in Linguistic Theory 322], F. Kiefer, M. Ladányi & P. Siptár (eds), 237–264. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.322.11dre>
- Dressler, W. U., Libben, G. & Korecky-Kröll, K. 2014. Conflicting vs. convergent vs. interdependent motivations in morphology. In *Competing Motivations in Grammar and Usage*, B. MacWhinney, A. Malchukov & E. Moravcsik (eds), 181–196. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780198709848.003.0011>
- Elsen, H. & Schlipphak, K. 2015. Word-formation in first language acquisition. In *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 3, P. Müller, I. Ohnheiser, S. Olsen & F. Rainer (eds), 2117–2137. Berlin: De Gruyter.
- Gahl, S. & Plag, I. 2019. Spelling errors in English derivational suffixes reflect morphological boundary strength: A case study. *The Mental Lexicon* 14: 1–36. <https://doi.org/10.1075/ml.19002.gah>
- Gentner, D. & Boroditsky, L. 2001. Individuation, relativity, and early word learning. In *Language Acquisition and Conceptual Development*, M. Bowerman & S. Levinson (eds), 215–256. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511620669.010>
- Goldberg, A. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: OUP.
- Harris, M. 1992. *Language Experience and Early Language Development: From Input to Uptake*. Hove: Psychology Press.
- Haspelmath, M. 1996. Word-class changing inflection and morphological theory. In *Yearbook of Morphology 1995*, G. Booij & J. van Marle (eds), 43–66. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-3716-6_3
- Kauschke, C. 2007. *Erwerb und Verarbeitung von Nomen und Verben* (Acquisition and Processing of Nouns and Verbs). Tübingen: Niemeyer. <https://doi.org/10.1515/9783110938692>
- Laaha, S., Ravid, D., Korecky-Kröll, K., Laaha, G. & Dressler, W. U. 2006. Early noun plurals in German: regularity, productivity or default? *Journal of Child Language* 33: 271–302. <https://doi.org/10.1017/S0305000906007379>
- Libben, G. 2014. The nature of compounds: A psychocentric perspective. *Cognitive Neuropsychology* 31: 8–25. <https://doi.org/10.1080/02643294.2013.874994>
- Lieber, R. 2004. *Morphology and Lexical Semantics*. Cambridge: CUP.

- Marchand, H. 1963. Die Ableitung desubstantivischer Verben mit Nullmorphem im Französischen und die entsprechenden Verhältnisse im Englischen und Deutschen (The derivation of denominal verbs with zero-morpheme in French and the respective relations in English and German). *Zeitschrift für französische Sprache und Literatur* (Journal of French Language and Literature) 73(3–4): 164–179.
- Meibauer, J., Guttropf, A. & Scherer, C. 2004. Dynamic aspects of German -er-nominals: A probe into the interrelation of language change and language acquisition. *Linguistics* 42(1): 155–193. <https://doi.org/10.1515/ling.2004.002>
- Nippold, M. 2007. *Later Language Development: School-age Children, Adolescents, and Young Adults*. Austin TX: Pro-Ed.
- Panagl, O. 2019. Wortbildung und Textsorte: Verbalabstrakta in der spät(er)en Latinität (Word formation and Text type: Verbal abstracts in the later Latinity). *Acta Antiqua Academiae Scientiarum Hungaricae* 59: 387–397. <https://doi.org/10.1556/068.2019.59.1-4.34>
- Rainer, F. 2014. Polysemy in derivation. In *The Oxford Handbook of Derivational Morphology*, R. Lieber & P. Štekauer (eds), 338–353. Oxford: OUP.
- Ravid, D. 2004. Derivational morphology revisited: Later lexical development in Hebrew. In *Language Development across Childhood and Adolescence* [Trends in Language Acquisition Research 3], R. A. Berman (ed.), 53–81. Amsterdam: John Benjamins. <https://doi.org/10.1075/tilar.3.07rav>
- Ravid, D., Keuleers, E. & Dressler, W. U. 2020. Emergence and early development of lexicon and morphology. In *Word Knowledge and Word Usage: A Cross-disciplinary Guide to the Mental Lexicon*, V. Pirrelli, I. Plag & W. U. Dressler (eds), 593–633. Berlin: De Gruyter.
- Salerni, N., Assanelli, A., D’Odorico, L. & Rossi, G. 2007. Qualitative aspects of productive vocabulary at the 200- and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language* 27: 75–87. <https://doi.org/10.1177/0142723707067545>
- Savickienė, I. & Dressler, W. U. 2007. *The Acquisition of Diminutives: A Cross-linguistic Perspective* [Language Acquisition and Language Disorders 43]. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43>
- Scherer, C. 2011. Polysemy and productivity in German. *STUF* 64(1): 41–52. <https://doi.org/10.1524/stuf.2011.0004>
- Skalička, V. 1979. *Typologische Studien* (Typological Studies). Braunschweig: Vieweg. <https://doi.org/10.1007/978-3-322-86349-2>
- Slobin, D. 1973. Cognitive prerequisites for the development of grammar. In *Studies of Child Language Development*, C. A. Ferguson & D. I. Slobin (eds), 175–208. New York NY: Holt, Rinehart & Watson.
- Slobin, D. 1985. Crosslinguistic evidence for the language-making capacity. In *The Crosslinguistic Study of Language Acquisition*, D. I. Slobin (ed.), 1157–1256. Hillsdale NJ: Lawrence Erlbaum Associates.
- Snow, C. 1995. Issues in the study of input: Finetuning, universality, individual and developmental differences, and necessary causes. In *The Handbook of Child Language*, P. Fletcher & B. MacWhinney (eds), 180–193. Oxford: Blackwell.
- Swan, D. 2000. How to build a lexicon: A case study of lexical errors and innovations. *First Language* 20: 187–204. <https://doi.org/10.1177/014272370002005904>
- Tardif, T. 1996. Nouns are not always learned before verbs: Evidence from Mandarin speakers’ early vocabularies. *Developmental Psychology* 32(3): 492–504. <https://doi.org/10.1037/0012-1649.32.3.492>

- Tolchinsky, L. 2004. The nature and scope of later language development. In *Language Development across Childhood and Adolescence* [Trends in Language Acquisition Research 3], R. A. Berman (ed.), 233–248. Amsterdam: John Benjamins. <https://doi.org/10.1075/tilar.3.15tol>
- Tomasello, M. 2000. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11(1–2): 61–82.
- Vogel, P. 2000. Grammaticalisation and part-of-speech systems. In *Approaches to the Typology of Word Classes*, P. Vogel & B. Comrie (eds), 259–284. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110806120.259>

Subject index

A

ablaut 24, 33, 104, 110–111
see also umlaut

abstract
noun 57, 71, 77, 120, 133–134, 143, 147–149, 180, 206, 297, 299
verb 112

adjective 57

meaning 4, 112, 133–134, 180, 185, 191–192, 219–220, 255, 298–299

action, activity 12, 30, 53, 56, 64–65, 68, 76–77, 90, 92, 110–111, 115, 133–134, 143–144, 147–150, 161, 184–185, 187, 190–191, 199–201, 204–205, 213, 219–220, 228–229, 242–243, 250, 255, 277, 282–283, 297

adjective/adjectival
base 38, 57, 227 *see also* deadjectival

derivation/formation
2, 12–13, 25, 40, 47, 58, 67, 78, 93–94, 102, 105, 113, 115, 117, 123–124, 130, 132–134, 170, 174–175, 180–181, 186, 188–189, 201, 204–208, 210, 213–214, 252–259, 295, 297

prefix 102, 171 *see also* prefixed adjective

suffix 12, 38, 40, 58, 67, 78, 133, 257 *see also* suffixed adjective

agent (noun) 3, 12, 14, 37, 40–41, 45, 47, 56, 61, 64–65, 68, 76–77, 91, 103, 111, 114, 120, 127, 133, 162–163, 185, 199, 220, 228–229, 239–240, 250, 258, 260, 263, 282–284, 296–297
see also agentive (noun)

agentive (noun) 9, 38, 61, 87, 117, 133, 179, 229, 267–268, 277, 283 *see also* agent (noun)

analogy 8, 103

aspect(ual) 13, 39, 43, 141, 143–144, 150–153, 159, 162–163, 170, 182–183, 187, 190, 192–193, 204–205, 297

C

causative, causativity 13–14, 113, 183, 187, 217, 219–220, 224–225, 229–230, 232, 234, 238, 246–247, 256–257, 259, 263–266, 271, 273–275, 277, 284, 292, 297

circumfix(ation) 3, 111, 113, 198, 200

composition, compound(ing)
3, 8–9, 11–14, 21–23, 36, 46, 55, 57, 79, 87–89, 95, 99, 104–105, 112, 114, 116, 119, 123, 130–132, 141–142, 169–171, 173, 178, 185–186, 189, 203–204, 206, 222, 224, 231–234, 237–240, 243, 245, 252, 255, 259–260, 263, 265–267, 270–271, 274–275, 277, 279–282, 284–285, 290–291, 295, 297–298, 300

complex(ity) 3–4, 6–7, 25, 29, 31, 43, 46, 48, 59–61, 68, 79–80, 89–90, 99, 116, 132, 135, 141, 145, 185, 187, 192–193, 218, 230–232, 234, 238, 266, 269–270, 274, 277, 290, 296, 298

conversion 4, 11–12, 24, 53–58, 70, 74, 78, 80, 87–88, 92–96, 99, 101–104, 110, 112–116, 118–122, 127–132, 134, 139–140, 143, 170–171, 178, 185, 192, 198, 221, 226–227, 238, 251, 257,

265–266, 280–282, 295–296, 298 *see also* zero-derivation, zero-suffixation

course of acquisition/
development 8, 10, 38, 42, 54, 86, 124–126, 129, 133, 135, 144, 212, 289, 294, 300

critical mass 131–132, 214

Croatian 10–11, 13, 141–164, 289, 291, 294–299

D

Danish 10–12, 85–105, 250, 289, 291, 295–296, 298

deadjectival 2, 56, 92, 113, 122, 124, 174, 180, 182, 185, 190, 192, 198–201, 204, 244 *see also* adjective/adjectival base

denominal 2, 25, 29, 55–58, 92–94, 99, 103, 111–113, 120, 124, 127, 130, 133, 173, 178–182, 185, 189–190, 193, 198–201, 204, 215, 220, 226, 246–247, 250, 295 *see also* nominal/noun base

deverbal 2, 12, 55–58, 74–75, 93–94, 111, 113, 122, 124, 127, 129, 133, 147–149, 161, 174, 179, 182, 185, 190, 193, 198–201, 203–204, 214–215, 220, 228–229, 231, 238, 244–246, 250, 255, 277 *see also* verb(al) base

diminutive 2, 12–13, 22, 28, 55–56, 60–61, 66, 77, 91, 118–119, 122, 143–144, 150, 154, 160, 163, 177, 179, 185–186, 190, 192–193, 199, 202–204, 206, 209, 214–215, 220, 222, 224–229, 231–232, 234, 256–257, 264, 266, 275, 278–281, 283, 290–291, 300

E

emergence

order of 9, 12, 22, 78, 80,
109, 129–131, 179, 181,
183, 185, 187, 203–204,
214, 224–225, 227–231,
234, 259–260, 273–275,
284–285, 299

age of 64, 71–74, 126–128,
130–132, 185, 203, 222,
225, 234, 258, 273–274,
278–279, 281–282

English 4, 9, 55, 87, 89, 91, 110,
113, 135, 141–142, 162–163, 229,
255, 257, 265, 267, 296

Estonian 10–11, 13, 217–235,
289–292, 294, 296–299

F

factitive 57, 200, 224, 229–230,
246

Finnish 10–11, 14, 237–260,
289–298

French 4, 10–12, 53–80, 89, 135,
289–291, 294–298

frequency

input frequency 7, 9, 42,
86, 104–105, 130–131, 133,
224, 296

lemma frequency 6–7, 12,
95, 131–132, 293, 299

token frequency 7, 10, 42,
46–47, 95, 98, 104, 110, 116,
122, 124, 129–131, 144, 162,
174, 188–190, 206, 238, 243,
258, 270–271, 279–280,
283–284, 294, 297, 299

type frequency 7, 10, 28, 47,
59, 87, 95, 98, 104, 116, 131,
144, 162, 243, 258, 260, 299

G

gender 25, 60, 62, 64–65, 68,
76–77, 88, 111, 220

German 2–4, 8, 10–12, 87–91,
109–135, 289, 291, 293–299

Greek 10–12, 21–49, 89, 289,
291–292, 294–299

I

implicit derivation 12, 87–88,
91–92, 94, 101–102, 104, 111 *see*
also ablaut, root (vowel)/stem
modification, umlaut

imitative 219–220, 225–226,
229, 234, 237, 239

inflection 1–3, 7–8, 12–14, 22,
25, 39, 41–44, 47–48, 88–89,
94–95, 128–130, 171, 174–175,
192–193, 198, 202–203, 214–
215, 222, 244, 255, 257, 259–
260, 264 265, 267, 273, 274,
281, 284, 289, 292, 295–297

innovation, innovative 12,
85, 87, 95, 99, 104–105, 133,
159–160, 237–238, 240, 242,
267–268, 285, 290, 292–293
see also neologism

instrument(al) (noun) 9, 12,
14, 34, 55–56, 68, 91, 103, 111,
114, 127, 133–134, 147–149, 179,
185, 200, 219–221, 228–229,
241–242, 257–258, 260, 263–
264, 279, 282–284, 296–297

interjection 99, 246–247, 265

L

lexicalised 45, 55, 59, 61, 69,
116, 133, 213, 220–221, 224, 231,
234, 239, 254

Lithuanian 10–11, 13, 197–215,
227, 289–291, 294–298

location (noun) 56, 151–153,
163, 185–186, 199–200, 245,
256, 297

M

mini-paradigm (criterion) 8,
13–14, 85, 95, 101–102, 105,
109, 116–118, 126–130, 133,
141, 146, 157–159, 161, 185,
207–209, 212, 214, 259, 274,
291, 293, 295, 298 *see also*

potential productivity,
productivity criterion

N

neologism 8, 12–13, 22, 28,
53, 55, 80, 95, 114, 127, 134,
144–145, 156, 159–161, 164,

237–240, 242–245, 247–255,
257–258, 281–282, 285,

292–293 *see also* innovation,
innovative

nominal/noun

base 55–57, 61, 226, 241, 267,
278 *see also* denominal
derivation/formation 13, 38,
48, 55–56, 90–92, 100–101,
105, 111–112, 114, 118, 127,
129, 133, 170, 177, 184–185,
188, 197–198, 204–209, 215,
218–219, 222, 224–227, 229,
231–235, 239, 243, 255, 264,
270–271, 275–276, 278,
292, 294–295, 297–299

prefix(ation) 8, 78, 101, 114,
299 *see also* prefixed noun

suffix(ation) 12, 38, 41, 56,
59, 62, 67, 76–79, 110, 132,
133, 134, 146–147, 193, 220,
224–225, 227, 229, 234, 300
see also suffixed noun

nominalised, nominalization 13,
44–45, 60, 66, 112, 115, 127, 219

O

object 13, 37, 40–41, 45, 53, 56,
68–69, 77, 92, 110, 127, 143,
147–149, 162, 175, 180, 182,
185–186, 188–190, 193, 220–
221, 228–229, 234, 276–277,
282–283, 297

onomatopoeia, onomatopoe(tic)
93, 182–183, 219, 226, 230, 265,
271–272, 275–276, 283

opacity, opaque 2, 12, 29–31, 33,
38, 42–43, 46–48, 55, 59, 61,
68–69, 113, 122, 127, 130, 135,
182, 218, 220, 239, 244, 257,
292, 298–299

P

part of speech 24, 37–38, 150,
206–207, 214, 218–219, 246,
250 *see also* word class

particle (verb) 3, 12, 112–113,
115, 120–122, 126, 128–129, 132,
140, 151, 294–295, 297, 299

passive 133, 183, 201, 248–250,
256, 264, 266, 271, 273–274, 284

- person 13, 37, 110, 127, 129, 143,
147–149, 162, 171, 180, 186,
198–199, 220, 228, 248, 297
- postfix(ation) 2, 170–171,
182–183, 187, 192
- potential productivity 67, 69,
101–102, 105, 118, 207, 269,
274, 279, 291, 293–294, 297–
298 *see also* mini-paradigm,
productivity criterion
- prefixed
- adjective 69, 130, 182,
201 *see also* adjective/
adjectival prefix
- noun 69, 80, 127, 151,
197–198, 201–202, 294
see also nominal/noun
prefix(ation)
- verb 12, 29, 39, 43, 54,
69–70, 78–89, 121–122,
126, 128–129, 131, 141, 144,
151, 153, 155, 161–162, 182,
187, 190, 192, 204–205, 207,
210–214, 298–299 *see also*
verb(al) prefix(ation)
- productive use/productivity
- 6, 9, 12–14, 24, 28, 33, 42, 45,
47–48, 80, 104–105, 114, 115,
120–121, 124, 126–130, 133–134,
142, 144–147, 151, 153, 159, 161,
163–164, 169, 176, 181–182, 185,
187, 190, 193, 207–209, 212,
218, 224–225, 229, 234–235,
238–242, 245, 247–248, 250,
252, 254, 256–260, 265–266,
268, 273, 277–279, 281, 285,
289–299 *see also* innovation,
innovative, mini-paradigm,
neologism, potential
productivity, productivity
criterion
- productivity criterion 67, 268,
275, 278–279, 285, 293 *see also*
mini-paradigm criterion,
potential productivity
- pseudo-affixation 59, 94, 116
- R
- recency effect 5, 7, 104, 111, 184,
292, 294
- reduplication 71, 265, 270, 281
- reflexive 170–171, 183, 187,
190–193, 219, 229, 238, 248–
249, 256, 264, 266, 271, 273
- repetitive action 90, 187, 191
- result(ative) 13, 56, 64–65,
76–77, 111, 133–134, 179–180,
182, 185, 187–189, 191, 193, 199,
219–220, 225, 228–229, 234,
242, 244, 297
- root (vowel)/stem modification
- 4, 58, 73–74, 110, 113–114,
128–129, 131–132, 238 *see also*
- ablaut, implicit derivation,
umlaut
- Russian 2, 10–11, 13, 169–193,
218, 227, 289, 291, 294–299
- S
- salience/salient 5, 61, 68,
111–112, 131, 184, 215, 264, 285,
289, 299
- suffixed
- adjective 40, 47–48, 54,
62, 67, 74, 78, 130, 175, 181,
204 *see also* adjective/
adjectival suffix
- noun 12–13, 294, 60, 62–65,
67–69, 71, 75–76, 79–80,
127, 141, 143, 161, 197,
201–204, 294
- verb 12–13, 38–39, 41, 43,
46–47, 54, 62, 69, 78, 122,
128–129, 131, 150, 155, 160,
203, 205
- synthetic compound 116, 132,
170–171, 178, 266, 280
- T
- transparency/transparent 2–4,
6, 8–9, 11, 13–14, 22, 26, 29–33,
35–36, 39–40, 42–43, 45–48,
55, 61, 68–69, 95, 104–105,
129, 133–135, 138, 186, 190–191,
193, 217–220, 232, 234, 238,
252, 256–257, 259–260,
266–267, 284–285, 289–296,
298–300
- Turkish 10–11, 14, 263–285,
289–292, 296–298
- U
- umlaut 104, 110, 132 *see also*
ablaut
- W
- word class 4, 13, 14, 54, 57, 70,
87–88, 93–94, 98, 104, 110–111,
143, 149, 171, 177, 184, 215,
255–256, 263, 265, 294–296
see also part of speech
- word formation 3, 5–6, 8,
10–11, 14, 36, 57, 87, 95, 103, 110,
114, 116, 132–134, 142–143, 169,
171, 178, 184, 193, 197–198, 203,
206, 213, 217, 265–267, 270–
271, 273, 275, 281, 284–285,
289–290, 293, 297, 299–300
- V
- verb(al)
- base 26, 60, 117, 132, 226,
228, 278 *see also* deverbal
derivation/formation 13–14,
55, 57, 92, 101, 112–115,
120–122, 128–129, 132–133,
137, 155, 170, 176, 182–184,
187, 190, 192, 197, 200,
214–215, 217–218, 224, 226,
232–235, 256, 263, 265, 268,
271, 273–274, 284, 291–292,
294–295, 297–298, 300
- prefix(ation) 12–13, 21, 36,
39, 47, 57, 78, 92, 112, 115,
131, 134, 142, 144, 151–152,
160–161, 171, 183–184, 210,
212, 214–215, 295 *see also*
prefixed verb
- suffix(ation) 2, 56–57, 92,
113, 122, 150, 183, 219–220,
224–225, 226, 228–230,
234, 256, 299 *see also*
suffixed verb
- verbalization/verbalizing suffix
- 25, 38–39, 42, 47, 219
- Z
- zero-derivation 4, 53, 55, 87, 92,
104, 112, 219, 221, 290, 296 *see*
also conversion
- zero-suffixation 170, 178, 185
see also conversion

This book offers the first systematic study of the early phases in the acquisition of derivational morphology from a cross-linguistic and typological perspective.

It presents ten empirical longitudinal studies in genealogically and typologically diverse languages (Indo-European, Finno-Ugric, Altaic) with different degrees of derivational complexity. Data collection, analysis and systematic comparison between child speech and parental child-directed speech are strictly parallel across the chapters. In order to identify the productivity of a derivational pattern, signalling the crucial developmental stage in its acquisition, the concept of the mini-paradigm criterion was applied.

Similar developmental processes can be observed in all children, independent of the language they acquire, but the children's courses of development also show obvious typological differences. This points towards an important impact of the structural properties of the specific language on emergence, use and the early course of development of derivational patterns.

ISBN 978 90 272 0982 5



9 789027 209825

JOHN BENJAMINS PUBLISHING COMPANY