



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Der frühe Erwerb der deutschen Derivationsmorphologie“

Eine korpuslinguistische Untersuchung der Kindersprache
in Interaktion mit dem elterlichen Input

verfasst von / submitted by

Mag. Sabine Sommer-Lolei MA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Doktorin der Philosophie (Dr. phil.)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

UA 792 327

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Sprachwissenschaft

Betreut von / Supervisor:

emer. o. Univ.-Prof. Dr. Wolfgang U. Dressler
o. Univ.-Prof. Mag. Dr. Franz Rainer

für all die Menschen,
die in den letzten Jahren
mit unzähligen
Wortneubildungen
überfordert wurden

...neben anderen Dingen

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS	10
TABELLENVERZEICHNIS	15
ABKÜRZUNGSVERZEICHNIS.....	19
SCHREIBKONVENTIONEN.....	22
VORWORT	23
1 EINLEITUNG.....	26
2 WORTBILDUNG UND DER ERWERB DER MORPHOLOGIE.....	32
2.1 WORTBILDUNGSARTEN IM DEUTSCHEN.....	32
2.1.1 <i>Komposition</i>	33
2.1.2 <i>Derivation</i>	34
2.2 WORTARTWECHSEL DURCH WORTBILDUNG	37
<i>Wortartbewahrende und wortartverändernde Wortbildungsprozesse</i>	38
2.3 GESCHICHTE DER SPRACHERWERBSFORSCHUNG	39
2.3.1 <i>Erhebungsmethoden</i>	39
2.3.2 <i>Wortbildungsmorphologie im Erstspracherwerb des Deutschen – Forschungsstand</i>	40
2.4 SPRACHERWERBSTHEORETISCHE ANSÄTZE	43
2.4.1 <i>Morphologische Entwicklungsphasen im frühen Spracherwerb</i>	43
2.4.2 <i>Morphologische Natürlichkeitstheorie</i>	46
2.5 GRUNDLEGENDE DEFINITIONEN DER SPRACHERWERBSFORSCHUNG.....	50
2.5.1 <i>Emergenz</i>	50
2.5.2 <i>Produktivität</i>	50
2.5.3 <i>Potenzielle Produktivität in CS</i>	51
2.6 EINFLUSSFAKTOREN AUF DEN MORPHOLOGIEERWERB	51
2.6.1 <i>Frequenz</i>	52
2.6.2 <i>Positionelle Salienz</i>	54
2.6.3 <i>Prosodische Salienz</i>	55
2.6.4 <i>Morphosemantische und morphotaktische Transparenz</i>	55
2.6.5 <i>Morphologische Komplexität</i>	56
2.6.6 <i>Erwartete Zusammenhänge</i>	56
2.7 FRAGESTELLUNGEN UND ANNAHMEN	57
2.7.1 <i>Forschungsfragen</i>	57
2.7.2 <i>Hypothesen</i>	58
3 SPONTANSPRACHDATEN.....	60

3.1	LONGITUDINALE SPONTANSPRACHDATEN.....	60
3.2	PANELDATEN	62
4	METHODEN.....	65
4.1	DATENERHEBUNG SPONTANSPRACHAUFNAHMEN.....	65
4.1.1	<i>Erhebung Longitudinale Spontansprachaufnahmen</i>	<i>65</i>
4.1.2	<i>Paneldatenerhebung Spontansprachaufnahmen</i>	<i>66</i>
4.2	ELTERNINTERVIEWS.....	68
4.2.1	<i>Erstellung und Durchführung der Interviews.....</i>	<i>68</i>
4.2.2	<i>Interviewdesign.....</i>	<i>69</i>
4.2.3	<i>Interviewdaten – Erfassung und Auswertung.....</i>	<i>71</i>
4.3	OPERATIONALISIERUNG DES SOZIOÖKONOMISCHEN STATUS.....	73
4.4	TRANSKRIPTION UND MORPHOLOGISCHE CODIERUNG.....	74
4.4.1	<i>Erweiterte Datencodierung.....</i>	<i>75</i>
4.4.2	<i>Datenauswahl</i>	<i>79</i>
4.4.3	<i>Datenbereinigung</i>	<i>80</i>
4.4.4	<i>Zählung.....</i>	<i>82</i>
4.5	MINI-PARADIGMA-KRITERIUM	84
5	WORTBILDUNG IN CS UND CDS	86
5.1	VERTEILUNG DER DERIVATIONEN JE WORTKLASSE	86
5.2	VERTEILUNG DER KOMPOSITA JE WORTKLASSE	87
5.3	WORTBILDUNG DER SUBSTANTIVE IN CS UND CDS	90
5.3.1	<i>Verteilung der Wortbildungsmuster bei Substantiven</i>	<i>90</i>
5.3.2	<i>Erwerbsverlauf derivierter Substantive und synthetischer Komposita.....</i>	<i>97</i>
5.4	WORTBILDUNG DER VERBEN IN CS UND CDS	102
5.4.1	<i>Verteilung der Wortbildungsmuster bei Verben</i>	<i>102</i>
5.4.2	<i>Erwerbsverlauf derivierter Verben.....</i>	<i>110</i>
5.5	WORTBILDUNG DER ADJEKTIVE IN CS UND CDS.....	116
	<i>Verteilung der Wortbildungsmuster bei Adjektiven</i>	<i>116</i>
5.6	EMERGENZ UND POTENZIELLE PRODUKTIVITÄT	124
5.6.1	<i>Derivierte und zusammengesetzte Substantive</i>	<i>126</i>
5.6.2	<i>Derivierte Verben</i>	<i>127</i>
5.6.3	<i>Derivierte Adjektive.....</i>	<i>128</i>
5.6.4	<i>Frühe Emergenz vs. späte Meisterung.....</i>	<i>129</i>
5.7	STATISTISCHE AUSWERTUNG DER LONGITUDINALEN SPONTANSPRACHDATEN.....	130

5.7.1	<i>Chi-Quadrat-Anpassungstest</i>	130
5.7.2	<i>Zusammenfassender Überblick der Signifikanzen</i>	132
5.8	STATISTISCHE AUSWERTUNG DER PANELDATEN	133
5.8.1	<i>Unabhängige T-Tests</i>	134
5.8.2	<i>Gepaarte T-Tests</i>	138
6	ABLEITUNGSBASEN BEI DERIVATIONEN	144
6.1	BASEN ABGELEITETER SUBSTANTIVE	145
6.2	BASEN DERIVIERTER VERBEN	146
6.3	BASEN DERIVIERTER ADJEKTIVE	148
7	KOMPLEXITÄT IN DER WORTBILDUNG	150
7.1	KOMPLEXITÄTSGRADE DER DERIVATE UND KOMPOSITA.....	150
7.2	MORPHOLOGISCHE KOMPLEXITÄT IM ERWERBSVERLAUF	151
7.2.1	<i>Anteile der Wortbildungsarten je Komplexitätsgrad</i>	151
7.2.2	<i>Anteile der Wortbildungsmuster je Komplexitätsgrad</i>	153
7.2.3	<i>Morphologische Komplexität der Komposita</i>	159
7.2.4	<i>Morphologische Komplexität der Derivate</i>	166
8	FEHLERAUSWERTUNG	180
8.1	FEHLER OHNE FEHLER	180
8.2	FEHLERANALYSE	181
8.2.1	<i>Fehler bei synthetischen Komposita</i>	183
8.2.2	<i>Fehler bei derivierten Substantiven</i>	185
8.2.3	<i>Fehler bei abgeleiteten Verben</i>	186
8.2.4	<i>Fehler bei derivierten Adjektiven</i>	190
9	ELIZITATIONSEXPERIMENT	191
9.1	FORSCHUNGSZIEL UND FORSCHUNGSFRAGEN.....	191
9.2	METHODE	192
	<i>Pilotexperiment</i>	192
9.3	TEILNEHMER.....	193
9.4	STIMULI	195
9.5	TESTDURCHFÜHRUNG	197
	<i>Ablauf des Experiments</i>	197
9.6	DATENAUSWERTUNG	198
9.7	FEHLERCODIERUNG	198
9.8	ERGEBNISSE DES ELIZITATIONSEXPERIMENTS	200
9.8.1	<i>Korrekte und semikorrekte Ableitungen</i>	201
9.8.2	<i>Fehleranalyse</i>	206

9.8.3	<i>Statistische Auswertung der Ergebnisse des Experiments</i>	213
9.9	ABSCHLIEßENDE BETRACHTUNGEN UND ZUSAMMENFASSUNG	216
10	KORRELATIONEN	219
10.1	KORRELATIONEN ERGEBNISSE DERIVATIONSTEST	219
10.1.1	<i>Primäre Zusammenhänge der Ergebnisse</i>	225
10.1.2	<i>Inputmenge</i>	226
10.1.3	<i>Neue Medien – YouTube, Internetnutzung, Smartphone, Fernsehen</i>	226
10.1.4	<i>Hörbücher und Radio</i>	226
10.1.5	<i>Büchervorlesen</i>	226
10.2	KORRELATIONEN ZU DEN SPONTANSPRACHDATEN	227
10.2.1	<i>Primäre Zusammenhänge</i>	231
10.2.2	<i>Inputmenge</i>	231
10.2.3	<i>Neue Medien – YouTube, Internetnutzung, Smartphone, Fernsehen</i>	232
10.2.4	<i>Hörbücher und Radio</i>	232
10.2.5	<i>Büchervorlesen</i>	232
10.2.6	<i>Elizitationsexperiment vs. Spontansprache</i>	233
11	KONKLUSION	234
12	AUSBLICK	240
13	LITERATURVERZEICHNIS	243
ANHANG 1 TRANSLITERATIONS- UND CODIERUNGSBEISPIELE ..		254
A1.1	AUSZUG AUS EINEM TRANSKRIPT (CHAT)	254
A1.2	BEISPIEL MORPHOLOGISCHE CODIERUNG (MOR)	254
A1.3	CODIERUNGSBEISPIEL ZITAT (CIT)	254
A1.4	CODIERUNGSBEISPIEL IMITATION (IMI)	255
ANHANG 2 CODIERUNGSSCHEMATA		255
A2.1	BASISCODIERUNGEN DER SPONTANSPRACHDATEN	255
A2.2	ZUSATZCODIERUNGEN DER SPONTANSPRACHDATEN	256
A2.3	CODIERUNGSSCHEMA ELTERNINTERVIEWS	258
ANHANG 3 KORPORA – GESAMTÜBERSICHTEN		263
A3.1	JAN-KORPUS	263
A3.2	KAT-KORPUS	264
A3.3	LEN-KORPUS.....	264
A3.4	INPUT-KORPUS.....	265
ANHANG 4 VERTEILUNG DER WORTBILDUNGSMUSTER		268
A4.1	NOMINALE WORTBILDUNGSMUSTER IN CS UND CDS.....	268

A4.2	VERBALE WORTBILDUNGSMUSTER IN CS UND CDS	270
A4.3	ADJEKTIVISCHE WORTBILDUNGSMUSTER IN CS UND CDS.....	271
ANHANG 5 DERIVATIONEN IN CS UND CDS – ÜBERSICHTEN		272
A5.1	JAN-KORPUS – N, V, ADJ DERIV JE MONAT	272
A5.2	KAT-KORPUS – N, V, ADJ DERIV JE MONAT	273
A5.3	LEN-KORPUS – N, V, ADJ DERIV JE MONAT	274
A5.4	INPUT-KORPUS – N, V, ADJ DERIV JE DATENZEITPUNKT	275
ANHANG 6 KOMPOSITA IN CS UND CDS – ÜBERSICHTEN.....		280
A6.1	JAN-KORPUS.....	280
A6.2	KAT-KORPUS.....	281
A6.3	LEN-KORPUS	282
A6.4	INPUT-KORPUS	283
ANHANG 7 ABLEITUNGSBASEN		283
A7.1	BASEN DERIVIERTER SUBSTANTIVE (TOKENS).....	283
A7.2	BASEN DERIVIERTER VERBEN (TOKENS)	284
A7.3	BASEN DERIVIERTER ADJEKTIVE (TOKENS).....	284
ANHANG 8 KOMPLEXITÄT DER WORTBILDUNGSMUSTER		285
A8.1	NOMINALE WORTBILDUNGSMUSTER IN CS UND CDS (TOKENS).....	285
A8.2	VERBALE WORTBILDUNGSMUSTER IN CS UND CDS (TOKENS).....	286
A8.3	ADJEKTIVISCHE WORTBILDUNGSMUSTER IN CS UND CDS (TOKENS) ..	287
ANHANG 9 KOMPLEXITÄT DER WORTBILDUNGSPRODUKTE.....		288
A9.1	JAN-KORPUS.....	288
A9.2	KAT-KORPUS UND LEN-KORPUS	290
A9.3	INPUT-KORPUS	291
ANHANG 10 ELIZITATIONSEXPERIMENT		292
A10.1	EXPERIMENTALBLATT	292
A10.2	CODIERUNGSSCHEMA FÜR DIE DATENERFASSUNG.....	295
A10.3	FREQUENZEN SEMIKORREKTER DERIVATE.....	296
ANHANG 11 PANELDATEN (KINDERGARTEN).....		296
ANHANG 12 ZUSAMMENFASSUNG.....		298
ANHANG 13 ABSTRACT		299

Abbildungsverzeichnis

Abbildung 1. Prozentuale Verteilung der nominalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS bei JAN, KAT und LEN (gesamter Zeitraum).....	91
Abbildung 2. Prozentuale Verteilung der nominalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS im gesamten INPUT-Korpus (HSES und LSES)	92
Abbildung 3. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum (1;3–4;11) nach absteigender Häufigkeit in CDS	94
Abbildung 4. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS.....	95
Abbildung 5. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS.....	96
Abbildung 6. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita in Lemmas (1;3–4;11)	97
Abbildung 7. JAN-Korpus CDS: Monatlicher Verlauf derivierter Substantive und synthetischer Komposita in Lemmas (1;3–4;11).....	97
Abbildung 8. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita in Tokens (1;3–4;11)	98
Abbildung 9. JAN-Korpus CDS: Monatlicher Verlauf derivierter Substantive und synthetischer Komposita in Tokens (1;3–4;11).....	98
Abbildung 10. INPUT-Korpus CS: Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita nach Erhebungszeitpunkt (DP1–4) in Lemmas	100
Abbildung 11. INPUT-Korpus CDS: Verlauf derivierter Substantive und synthetischer Komposita nach Erhebungszeitpunkt (DP1–4) in Lemmas ..	100
Abbildung 12. INPUT-Korpus CS: Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita nach Erhebungszeitpunkt (DP1–4) in Tokens	101
Abbildung 13. INPUT-Korpus CDS: Verlauf derivierter Substantive und synthetischer Komposita nach Erhebungszeitpunkt (DP1–4) in Tokens	101
Abbildung 14. Prozentuale Verteilung verbaler Wortbildungsmuster (Lemmas und Tokens) in CS und CDS in den Spontansprachkorpora von JAN, KAT und LEN (gesamter Zeitraum).....	103

Abbildung 15. Prozentuale Verteilung der verbalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS im gesamten INPUT-Korpus (HSES und LSES).....	104
Abbildung 16. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum (1;3–4;11) nach absteigender Häufigkeit in CDS	107
Abbildung 17. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS	108
Abbildung 18. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS	109
Abbildung 19. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Verben (V DERIV) vs. Simplex-Verben (V SIMP) in Lemmas (1;3–4;11)	110
Abbildung 20. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Verben (V DERIV) vs. Simplex-Verben (V SIMP) in Tokens (1;3–4;11)	110
Abbildung 21. JAN-Korpus CS: Monatlicher Erwerbsverlauf von Präfixverben (PREFV), Partikelverben (PTLV) und Partikel (PTL) in Lemmas (1;3–4;11)	112
Abbildung 22. JAN-Korpus CDS: Verlauf derivierter Verben (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) in Lemmas (1;3–4;11)	112
Abbildung 23. JAN-Korpus CS: Monatlicher Erwerbsverlauf von Präfixverben (PREFV), Partikelverben (PTLV) und Partikel (PTL) in Tokens (1;3–4;11)	113
Abbildung 24. JAN-Korpus CDS: Verlauf derivierter Verben (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) in Tokens (1;3–4;11)	113
Abbildung 25. INPUT-Korpus CS: Erwerbsverlauf derivierter Verben gesamt (V DERIV) vs. Simplex-Verben (V SIMP), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Lemmas	113
Abbildung 26. INPUT-Korpus CDS: Verlauf derivierter Verben gesamt (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Lemmas	114
Abbildung 27. INPUT-Korpus CS: Erwerbsverlauf derivierter Verben gesamt (V DERIV) vs. Simplex-Verben (V SIMP), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Tokens	114
Abbildung 28. INPUT-Korpus CDS: Verlauf derivierter Verben gesamt (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Tokens	114
Abbildung 29. Prozentuale Verteilung der Wortbildungsmuster (Lemmas und Tokens) bei Adjektiven in CS und CDS in den Spontansprachkorpora von JAN, KAT und LEN (gesamter Zeitraum).....	117

Abbildung 30. Prozentuale Verteilung der Wortbildungsmuster (Lemmas und Tokens) bei Adjektiven in CS und CDS im gesamten INPUT-Korpus (HSES und LSES)	118
Abbildung 31. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum (1;3–4;11) nach absteigender Häufigkeit in CDS	121
Abbildung 32. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS.....	122
Abbildung 33. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS.....	123
Abbildung 34. Ableitungsbasen derivierter Substantive (Lemmas) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS	146
Abbildung 35. Ableitungsbasen derivierter Verben (Lemmas) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS	147
Abbildung 36. Ableitungsbasen derivierter Adjektive (Lemmas) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS	149
Abbildung 37. Prozentuale Anteile nominaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	153
Abbildung 38. Prozentuale Anteile nominaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	154
Abbildung 39. Prozentuale Verteilung verbaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	155
Abbildung 40. Prozentuale Verteilung verbaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	155
Abbildung 41. Prozentuale Verteilung adjektivischer Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	157
Abbildung 42. Prozentuale Verteilung adjektivischer Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	158
Abbildung 43. JAN-Korpus: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS	160
Abbildung 44. KAT- und LEN-Korpora: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS	160
Abbildung 45. INPUT-Korpus: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS	161

Abbildung 46. JAN- und INPUT-Korpus: Komplexitätsgrade von Adjektivkomposita (Lemmas) in CS und CDS.....	166
Abbildung 47. JAN-Korpus: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS	167
Abbildung 48. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS.....	167
Abbildung 49. INPUT-Korpus: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS	168
Abbildung 50. JAN-Korpus: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS.....	172
Abbildung 51. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS	173
Abbildung 52. INPUT-Korpus: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS	173
Abbildung 53. JAN-Korpus: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS	176
Abbildung 54. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS.....	177
Abbildung 55. INPUT-Korpus: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS	177
Abbildung 56. Fehlerhäufigkeiten je Fehlercode bei synthetischen Komposita in CS im JAN- und INPUT-Korpus	183
Abbildung 57. Fehlerhäufigkeiten je Fehlercode bei derivierten Substantiven in CS im JAN- und INPUT-Korpus	185
Abbildung 58. Fehlerhäufigkeiten je Fehlercode bei derivierten Verben in CS im JAN- und INPUT-Korpus	186
Abbildung 59. Häufigkeiten (Mittelwerte) der korrekten und semikorrekten Ableitungen im Elizitationsexperiment bei HSES und LSES	204
Abbildung 60. Häufigkeiten (Mittelwerte) der Wortbildungsmuster in semikorrekten Derivationen bei HSES und LSES je Wortbildungskategorie	205
Abbildung 61. Prozentuale Fehlerhäufigkeiten und fehlende Antworten in den einzelnen Wortbildungskategorien gemessen an der Gesamtanzahl	206
Abbildung 62. Mittelwerte der Fehlerhäufigkeiten semikorrekter und inkorrektor Bildungen, sowie der fehlenden Antworten (schraffiert) im Elizitationsexperiment bei HSES und LSES je Wortbildungskategorie.....	207
Abbildung 63. Verteilung (Mittelwerte) der verwendeten Ableitungsbasen in semi- und inkorrekten Derivationen bei HSES und LSES je Wortbildungskategorie.....	212
Abbildung 64. Ableitungsbasen derivierter Substantive (Tokens) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS....	283

Abbildung 65. Ableitungsbasen derivierter Verben (Tokens) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS	284
Abbildung 66. Ableitungsbasen derivierter Adjektive (Tokens) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS	284
Abbildung 67. Prozentuale Verteilung nominaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	285
Abbildung 68. Prozentuale Verteilung nominaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	285
Abbildung 69. Prozentuale Verteilung verbaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	286
Abbildung 70. Prozentuale Verteilung verbaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	286
Abbildung 71. Prozentuale Verteilung adjektivischer Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	287
Abbildung 72. Prozentuale Verteilung adjektivischer Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4).....	287
Abbildung 73. JAN-Korpus: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS.....	288
Abbildung 74. JAN-Korpus: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS	288
Abbildung 75. JAN-Korpus: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS	289
Abbildung 76. JAN-Korpus: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS	289
Abbildung 77. JAN- und INPUT-Korpus: Komplexitätsgrade von Adjektivkomposita (Tokens) in CS und CDS	289
Abbildung 78. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS	290
Abbildung 79. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS.....	290
Abbildung 80. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS	290
Abbildung 81. KAT- und LEN-Korpora: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS	291
Abbildung 82. INPUT-Korpus: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS.....	291

Abbildung 83. INPUT-Korpus: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS	291
Abbildung 84. INPUT-Korpus: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS	292
Abbildung 85. INPUT-Korpus: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS	292
Abbildung 86. Experimentalblatt mit detaillierter Anleitung für den Experimentator zum Elizitationstest nominaler Ableitungen, Seite 1	293
Abbildung 87. Experimentalblatt zum Elizitationstest nominaler Ableitungen, Seite 2.....	294

Tabellenverzeichnis

Tabelle 1. Derivationsmuster im Deutschen – Überblick	37
Tabelle 2. Longitudinale Spontansprachdaten (CS und CDS).....	61
Tabelle 3. Paneldaten zu Hause CS und CDS (INPUT-Korpus)	63
Tabelle 4. Beispiel einer Äußerung nach dem Import in Excel	76
Tabelle 5. Morphologische Komplexitätsgrade der Derivate und Komposita.....	78
Tabelle 6. Beispiele für fehlerhafte morphologische Codierungen in den Korpora	80
Tabelle 7. Frequenzen (Lemmas, Tokens und Prozente) derivierter Substantive (ohne DIM), Verben und Adjektive in Relation zur Gesamtzahl an Substantiven, Verben und Adjektiven in CS und CDS in allen Korpora im gesamten Zeitraum	86
Tabelle 8. Frequenzen (Lemmas, Tokens und Prozente) nominaler Komposita (N KOMP), synthetischer Komposita (synth. KOMP) und Adjektivkomposita (ADJ KOMP) in Relation zur Gesamtzahl an Substantiven und Adjektiven in CS und CDS in allen Korpora	88
Tabelle 9. Emergenzen (Em.) und Zeitpunkte der Erfüllung des Mini- Paradigma-Kriteriums (MPK) der relevantesten Derivationsmuster bei Substantiven, Verben und Adjektiven in CS im gesamten Zeitraum	125
Tabelle 10. Überblick Signifikanzen der Chi-Quadrat-Anpassungstests in den Korpora JAN, KAT und LEN: Simplex- vs. derivierte/zusammengesetzte Lexeme (Lemmas, Tokens) in CS vs. CDS im gesamten Zeitraum je Wortklasse.....	132
Tabelle 11. Mittelwertdifferenzen in CS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)	135
Tabelle 12. Mittelwertdifferenzen in CS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Tokens)	135

Tabelle 13. Mittelwertdifferenzen in CDS nach SES-Zugehörigkeit, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas)	137
Tabelle 14. Mittelwertdifferenzen in CDS nach SES-Zugehörigkeit, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Tokens)	137
Tabelle 15. Gruppenvergleiche HSES, CS vs. CDS, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas und Tokens), gesamter Zeitraum.....	139
Tabelle 16. Gruppenvergleiche LSES, CS vs. CDS, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas und Tokens), gesamter Zeitraum.....	140
Tabelle 17. Mittelwertdifferenzen zweier Datenzeitpunkte in HSES-CS, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas)	142
Tabelle 18. Mittelwertdifferenzen zweier Datenzeitpunkte in LSES-CS, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas)	142
Tabelle 19. Mittelwertdifferenzen zweier Datenzeitpunkte in HSES-CDS, Signifikanzen (<i>p</i> , 2-seitig) und Effektstärken (<i>r</i>) der Derivationen und Komposita je Wortklasse (Lemmas)	143
Tabelle 20. Mittelwertdifferenzen zweier Datenzeitpunkte in LSES-CDS und Signifikanzen (<i>p</i> , 2-seitig) der Derivationen und Komposita je Wortklasse (Lemmas).....	143
Tabelle 21. Mittelwerte aller Korpora (in Prozent) der Komplexitätsgrade nominaler Komposita, derivierter Substantive, Verben und Adjektive in CS und CDS (Altersspanne 2;9–3;3).....	152
Tabelle 22. Prozentualer Anteil mehrgliedriger Substantivkomposita in Relation zur Gesamtanzahl nominaler Komposita in CS und CDS in allen Korpora	163
Tabelle 23. Übersicht der Fehlerhäufigkeiten derivierter und zusammengesetzter Lexeme in Lemmas, Tokens und Prozenten in CS in allen Korpora im gesamten Zeitraum	181
Tabelle 24. Übersicht der Teilnehmer des Pilotexperimentes und zusammenfassende Ergebnisse.....	193
Tabelle 25. Teilnehmer Derivationsexperiment	194
Tabelle 26. Übersicht der Stimuli und Targets nach Ableitungssuffix, Um-/Ablaut, Basis und semantischer Wortbildungskategorie.....	196
Tabelle 27. Übungssätze zur Vorbereitung auf das Elizitationsexperiment.....	197
Tabelle 28. Fehlercodierung für semikorrekte und inkorrekte Derivationen mit Beispielen	199
Tabelle 29. Überblick über die Ergebnisse des Elizitationsexperiments je Teilnehmer.....	201

Tabelle 30. Individuelle und Gesamthäufigkeiten der korrekten Derivationen im Elizitationsexperiment in den Wortbildungskategorien nach SES-Zugehörigkeit.....	202
Tabelle 31. Gesamte und individuelle Fehlerhäufigkeiten je Fehlercode bei HSES und LSES.....	209
Tabelle 32. Übersicht und Häufigkeiten der verwendeten Suffixe in semi- und inkorrekten Ableitungen bei HSES und LSES mit Beispielen	211
Tabelle 33. Mittelwertdifferenzen nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) verschiedener Variablen des Elizitationsexperiments.....	214
Tabelle 34. Mittelwertdifferenzen nach Geschlecht und Signifikanzen (p , 2-seitig) verschiedener Variablen des Elizitationsexperiments.....	215
Tabelle 35. Korrelationen Ergebnisse des Derivationsexperiments – INPUT....	221
Tabelle 36. Korrelationen Ergebnisse des Derivationsexperiments – YISF.....	222
Tabelle 37. Korrelationen Ergebnisse des Derivationsexperiments – HBR	223
Tabelle 38. Korrelationen Ergebnisse des Derivationsexperiments – Büchervorlesen.....	224
Tabelle 39. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): Verbale Derivationen – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)	229
Tabelle 40. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): derivierte und zusammengesetzte Adjektive – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)	229
Tabelle 41. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): Nominale Derivationen und Komposita – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)	230
Tabelle 42. Übersicht über die wesentlichen Basiscodierungen der Spontansprachdaten	255
Tabelle 43. Übersicht über die erforderlichen Zusatzcodierungen für die Analyse der Derivationen und Komposita, sowie für die Fehlerauswertung, mit Beispielen.....	256
Tabelle 44. Auszug aus dem Codierungsschema zur Auswertung der Elterninterviews mit dem Fokus auf relevante Variablen für den Spracherwerb	258
Tabelle 45. JAN-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS).....	263
Tabelle 46. KAT-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS).....	264

Tabelle 47. LEN-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS).....	264
Tabelle 48. INPUT-Korpus HSES - Übersicht über Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) je Kind und Erhebungszeitpunkt (DP 1–4) im Input (CDS) und Output (CS).....	265
Tabelle 49. INPUT-Korpus LSES - Übersicht über Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) je Kind und Erhebungszeitpunkt (1–4) im Input (CDS) und Output (CS).....	267
Tabelle 50. Frequenzen in Lemmas und Tokens und prozentuale Verteilung der Wortbildungsmuster der Substantive in CS und CDS in allen Korpora im gesamten Zeitraum	268
Tabelle 51. Frequenzen in Lemmas und Tokens, und prozentuale Verteilung der Wortbildungsmuster der Verben in CS und CDS in allen Korpora im gesamten Zeitraum	270
Tabelle 52. Frequenzen in Lemmas und Tokens, und prozentuale Verteilung der Wortbildungsmuster der Adjektive in CS und CDS in allen Korpora im gesamten Zeitraum	271
Tabelle 53. JAN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS	272
Tabelle 54. KAT-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS	273
Tabelle 55. LEN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS	274
Tabelle 56. INPUT-Korpus - Gesamtübersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Erhebungszeitpunkt (1–4) in CS und CDS (HSES und LSES)	275
Tabelle 57. INPUT-Korpus HSES - Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Kind und Erhebungszeitpunkt (1–4) in CS und CDS.....	275
Tabelle 58. INPUT-Korpus LSES - Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Kind und Erhebungszeitpunkt (1–4) in CS und CDS.....	278

Tabelle 59. JAN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS	280
Tabelle 60. KAT-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS	281
Tabelle 61. LEN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS	282
Tabelle 62. INPUT-Korpus - Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita je Erhebungszeitpunkt (1–4) in CS und CDS (HSES und LSES)	283
Tabelle 63. Codierungsschema für die Datenerfassung des Derivationsexperiments mit Beispielen	295
Tabelle 64. Individuelle und Gesamthäufigkeiten der semikorrekten Ableitungen im Elizitationsexperiment in den Wortbildungskategorien nach SES-Zugehörigkeit	296
Tabelle 65. Paneldaten im Kindergarten CS und CDS (INPUT-Korpus).....	297

Abkürzungsverzeichnis

1	erste Person
2	zweite Person
3	dritte Person
AA	Adjektiv-Adjektiv Kompositum
ACC	<i>accusative</i> , Akkusativ
ADJ	Adjektiv
ADJ KOMP	Adjektivkompositum
ADS	<i>adult-directed speech</i> , an Erwachsene gerichtete Sprache
ADV	Adverb
Anz.	Anzahl
art:def	definiter Artikel
CDS	<i>child-directed speech</i> , an das Kind gerichtete Sprache; Input des Kindes
CHAT	<i>Codes for the Human Analysis of Transcripts</i> , Codes für die menschliche Analyse von Transkripten
CHI	(<i>Target</i>) <i>child</i> , Zielkind, Fokuskind

CHILDES	<i>Child Language Data Exchange System</i> , Datenaustauschsystem für Kindersprache
CIT	<i>citation</i> , Zitat
CLAN	<i>Computerized Language Analysis</i> , Computergestützte Sprachanalyse
CONJ	<i>conjunction</i> , Konjunktion
CS	<i>child speech</i> , Sprache des Kindes
CSV	<i>comma-separated values</i> , kommasetrennte Werte
DERIV	Derivat, deriviert
DET	Determinator, Determinierer, Artikel
df	<i>degrees of freedom</i> , Freiheitsgrade
DIM	Diminutiv
DP	<i>data point</i> , Daten(zeit)punkt
Em.	Emergenz, erstes Auftreten
Err.	Error
FSUFF	Fremdsuffix
GEN	Genitiv
HBP	Hauptbezugsperson, jene erwachsene Person, die am meisten Zeit mit dem Fokuskind verbringt
HBR	Hörbuch und Radio
HSES, H	<i>high(er) SES</i> , hoher/höherer sozioökonomischer Status
IMI	Imitation
IMPL	Implizit(e Ableitung)
INF	Infinitiv
INPUT	<i>Investigating Parental and Other Caretakers' Utterances to Kindergarten Children</i>
INTF/intf	Interfix
KI	Konfidenzintervall
KOMP	Kompositum, Komposita
KONV	Konversion (Nullderivation)
KONV INF	Infinitivkonversion, syntaktische Konversion
KONV mor	morphologische Konversion
L1	Erstsprache (in dieser Arbeit: Österreichisches Deutsch)
L2	Zweitsprache (in dieser Arbeit: Österreichisches Deutsch)

Lem.	Lemmas
LSES, L	<i>low(er) SES</i> , niedriger(er) sozioökonomischer Status
m	männlich (Geschlecht)
M	<i>mean</i> , Mittelwert
MLU	<i>mean length of utterance</i> , mittlere Äußerungslänge
N	Substantiv
NA	Substantiv-Adjektiv Kompositum
NAC, N.Ac.	Nomen Acti/Actionis
NAf	Nominalaffix
NAG, N.Ag.	Nomen Agentis
NIN, N.In.	Nomen Instrumenti
NLOC	Nomen Loci
NN	Substantiv-Substantiv Kompositum
NQU	Nomen Qualitatis
n. sig.	nicht signifikant
NUM	Zahlwort
NUMA	Zahlwort-Adjektiv Kompositum
PDS	<i>pet-directed speech</i> , an Haustiere gerichtete Sprache
PL, P	Plural
PP	Perfektpartizip (Partizip II)
PPräs	Präsenspartizip (Partizip I)
PREF	<i>prefix</i> , Präfix, untrennbar
PREFV	Präfixverb (mit untrennbarem Präfix)
PREP, P	<i>preposition</i> , Präposition
PRO	Pronomen
PTL	Präfix, trennbar (= Partikel)
PTLV	Partikelverb (mit trennbarem Präfix)
SD	<i>Standard deviation</i> , Standardabweichung
SES	<i>socioeconomic status</i> , sozioökonomischer Status
SG, S	Singular
Sig., sig.	Signifikanz, signifikant
SIMP	Simplex, Simplizia
SUFF, suff	Suffix
synth. KOMP	synthetisches Kompositum

Tok.	Tokens
w	weiblich (Geschlecht)
V	Verb (Vollverb)
VN	Verb-Substantiv Kompositum
VSt	Verbstamm
YISF	YouTube, Internetnutzung, Smartphone, Fernsehen
ZIRK	Zirkumfix

Schreibkonventionen

<i>kursiv</i>	Fremdwörter; Beispiele zur Hervorhebung im Text	<i>language</i>
,...‘	einfache Anführungszeichen beinhalten Glossen	‚sag-2SG‘
GLOSSE	Glossen werden in Kapitälchen dargestellt	‚sag-3PL‘
→ / ←	Pfeile geben die Ableitungsrichtung an	rufen → Ruf
+	markiert die Kompositionsfuge	Holz+haus
#	markiert trennbare Verbpräfixe (Partikel)	an#halt-en
-	markiert die Morphemgrenze(n)	ver-such-en
∅	markiert ein Nullmorphem	zu#∅macht
ˈ	markiert den Hauptakzent eines Lexems	Ún-tier
˘	markiert den Nebenakzent eines Lexems	Háus+tiér
*	markiert im Fließtext ein inexistentes Wort	*bahnig

Vorwort

Allen voran möchte ich mich bei der Österreichischen Akademie der Wissenschaften für die Zuerkennung eines DOC-team Stipendiums herzlich bedanken, ohne deren finanzielle Unterstützung die Umsetzung meines Projekts nicht möglich gewesen wäre, sowie bei allen Befürwortern und Unterstützern in diesem Zusammenhang. Ich möchte mich auch für die wunderbare Zusammenarbeit mit Eva Gutknecht und Joanna Kölbl bedanken, die mir bei allen Fragen und Problemen hilfreich und mit viel Verständnis zur Seite standen, sowie bei Barbara Haberl, der Leiterin der Abteilung für Stipendien und Preise für ihre Unterstützung.

Außerdem möchte ich mich bei Alexandra N. Lenz bedanken, Direktorin des Austrian Centre for Digital Humanities and Cultural Heritage (ACDH-CH) der Österreichischen Akademie der Wissenschaften, sowie bei ihrem Vorgänger Karlheinz Mörth für die Ansiedlung meines Dissertationsprojektes an ihrem Institut und in ihrer Forschungsabteilung Variation und Wandel des Deutschen in Österreich.

Überdies danke ich Peter Schneider von der Forschungsgruppe Musik und Gehirn an der Neurologischen Universitätsklinik Heidelberg für seine Unterstützung und die Möglichkeit eines Auslands- und Forschungsaufenthalts an der Abteilung für Neuroradiologie.

Ebenfalls bedanke ich mich sehr herzlich bei meinem großartigen Betreuer und langjährigen Mentor Wolfgang U. Dressler, der immer für meine Anliegen offen war und ist und mich in diesem Entstehungsprozess durchgehend begleitet hat. Sein beständiger Einsatz und seine immerwährende Unterstützung haben zum Gelingen dieser Arbeit beigetragen. Insbesondere danke ich ihm für die Bereitstellung der Daten des INPUT-Projekts für die Auswertung.

Meinem Zweitbetreuer Franz Rainer danke ich sehr herzlich für seine Bereitschaft die Begutachtung meiner Arbeit zu übernehmen. Außerdem bin ich ihm zu Dank verpflichtet für seine wertvollen Kommentare und seine veränderte Sichtweise auf

die Dinge. Seine Vorschläge und Anmerkungen haben meine Arbeit noch weiter verbessert und in bestimmten Bereichen vertieft.

Bei meinen lieben Kolleginnen Veronika Mattes und Katharina Korecky-Kröll bedanke ich mich für die intensive Zusammenarbeit in den letzten Jahren und für ihre raschen Antworten auf meine oft spontanen Eingebungen. Meinem Kollegen Markus Christiner danke ich für sein offenes Ohr, dass zu jeder Zeit für all meine Fragen offenstand, seine Geduld und seine Unterstützung vor allem in statistischen Belangen.

Mein besonderer Dank gilt allen Müttern und Vätern, sowie ihren Familien, für die jahrelangen Aufnahmen ihrer Kinder. Zunächst geht mein herzlicher Dank an Andrea Kunschert (Lenas Mutter), Katharina Korecky-Kröll (Jans Mutter), Kathis Mutter, deren Name mir leider nicht bekannt ist, sowie Brigitta Müller (Kathis Experimentatorin) für das regelmäßige Erstellen der longitudinalen spontansprachlichen Aufnahmen, die über Jahre hinweg in wechselnden Situationen durchgeführt wurden. Ebenso bedanken möchte ich mich bei Sabine Laaha, die Lenas Mutter zur Datenerhebung motiviert und die Transkription der Aufnahmen organisiert hat. Ich danke allen Eltern des INPUT-Projekts, die es uns ermöglicht haben ihre Kinder auf Audio und Video aufzunehmen. In besonderer Weise danke ich allen Kindern für ihre Gesprächigkeit, denn ohne Euch wäre diese Arbeit nicht möglich gewesen.

Sämtlichen an dem langen Entstehungsprozess von der Aufnahme über Transliteration, Erstellung des Lexikons, morphologische Codierung bis hin zur Disambiguierung der Daten, beteiligten Kollegen und Kolleginnen, Studierenden, Mitwirkenden und Unterstützern dieser umfangreichen longitudinalen Studien und Paneldatenerhebungen gebührt mein aufrichtiger Dank.

Ich bin sehr dankbar, dass mir die longitudinalen Korpora und Paneldaten für die Auswertung der Derivationen und Komposita zur Verfügung gestellt wurden.

Des Weiteren danke ich Paul C. Korecky für das Programmieren und die kostenfreie Bereitstellung der CLANTOCSV-Software, und die damit verbundene Vereinfachung der Lemmasuche und Auswertung.

Überdies möchte ich meinen lieben Kolleginnen Marina Camber, Maria Weichselbaum, Lisa Buchegger und Gülsah Hizli für die Durchführung des Elizitationsexperimentes mit den L1- und L2-Kindern recht herzlich danken, sowie allen Kindern und ihren Eltern für die Bereitschaft und das Einverständnis bei diesem Folgeprojekt des INPUT-Projekts erneut mitzuwirken.

Ebenfalls möchte ich der Direktion der Volksschule OVS Zeltgasse im 8. Wiener Gemeindebezirk recht herzlich für die Erlaubnis danken, das Pilotexperiment zur Elizitation nominaler Ableitungen an ihrer Schule durchführen zu dürfen, sowie den Lehrern, die dies in ihren Klassen ermöglicht haben. Insbesondere gebührt mein herzlicher Dank allen Kindern, die mit so viel Begeisterung mitgemacht haben. Überdies danke ich meinen lieben Nachbarskindern für ihr reges Interesse an meiner Arbeit und ihre Bereitschaft an meiner Pilotstudie aktiv mitzuwirken.

Ich danke auch dem DoktorandInnenzentrum der Universität Wien für die äußerst lehrreiche Schreibwoche, die es mir ermöglicht hat, fokussierter und konzentrierter zu arbeiten.

Mein innigster Dank gilt jedoch meiner Familie, die mich wie schon zuvor jahrelang auf vielfältige Weise unterstützt und motiviert hat. Allen voran danke ich meinem Sohn Ralph für sein großartiges Selbst und für die Musik, die er in mein Leben bringt zusammen mit immer neuen Perspektivenwechseln. Meinem Lebensmenschen Markus, der mich angespornt und mich geduldig durch den analytischen Teil meiner Arbeit begleitet hat, danke ich in besonderer Weise. Ohne Eure großartige Unterstützung in allen Bereichen meines Lebens wäre es mir nicht möglich gewesen, diese Arbeit und die damit einhergehenden Herausforderungen zu bewältigen. Ich danke Euch von ganzem Herzen.

1 Einleitung

Die Grundlagenforschung der Wortbildung im Erstspracherwerb des Deutschen steht im Mittelpunkt dieser Arbeit. Bereits vorhandene Arbeiten zur Derivation und Komposition sollen durch meine Analysen erweitert werden und den Grundstein für nachfolgende Forschungen legen.

Die Wortbildung ist für den Wortschatzerwerb bzw. dessen -erweiterung essenziell und stellt Kinder, die gerade ihre Erstsprache erwerben, vor eine Herausforderung. Hauptsächlich dient die Wortbildung der Bildung neuer, komplexer Lexeme auf Basis bereits bekannter Wörter und erweitert auf diese Weise den Wortschatz, auch der Kinder. Komplexe Lexeme bestehen aus zwei oder mehr morphologischen Elementen und sind charakterisiert durch die Anzahl der zu ihrer Bildung nötigen Wortbildungsmuster. Ihre interne Struktur und damit die zugrundeliegenden Muster und Gesetzmäßigkeiten sind Gegenstand der Wortbildungsforschung.

Der Erstspracherwerb von Kindern vollzieht sich über einen langen Zeitraum und ist von vielen ineinandergreifenden Mechanismen abhängig. Er besteht aus verbaler und nonverbaler Kommunikation und verläuft in großem Maße individuell. Nur wenige Studien untersuchen die Kindersprache in Relation zu ihrem Input (siehe Dressler et al. 2017; Dressler, Sommer-Lolei, et al. 2019; Mattes et al. 2021; Sommer-Lolei & Dressler akzeptiert). Speziell in den frühen Phasen des Spracherwerbs unterscheidet sich die Sprache der Mutter, je nachdem ob diese an das Kind (*child-directed speech*, CDS) oder an Erwachsene gerichtet ist (*adult-directed speech*, ADS).

Dabei selektiert das Kind, welche Muster es vom Input aufnimmt, eine Weile verfolgt, wieder verwirft oder beibehält, denn nicht alles vom Input wird auch in gleicher Weise vom Kind übernommen. In der sprachwissenschaftlichen Forschung wird daher zwischen *Intake* und *Uptake* unterschieden (vgl. Harris 1992). Dabei bezieht sich der Begriff *Intake* auf all das, was angenommen wird, dass ein Kind vom Input wahrnimmt. Etwa wird beim frühen Spracherwerb davon ausgegangen, dass wortfinale und -initiale Elemente besonders gut von den Kindern wahrgenommen werden, aufgrund ihrer positionellen oder prosodischen Salienz.

Uptake vereint als Überbegriff sämtliche inputrelevanten Merkmale, die vom Kind abstrahiert und gespeichert werden. Auf Basis dieses angeeigneten Wissens konstruieren die Kinder Muster, die dem Input entsprechen können, oder aber, wie im Fall von Übergeneralisierungen und Sackgassenentwicklungen, sich von diesem maßgeblich unterscheiden können, wie dies beispielsweise für Komposition untersucht wurde von Korecky-Kröll et al. (2017: 30 f.). Was Kinder tatsächlich von ihrem Input abstrahieren und wie sie diese Informationen weiterentwickeln, ist naturgemäß nicht in direkter Weise untersuchbar, sondern kann nur auf Basis der Sprachproduktion und des Sprachverständnisses rekonstruiert werden.

In dieser Arbeit werden alle Wortbildungsmuster insbesondere der Derivationsmorphologie festgestellt, die in der Spontansprache von Kindern unter 5 Jahren und ihren Hauptbezugspersonen vorkommen. Dabei liegt der Fokus auf Wortbildungsmustern, die in dieser frühen Phase des Spracherwerbs bevorzugt verwendet werden, mit einem zusätzlichen Schwerpunkt auf einem Alter zwischen 3 bis knapp 5 Jahren. Die Untersuchung der kindlichen Spontansprache zeigt sich deshalb als so wesentlich, da häufige Fehler, die Kinder im Rahmen von Tests machen, aufgrund von Nervosität und Erfolgsdruck sowie metalinguistischen Anforderungen, in ihrer Spontansprache jedoch nicht vorzufinden sind. Daher ist eine sinnvolle und valide Methode die Untersuchung der Spontansprache in Kombination mit Ergebnissen von Testdaten. Aus diesem Grund hat die Verfasserin für den zweiten empirischen Teil der Arbeit einen Derivationstest entwickelt, der die gleichen Kinder zu einem späteren Zeitpunkt in ihrer Entwicklung untersucht. Die zusätzliche Testung soll insbesondere Aufschluss über das Sprachbewusstsein der Kinder im Alter von durchschnittlich 8 Jahren geben (vgl. dazu die Sprachentwicklungstheorie von Karmiloff-Smith 1996), da jedes Experiment ein hohes Sprachbewusstsein verlangt. Außerdem zeigt das Testergebnis die Unterschiede in der lexikalischen Weiterentwicklung der Kinder mit höherem oder niedrigerem sozioökonomischen Hintergrund im Bereich nominaler Ableitungen.

Die vorliegende systematische Studie beinhaltet Daten von insgesamt 32 Wiener Kindern und ihren Hauptbezugspersonen und untersucht den elterlichen Input und den kindlichen Output. Die Alterszeiträume erstrecken sich von 1,5 Jahren (3 Kinder) bis zu einem Alter von 3 Jahren (2 Kinder), sowie von 3 bis knapp 5 Jahren

(30 Kinder). Die sehr frühen Daten der drei Kinder wurden bereits in Sommer-Lolei et al. (2021) für den Zeitraum bis 3 Jahre untersucht. Diese Daten werden durch diese Arbeit detaillierter behandelt und in bestimmten Bereichen vertieft und erweitert. Vor allem wird der Zeitraum zwischen 3–5 Jahren untersucht, um eine Verbindung zu den in Mattes (2018) analysierten, späteren Daten eines der Kinder ab 6 Jahren herzustellen. Das Hauptziel dieser Arbeit besteht darin, die bestehenden Lücken, die es im Bereich der Derivationsmorphologie im Spracherwerb gibt, zu minimieren. Zunächst soll ein Überblick über die tendenzielle sprachliche Entwicklung von Kindern mit unterschiedlichem sozioökonomischen Hintergrund (höherer und niedrigerer sozioökonomischer Status) ab einem Alter von 3 Jahren bis knapp vor den fünften Geburtstag (4 Jahre 11 Monate) gegeben werden. Außerdem werden die Dyaden in zwei Gruppen aufgeteilt, um sie gegenüberstellend zu analysieren. Dies geschieht unter Berücksichtigung verschiedener Einflussfaktoren, wie Frequenz, Transparenz, Salienz und Komplexität. Daher werden Lemmafrequenzen¹ und Tokenfrequenzen der Derivationen und Komposita sowie die quantitative Verteilung der Wortbildungsarten und -muster im Input als auch im Output festgestellt. In einem weiteren Schritt werden die spontansprachlichen Daten einerseits den Ergebnissen der Elterninterviews gegenübergestellt, um den Einfluss der Inputmenge, sowie des Medienkonsums auf das Sprachverhalten der Kleinkinder, ihren Erwerbsverlauf und Wortschatz darzustellen. Andererseits werden die oben genannten Häufigkeiten mit den Ergebnissen des rund vier Jahre später durchgeführten Derivationstests verglichen, um die weitere sprachliche Entwicklung der Kinder im Bereich der deverbale nominalen Suffigierung zu zeigen. Dabei wird insbesondere festgestellt, welche Unterschiede und Zusammenhänge zwischen der Sprachproduktion der Kinder und jener ihrer Hauptbezugspersonen bestehen, wobei nur eine Aussage zur Sprachproduktion der Kinder und ihrer Hauptbezugspersonen getroffen werden kann, nicht jedoch zu deren Sprachverständnis.

¹ Dies entspricht in anderen Publikationen dem Begriff der Typenfrequenz (*type frequency*).

In meiner Arbeit verwende ich die gebräuchlichen Abkürzungen CS (*child speech*) für die Sprache des Kindes und CDS für die an das Kind gerichtete Sprache der Hauptbezugsperson (kurz HBP; Mutter oder Vater).

Das einleitende Kapitel 2 beschreibt die unterschiedlichen Arten der Wortbildung im Deutschen, geht insbesondere auf Derivationsmuster ein und gibt einen Überblick über wortartbewahrende und -verändernde Wortbildungsprozesse. Außerdem gewährt es einen kurzen Einblick in die Geschichte der Spracherwerbsforschung sowie in die verschiedenen Erhebungsmethoden von sprachlichem Material. Ebenso wird der aktuelle Forschungsstand im Bereich der Wortbildungsmorphologie gezeigt, die für diese Arbeit wesentlichen spracherwerbstheoretischen Ansätze werden dargestellt und grundlegende Termini der Spracherwerbsforschung definiert. Überdies behandelt das Kapitel verschiedene mögliche Einflussfaktoren auf den frühen Morphologierwerb. Abschließend finden sich die Forschungsfragen und Hypothesen.

Das gesamte analysierte Datenmaterial wird in Kapitel 3 umfassend und detailliert beschrieben und zeigt den Umfang der longitudinalen Spontansprachdaten und der Paneldaten, die zu Hause erhoben wurden.

In Kapitel 4 werden sämtliche verwendeten und angewendeten Methoden ausführlich dargestellt. Diese umfassen die Bereiche der Datenerhebung, Transkription, morphologischen Codierung, Datenauswahl und -bereinigung, Zählung, sowie der Operationalisierung des sozioökonomischen Status, der Erstellung und Durchführung der Elterninterviews, und des theoretischen Hintergrunds des Mini-Paradigma-Kriteriums für die Feststellung der potenziellen Produktivität eines Wortbildungsmusters in der Kindersprache.

Kapitel 5 zeigt die Verteilung der Derivationen im Vergleich zu Komposita in den Korpora in CS und CDS und geht auf die relevantesten Wortbildungsmuster bei Substantiven, Verben und Adjektiven ein und stellt diese umfassend dar. Dabei wird speziell auf jene Muster eingegangen, die sich in der Kindersprache als besonders produktiv herausstellen. Außerdem werden die jeweiligen Emergenzen derivierter und zusammengesetzter Lexeme in allen drei untersuchten Wortklassen, sowie die potenzielle Produktivität verschiedener Wortbildungsmuster festgestellt.

Überdies werden die Erwerbsverläufe derivierter Substantive und Verben und synthetischer Komposita bei CS im Vergleich zu den Verläufen in CDS dargestellt. Abschließend folgen die statistischen Auswertungen der longitudinalen Spontansprachdaten in Kapitel 5.7, um Zusammenhänge und Unterschiede zwischen CS und CDS aufzuzeigen. Die Paneldaten werden in Kapitel 5.8 behandelt, wo Gruppenvergleiche nach sozioökonomischem Status der Kindersprache (CS vs. CS) und des Inputs (CDS vs. CDS), sowie zwischen CS und CDS, als auch Veränderungen zwischen zwei Datenzeitpunkten ausgewertet werden. Die spontansprachlichen Daten werden in Kapitel 10.2 der Menge des Inputs, des Medienkonsums und Büchervorlesens gegenübergestellt und auf statistisch signifikante Zusammenhänge überprüft.

Die Ableitungsbasen der Derivationen werden in Kapitel 6 beleuchtet. Dabei werden die Basen in nominalen, verbalen und adjektivischen Ableitungen nach Wortklasse getrennt behandelt. Überdies werden die prozentualen Verteilungen der verschiedenen Basen in CS und CDS je Korpus dargestellt und anhand von Beispielen illustriert.

Kapitel 7 geht auf die Komplexität in der Wortbildung ein und auf die Entwicklung komplexer Wortformen im Erwerbsverlauf. Dabei wird zunächst ein Überblick über die Verteilung einfacher und komplexer Derivationen und Komposita in CS und CDS nach Wortbildungsart gegeben, und in einem weiteren Schritt werden die Anteile der Wortbildungsmuster je Komplexitätsgrad nach Wortklasse untersucht. Überdies wird die Verteilung der morphologischen Komplexitätsgrade derivierter und zusammengesetzter Substantive, Verben und Adjektive in allen Korpora nach Wortklasse und Wortbildungsart unter Verwendung zahlreicher Beispiele anschaulich dargestellt und analysiert. Dabei wird insbesondere auf die Lemma- und Tokenfrequenzen in CS und CDS eingegangen.

Kapitel 8 beinhaltet eine umfangreiche Fehleranalyse der Ableitungen in der Kindersprache in allen untersuchten Wortklassen (Substantive, Verben und Adjektive). Der Schwerpunkt liegt dabei auf den Fehlerhäufigkeiten und hauptsächlichsten Fehlermustern bei derivierten Substantiven und Verben sowie bei synthetischen Komposita.

Kapitel 9 beschreibt das entwickelte und durchgeführte Experiment zur Elizitation nominaler Ableitungen und gibt eine Übersicht über die Teilnehmer des Experiments, die verwendeten Stimuli sowie den Testablauf. Ebenso werden die Daten des Tests umfassend ausgewertet und eine genaue Fehleranalyse durchgeführt. Die Ergebnisse des Derivationsexperiments werden im Anschluss in Kapitel 10.1 den Daten der Elterninterviews gegenübergestellt und die Zusammenhänge zu den Variablen Inputmenge, Medienkonsum und Büchervorlesen errechnet.

Die Korrelationen zur Menge des Inputs, des Medienkonsums und des Büchervorlesens werden in Kapitel 10 ausführlich behandelt. Dabei werden die Korrelationen zu den Ergebnissen des Elizitationsexperiments in Kap. 10.1, und zu verschiedenen Wortbildungskategorien der Spontansprache in Kap. 10.2 berechnet und signifikante Zusammenhänge und Trends festgestellt.

In Kapitel 11 werden die Forschungsfragen und Hypothesen beantwortet und die wesentlichsten Erkenntnisse zusammenfassend dargestellt.

Im abschließenden Kapitel 12 wird ein Ausblick auf zukünftige Forschungsvorhaben sowie auf Sprachfördermaßnahmen gegeben.

Diese Arbeit soll durch die Grundlagenforschung Licht in bisher vernachlässigte Derivationsmuster im deutschen Erstspracherwerb bringen. Überdies soll sie einen Einblick in die Zusammenhänge und Unterschiede zwischen CS und CDS geben, sowie wesentliche Einflussfaktoren zu Tage fördern, die sich hemmend oder förderlich auf die morphologische Sprachentwicklung auswirken. Die Ergebnisse der Untersuchungen sollen zeigen, welche Variablen den Spracherwerb begünstigen, um in weiterer Folge auf Basis dessen dazu beizutragen, gezielte Sprachfördermaßnahmen bei Kindern und Jugendlichen mit individuellem und/oder sozialem Handicap anzuregen.

2 Wortbildung und der Erwerb der Morphologie

2.1 Wortbildungsarten im Deutschen

Das folgende Kapitel soll als Grundlage dienen und stellt eine kurze Einführung in die Begrifflichkeiten und Definitionen der Wortbildungslehre dar. Es gibt einen Überblick über die wesentlichsten Wortbildungsarten und -muster, die für die spätere Untersuchung der Spontansprachdaten in dieser Arbeit entscheidend sind. Für eine umfassende und ausführliche Beschreibung der Wortbildung in der deutschen Gegenwartssprache siehe Kühnhold und Wellmann (1973), Wellmann (1975), Kühnhold et al. (1978), Barz (2009), Fleischer und Barz (2012: 1–116), Eisenberg (2013) und Elsen (2014).

Die deutsche Sprache ist als westgermanische Sprache klassifiziert und gehört den flektierenden Sprachen an. Das österreichische Standarddeutsch, dessen Erwerb und Entwicklung Untersuchungsgegenstand dieser Arbeit darstellt, unterscheidet sich vom deutschen Standarddeutsch in einigen Lexemen, jedoch weder in Bezug auf seine Wortbildungsmorphologie noch seinen Erwerb. Das Deutsche verfügt über mehrere Möglichkeiten zur Wortbildung. Per Definition ist die Wortbildung als Prozess zur Bildung neuer Wörter einerseits, sowie als Resultat dieses Prozesses andererseits zu verstehen. Neue Lexeme, die nicht als Lehn- oder Fremdwort in den deutschen Wortschatz eingehen, werden hauptsächlich mittels **Komposition** und **Derivation** gebildet. Die auf diese Weise neu gebildeten Lexeme gehören meist den Wortarten Substantiv, Verb, Adjektiv und Adverb an (vgl. Fleischer & Barz 2012: 2). Im Rahmen dieser Arbeit werden sämtliche Lexeme in verschiedenen Spontansprachkorpora morphologisch analysiert (siehe Kap. 3), die durch eine der oben genannten Wortbildungsarten entstehen und den Wortarten Substantiv, Verb oder Adjektiv zuzuordnen sind. Verschiedene Wortformen, die durch das Anhängen von Flexionsmorphemen gebildet werden, wie z. B.: leg-st ‚leg-2SG‘, Versteck-e ‚Versteck-PL‘, werden in diesem Zusammenhang weder berücksichtigt noch untersucht, da diese für den frühen Erwerb der Wortbildungsmorphologie kaum relevant sind.

2.1.1 Komposition

Bei der Komposition handelt es sich um die im Deutschen produktivste Art der nominalen Wortbildung. Typischerweise werden zwei freie, lexikalische Morpheme zu einem neuen Lexem verbunden. Dabei können generell Lexeme aus allen Wortarten als Erstglied dienen, sowie Konfixe und sogar einzelne Buchstaben. In der überwiegenden Mehrheit der Komposita nimmt allerdings das Substantiv diese Position ein. Die Komposition dient hauptsächlich der Bildung von Substantiven, in geringerem Ausmaß von Adjektiven, und nur in Einzelfällen von Verben, die zudem nur selten genutzt werden (vgl. Barz 2009: 710). Charakteristisch für nominale Komposita im Deutschen ist ihre binäre Struktur, und dass sie häufig endozentrisch² und rechtsköpfig sind. Nur selten verfügen sie über mehr als drei, einfache oder komplexe, Konstituenten. Das Letztglied eines typischen Kompositums ist der Kopf des Kompositums, das Determinatum, das Wortart und Flexionskategorie des Zielwortes bestimmt, während das Erstglied, als abhängiges Element, das Determinans, den Kopf näher spezifiziert bzw. modifiziert (vgl. Rettung-s+auto, Müll+auto, Renn+auto, Polizei+auto, Lieb-ling-s+auto).

Die Schnittstelle zwischen den beiden Konstituenten wird Kompositionsfuge genannt. Diese kann Fugenelemente, sogenannte Interfixe aufweisen, die in Abhängigkeit des Erstglieds eines Kompositums stehen. Interfixe sind grammatische, gebundene Morpheme, die an dieser Stelle in ihrer Funktion als Bindeglied zwischen den unmittelbaren Konstituenten auftreten. Interfixe kommen nach Barz (2009: 712) in rund 30 % aller Komposita vor, woraus resultiert, dass der überwiegende Großteil der Komposita interfixlos gebildet wird. Interfixe für nominale und adjektivische Komposita sind: -e-, -s-, -es-, -n-, -en-, -ens-, -er-.

Der Komplexität von Komposita sind formal kaum Grenzen gesetzt, da sie sich beinahe unbegrenzt erweiterbar zeigen. Je komplexer ein Kompositum jedoch wird, desto schwieriger ist es auch zu erfassen. Komposita können ihrerseits ebenfalls als Basis für andere Komposita dienen. In ihrer einfachsten Zusammensetzung, wie

² Bei endozentrischen Komposita bringen die beiden Konstituenten die Bedeutung des Gesamlexems mit, während bei exozentrischen Komposita die Bedeutung von außerhalb hinzukommt und nicht von den beiden Konstituenten beigesteuert wird. Exozentrische Komposita vom Typ Lang+finger, als Bezeichnung für eine Person, die stiehlt und nicht einfach lange Finger hat, kommen in den Daten nicht vor.

dies bei der Verbindung zweier Lexeme oder eines Wortstamms und eines Lexems, ohne Hinzukommens eines Interfixes, der Fall ist, ist die Komposition sehr produktiv: Haus_N+tier_N, Wohn_{VS}+haus_N, hell_{ADJ}+grün_{ADJ}. Kommen ein Interfix oder mehrere Konstituenten hinzu, nimmt das Kompositum naturgemäß an Komplexität zu: Not_N+arzt_N+auto_N, Kind_N-er_{INTF}+garten_N+kind_N, Flug_N+hafen_N+feuer_N+wehr_N.

Einigkeit herrscht über die besondere Stellung, die synthetische Komposita³ innerhalb der Wortbildung einnehmen. Kontrovers diskutiert wird hingegen deren eindeutige Zuordenbarkeit und Definition aufgrund ihrer komplexen, internen Struktur. Ob synthetische Komposita zur Komposition oder Derivation zu zählen sind, ist nach wie vor umstritten (Gaeta 2010; Olsen 2017; für eine diachrone Betrachtungsweise siehe Werner 2017). Sie vereinen die Wortbildungsart der Komposition mit der Bildung eines deverbalen Substantivs durch ex- oder implizite Derivation, wie in Schach_N+spiel_{VS}-er_{NAF}. Dabei dient das Erstglied als Argument des Verbs. Neben der Ableitung mittels eines overtten Ableitungssuffixes gibt es im Deutschen auch die Möglichkeit, durch Konversion deverbale Substantive zu bilden: Kind_N-er_{INTF}+sitz_N, Wasser_N+fall_N. Innerhalb dieser Arbeit betrachtet die Verfasserin die synthetischen Komposita separat, um ihrem besonderen Status innerhalb der Wortbildung Rechnung zu tragen.

2.1.2 Derivation

Die Derivation (Ableitung) ist nach der Komposition die zweite Wortbildungsart zur Bildung neuer Lexeme im Deutschen. Dabei wird zwischen expliziter und impliziter Derivation unterschieden. Im Vergleich zur Verb- und Adjektivderivation ist die Substantivableitung mit Abstand am umfangreichsten ausgeprägt. Zur Bildung neuer Substantive stehen wesentlich mehr Muster und unterschiedliche Affixe zur Verfügung, als dies für andere Wortklassen der Fall ist.

Die explizite Derivation ist gekennzeichnet durch das Anhängen eines Ableitungssuffixes. Das Affix kann dabei dem Wort, dem Wortstamm, der Wurzel, also einem freien oder gebundenen Morphem, vorangehen (Präfix), diesem folgen

³ In der Literatur finden sich auch die Termini Zusammenbildungen und Rektionskomposita als Bezeichnung.

(Suffix), oder dieses umschließen (Zirkumfix). Davon zu unterscheiden ist die implizite Derivation, im Zuge derer der Ableitungsbasis kein Derivationsaffix hinzugefügt wird, sondern die Basis durch Modifikation und/oder Ab- oder Umlaut des Stammvokals verändert wird (find-env → Fund_N, ziehenv → Zug_N, warm_{ADJ} → wärm-env), wobei im Fall von steh-env zu stell-env, lieg-env zu leg-env und sitz-env zu setz-env diese Ableitungsbeziehung synchron kaum mehr erkennbar ist. In manchen Fällen wird die implizite Derivation als morphologische Konversion oder auch Verbstammkonversion angesehen (vgl. die Beispiele Flug, Griff, Aufstieg [u. a.] in Fleischer & Barz 2012: 267–269; und Rainers Anmerkung dazu 2010: 172), obwohl eine formale Änderung vorliegt. Dies ist im Bereich der Untersuchung der Kindersprache verständlich, da die Kinder falls überhaupt nur über wenige Bildungen verfügen und eine Differenzierung zur Konversion dann nicht lohnenswert erscheint. In der vorliegenden Arbeit wird allerdings nicht allein CS, sondern auch CDS untersucht, weshalb diese Unterscheidung jedenfalls berücksichtigt wird.

Ableitungsaffixe sind grammatische, gebundene Morpheme, die nicht für sich selbst wortfähig sind und an andere, freie oder gebundene lexikalische Morpheme, sowie Konfixe, angefügt werden. Dabei kann es sich um indigene Suffixe, Präfixe und Zirkumfixe, sowie um Fremdaffixe handeln. Was die Suffigierung für die Bildung neuer Substantive ist, ist die Präfigierung für die Bildung von Verben. Die Zirkumfigierung ist ebenso ein Muster zur Ableitung von Substantiven, Verben und Adjektiven, nimmt jedoch eine Randposition ein. Bei den Präfixen sind trennbare (= Partikel, z. B. auf#, um#, weg#, an#) und untrennbare Präfixe (z. B. be-, ent-, ver-, zer-) zu unterscheiden. Generell sind Affixe an bestimmte Beschränkungen und Restriktionen gebunden, beispielsweise welche Wortarten als Basen für die Ableitung dienen können.

Konversion

Die Konversion (auch Nullderivation) bezeichnet einen Wortbildungsprozess, der ohne die offensichtliche Verwendung eines Wortbildungsmerkmals (Affix, Ab- oder Umlaut, Stammmodifikation) stattfindet und dabei zu einem Wortartwechsel führt (siehe dazu Kap. 2.2). Ob die Konversion als Wortbildungsprozess innerhalb der Derivation, als sogenannte Nullableitung, also mittels eines Nullaffixes, oder

als eigenständige Wortbildungsart anzusehen ist, wird kontrovers diskutiert (für die Theorie der Nullableitung siehe u. a. Manova & Dressler 2005; Don 2005; Valera 2015; für eine vorrangig syntaktische Funktion unterspezifizierter Lexeme siehe u. a. Bybee 1995; Haspelmath 1996; Eschenlohr 1999).

Bei der Nullableitung wird zwischen morphologischer und syntaktischer Konversion unterschieden (vgl. dazu Erben 2006; Eisenberg 2013). Behält das Wortbildungsprodukt ein Flexionselement seiner Basis, handelt es sich um die syntaktische Konversion. Dazu zählen nominalisierte Infinitive (laufenv → das Laufenn), also substantivierte Verben, die die Infinitivendung der Basis unverändert beibehalten, sowie Partizipien in ihrer Verwendung als Adjektive (Präsenspartizip *sprechend*_{ADJ} ← Partizip I, Perfektpartizip *gebrochen*_{ADJ} ← Partizip II, *entzückt*_{ADJ} ← Partizip II) oder als Substantive (Präsenspartizip *der/die Wütenden* ← Partizip I, Perfektpartizip *der/die Vorgesetzten* ← Partizip II).

Die morphologische Konversion umfasst nach Fleischer und Barz (2012: 88) all jene Bildungen, in denen „zwei phonologisch gleiche Stämme in einer Motivationsbeziehung zueinander“ stehen, wie in: *der Stoß*_N ← *stoß*_{ENV}, *der Besuch*_N ← *besuch*_{ENV}, *fisch*_{ENV} ← *Fisch*_N, *das Tief*_N ← *tief*_{ADJ}. Der verbale Wortstamm oder das Ausgangslexem, das als Basis dient, wird in eine andere Wortart überführt und bleibt dabei formal und in seiner Bedeutung unverändert im Konversionsprodukt vorhanden. Nachdem Konfixe keiner Wortart angehören, sind diese als Basis für Konversionen ausgeschlossen.

Nicht immer ist die Ableitungsrichtung klar erkennbar. Ob etwa *Blitz* eine Konversion von *blitzen* darstellt oder ob es sich umgekehrt verhält, kann synchron nicht geklärt werden. Weitere Beispiele für die Schwierigkeit der Interpretation der Derivationsbeziehung sind die folgenden Wortpaare: *Spiel/spielen*, *Teil/teilen*, *Donner/donnern*, *Feier/feiern*, *Antwort/antworten*, *Schaukel/schaukeln*.

Derivationsmuster

Ein Derivationsmuster ist das einem Derivat zugrundeliegende Strukturschema. Wie bereits in Kap. 2.1.2 erläutert, handelt es sich bei den derivationellen Bildungsmustern um Affigierung (Suffigierung, Präfigierung, Zirkumfigierung),

Stammmodifikation und/oder Stammvokaländerungen sowie Konversion hauptsächlich verbaler, nominaler und adjektivischer Basen.

Tabelle 1 gibt einen umfassenden Überblick über die Derivationsmuster zur Bildung von Substantiven, Verben und Adjektiven im Deutschen basierend auf Fleischer und Barz (2012: 117–359, 373–441) und in Anlehnung an Mattes (2018: 90).

Tabelle 1. Derivationsmuster im Deutschen – Überblick
(in Anlehnung an Mattes 2018: 90)

		Suffixe, Präfixe, Andere	
S u b s t a n t i v	Derivation ca. 70 Affixe	Suffigierung -chen, -e, -ei/-er-ei, -el, -er/-ler/-ner, -heit/-keit/ -igkeit, -i, -ian, -in, -lein, -ling, -nis, -rich/-erich, -s, -sal, -schaft, -sel, -tel, -tum, -ung, -werk, -bold, mind. 26 Fremdsuffixe	
		Präfigierung Erz-, Ge-, Haupt-, Miss-, Un-, Ur-, mind. 15 Fremdpräfixe	
		Zirkumfigierung morph. Konversion syntakt. Konversion	Ge-...-e, Ge-...-t, Ge-...-st, Ge-...-de Wortartwechsel (fallen → der Fall) Nominalisierung des Infinitivs (das Lesen)
		V e r b	Derivation ca. 30 Affixe
be-, ent-, er-, (ge-), miss-, ver-, zer-, durch-, hinter-, über-, um-, unter-, wider-, mind. 9 Fremdpräfixe			
Präfigierung mittels einer trennbaren Partikel (an#, auf#, her#, hin#, um#, weg#, zu#, herum#, etc.)			
-(e)l, -(e)r, -ier/-isier/-ifizier, -ig, -s, -z/-enz be-...-ig, (v)er-...-ig, ent-...-ig, ver-...-ier Wortartwechsel (Angel → angeln)			
A d j e k t i v	Derivation ca. 40 Affixe	Suffigierung -bar, -e(r)n, -fach, -haft, -ig, -isch, -lich, -los, -mäßig, -sam, mind. 8 Fremdsuffixe	
		Präfigierung Zirkumfigierung	erz-, miss-, un-, ur, mind. 16 Fremdpräfixe un-...-bar/-lich/-sam, be-/ge-/ent-/zer-...-(e)t ge-...-ig
		morph. Konversion syntakt. Konversion	Wortartwechsel (Sand → sand) Departizipiale Konv. (PP erstaunt, PPräs glänzend)

2.2 Wortartwechsel durch Wortbildung

Häufig kommt es im Zuge des Wortbildungsprozesses zu einem Wortartwechsel. Dabei wird eine Basis (Lexem oder Wortstamm) in eine andere Wortart überführt. Dies betrifft in CS und CDS vor allem Bildungen durch Konversion, aber auch verschiedene Derivationsmuster.

Wortartbewahrende und wortartverändernde Wortbildungsprozesse

Komposition

Im Fall der Komposition gibt das Letztglied des Kompositums, in seiner Funktion als Kopf, die Wortart vor (siehe Kap. 2.1.1). Das Determinatum behält somit seine Wortklasse, sein Genus und seine Flexionsklasse bei, vgl. die Bahn_N → die Eisen_N+bahn_N, was auch bei synthetischen Komposita der Fall ist: der Fahr-er_N → der Auton+fahr-er_N. Die Komposition zählt damit zu den wortartbewahrenden Wortbildungsverfahren.

Konversion

Wie in Kap. 2.1.2 (S. 35) behandelt, gehören Lexeme, die durch Konversion entstehen, einer anderen Wortart an als ihre Basen. Die Ursprungseinheiten gleichen den Konversionsprodukten phonologisch und werden durch die Konversion in eine andere Wortart überführt: blau_{ADJ} → das Blaun_N, weit_{ADJ} → weit-_{env}, Angel_N → angel-_{nv}. Sie zählt daher zu den wortartverändernden Wortbildungsverfahren.

Derivation

Bei der Derivation verhalten sich die verschiedenen Präfixe, Suffixe und Zirkumfixe unterschiedlich.

Zirkumfixe zählen zu den wortartverändernden Affixen, da sie die Wortart ihrer Basis meist ändern: schön_{ADJ} → be-schön-ig-_{env}, heul-_{env} → Ge-heul-_{en}, Berg_N → Ge-birg-_{en}.

Suffixe können sowohl zu einem Wortartwechsel führen wie in: fahr-_{env} → Fahr-er_N, heiz-_{env} → Heiz-ung_N, Angst_N → ängst-lich_{ADJ}, Marsch_N → marsch-ier-_{env}, intensiv_{VADJ} → intensiv-ier-_{env}, oder wortartbewahrend sein: streich-_{env} → streichel-_{nv}, steig-_{env} → steig-er-_{nv}. Motionssuffixe hingegen behalten die Wortart der Basis bei, verändern jedoch deren Genus: Freund_N → Freund-_{inn}.

Verbale Präfixe, die denominal, deverbal und deadjektivisch ableiten, können wie Suffixe, nicht nur wortartverändernd sein, sondern auch wortartbewahrend. Dies trifft auf untrennbare Präfixe (be-, ent-, er-, ver-, zer-): Heimat_N → be-heimat-_{env},

suchen_V → be-such-en_V, arm_{ADJ} → ver-arm-en_V, als auch auf trennbare Präfixe (= Partikel, auf#, an#, um#, weg#, zu#, etc.) zu: fall-en_V → auf#fall-en_V, um#fall-en_V, Leinen → an#leine-n_V, breit_{ADJ} → aus#breit-en_V.

Keinen Wortartwechsel hingegen bewirken nominale und adjektivische Präfixe (un-, ur-, miss-, erz-), die Substantive und Adjektive als Basen nehmen: Un-wetter_N, Erz-feind_N, Miss-erfolg_N, ur-schön_{ADJ}, un-genau_{ADJ}.

2.3 Geschichte der Spracherwerbsforschung

2.3.1 Erhebungsmethoden

Um Einblicke in die Sprachverarbeitungs- und Spracherwerbsprozesse zu erhalten, vor allem in früheren Phasen der sprachlichen Entwicklung, kommen unterschiedliche Erhebungsmethoden zum Einsatz.

Zum einen handelt es sich um Befragungen, im Rahmen derer Bezugspersonen eines Kindes Auskunft über dessen Wortschatz z. B. mittels Wortlisten (Vokabellisten) erteilen oder gezielte Fragen in Elternfragebögen beantworten zum frühen Spracherwerb (Sprachverständnis und -produktion) ihres Kindes (siehe dazu Kauschke 2012: 8 f.). Die Kinder selbst werden bei dieser Methode weder miteinbezogen noch direkt befragt. Die Verlässlichkeit solcher Elternbefragungen hat sich bei einem Vergleich von Spontansprachdaten und den Angaben der zugehörigen Elternfragebögen in Studien zum Deutschen bestätigt. Die Hauptbezugspersonen sind reliabel, wenn es um die Einschätzung des sprachlichen Vermögens ihres Kindes geht (vgl. Szagun et al. 2006; Szagun 2007).

Zum anderen werden Beobachtungsstudien durchgeführt. Diese können Tagebuchstudien, longitudinale Spontansprachuntersuchungen oder auch Tests sein, die punktuell zu einem bestimmten Zeitpunkt oder über einen längeren Zeitraum die Sprache der Kinder dokumentieren. Bei den Tagebuchaufzeichnungen handelt es sich um die Mitschrift der sprachlichen Äußerungen der eigenen Kinder, wie jenen von Neugebauer-Kostenblut (1914), Stern und Stern (1928), oder in jüngerer Vergangenheit Meibauer (1995a, 1999), Elsen (1999) und Rainer (2010). Dabei werden die kindlichen Wörter und Kreationen zum Zeitpunkt ihrer Äußerung einzeln und manuell erfasst. Der Input durch die Bezugsperson wird, wie auch bei

den Tests und Befragungen, nicht erhoben. Häufig rückt die Aufnahme von Neubildungen in den Vordergrund, da diese nicht vom Input übernommen werden können und so das produktive Anwenden eines morphologischen Musters zeigen. Aus dem gleichen Grund wird in Testverfahren ebenso die Elizitation neuartiger Bildungen auf Basis von Nonsenswörtern angestrebt. Auf derartigen Testverfahren beruhen die Studien von Bartke (1998) und Clahsen et al. (1996). Während bei der Tagebuchstudie das Kind unbeeinflusst von einer ihm vertrauten Person beobachtet wird, werden bei jedem Experiment oder Test bestimmte metalinguistische Anforderungen an das Kind gestellt. Zudem kann sich das ungewohnte Umfeld, die Testsituation oder die Anwesenheit des Experimentators oder der Experimentatorin nachteilig auf die Leistung des Kindes auswirken.

Im Gegensatz zu Testdaten und Tagebuchstudien erlauben die über einen längeren Zeitraum regelmäßig erhobenen, aufgezeichneten und transkribierten Spontansprachdaten nicht nur eine statistische Auswertung, sondern die gleichzeitige Analyse der an das Kind gerichteten Sprache (CDS) durch die anwesende Hauptbezugsperson. Diese Methode weist eine höhere ökologische Validität auf, hat jedoch den Nachteil, die mitunter spärlich vorkommenden Neubildungen des Kindes zu verpassen, weil gerade nicht aufgenommen wird. Andererseits wird der Output des Kindes nicht selektiert, wie dies bei den Tagebuchstudien der Fall ist, sondern über den Aufnahmezeitraum in dessen Gesamtheit erfasst.

In dieser Arbeit werden verschiedene Methoden herangezogen, um ein möglichst ganzheitliches Bild des Spracherwerbs der Kinder abzubilden. Es wird die transkribierte, longitudinal erhobene Spontansprache analysiert, um Testdaten ergänzt, mit den Ergebnissen der Elterninterviews in Relation gesetzt und statistisch ausgewertet.

2.3.2 Wortbildungsmorphologie im Erstspracherwerb des Deutschen – Forschungsstand

Die Wortbildung stellt ein wichtiges und wesentliches Funktionsmerkmal des Deutschen dar, ist jedoch im Bereich der frühen Sprachentwicklung von Kindern, vor allem den Erwerb der Derivationsmorphologie betreffend, bislang nur marginal erforscht (vgl. Rainer 2010: 9; Schipke & Kauschke 2011: 69; Kauschke 2012: 72).

Da es nur vereinzelte Publikationen auf diesem Gebiet im Vergleich zu anderen Bereichen der Erstspracherwerbsforschung gibt, bestehen große Forschungslücken.

Die in Kap. 2.3.1 beschriebenen Tagebuchstudien haben auf Basis von Neologismen gezeigt, dass Kinder rund um ihren zweiten Geburtstag die Möglichkeit für sich entdecken, Wörter abzuleiten (siehe Neugebauer-Kostenblut 1914; Stern & Stern 1928; Elsen 1999; Rainer 2010), wengleich sich diese Bildungen auf wenige Muster beschränken.

Die bei Substantiven am häufigsten untersuchten Wortbildungsmuster im Spracherwerb sind die Komposition und die Suffigierung, letztere eingeschränkt auf Diminutivbildungen sowie Ableitungen auf -er. Abgesehen davon gibt es kaum oder keine relevanten Untersuchungen zu Konversion, Präfigierung und Suffigierung. Ebenso verhält es sich mit der Verbalderivation, über deren Entwicklung im Erstspracherwerb noch weniger berichtet wurde (Behrens 1998; Schulz et al. 2001). Nahezu völlig unbeachtet ist die Adjektivderivation, da diese im frühen Spracherwerb nur eine äußerst geringe Rolle spielt.

Hauptsächlich untersucht wurde der frühe Erwerb der Nominalkomposita (vgl. Dressler et al. 2010; Korecky-Kröll et al. 2017), der Zusammenhang zwischen Komposita und Pluralflexion im Erwerbsverlauf 3- bis 9-jähriger Kinder (vgl. Clahsen et al. 1992; Clahsen et al. 1996; Bartke 1998) und die Emergenz verschiedener Kompositionsmuster in der protomorphologischen Phase des Spracherwerbs (vgl. Lettner et al. 2011). Innerhalb der Derivation wurde der Erwerb von Diminutivsuffixen ausführlich behandelt, da diese sowohl in CS als auch in CDS eine zentrale und wesentliche Rolle einnehmen (vgl. Korecky-Kröll & Dressler 2007; Dressler et al. 2012). Synthetische Komposita wurden in einem sprachübergreifenden Vergleich in sieben Sprachen von Dressler, Sommer-Lolei, et al. (2019) untersucht. Den Erwerb der Nominalmorphologie von zwei Wiener Kindern behandelt die Arbeit von Korecky-Kröll (2011), mit dem Ergebnis, dass sich nur das Ableitungssuffix -er bei den Kindern als produktiv herausstellte und sich nur wenige Neologismen mit -ung, sowie Konversionen zeigten. Präfixe zur Ableitung von Substantiven wurden bis zu einem Alter von 6 Jahren nicht produktiv verwendet. Bisläng wurde das in der Kindersprache frequenteste Derivationsmuster -er, vorrangig zur Ableitung von Nomina Agentis und Instrumenti, intensiv

erforscht (vgl. Meibauer 1995b, 1995a; Meibauer et al. 2004), da es in CS für gewöhnlich sehr früh auftritt und tendenziell vor anderen Ableitungsmustern erworben wird. Rainer (2010) untersuchte den Spracherwerb seiner Tochter bis zum Schuleintritt und gibt einen umfassenderen Einblick in die Entwicklung der Komposition und Derivation in ihrer Sprachproduktion. Mattes (2018) ist eine eingehende Spracherwerbsstudie, die sich auf den Erwerb derivationeller Muster in einem Alter zwischen 5 und 7 Jahren fokussiert. Erste Emergenzen sowie die Entwicklung der Derivationsmorphologie und ihre produktive Verwendung im frühen Spracherwerb (bis zu einem Alter von 3 Jahren) wurde für das Deutsche erstmalig für Substantive, Verben und Adjektive von Sommer-Lolei et al. (2021) untersucht. Dabei wurde festgestellt, dass andere Muster, wie das nur schwach produktive Suffix -e, und morphologische Konversion ebenfalls früh auftreten. Das im Deutschen frequente Suffix -ung tritt für gewöhnlich später als die zuvor genannten auf, meist im Verlauf des dritten Lebensjahres des Kindes. Vereinzelt Bildungen treten zwar auch bei jüngeren Kindern auf, werden jedoch nicht vor einem Alter von 5 oder 6 Jahren als voll produktiv betrachtet.

Das wichtigste und mit Abstand frequenteste Ableitungsmuster im frühen Spracherwerb im Bereich der Verbderivation ist die Partikelverbbildung. Ebenfalls wesentlich, wenngleich in geringerem Ausmaß, ist die Konversion. Untrennbare Präfixe zur Ableitung von Verben sind irrelevant bis zu einem Alter von 3 Jahren (vgl. Sommer-Lolei et al. 2021), nehmen aber im Erwerbsverlauf (bis 6 Jahre und danach) eine immer größere Position ein (vgl. Behrens 1998; Rainer 2010; Schipke & Kauschke 2011; Mattes 2019).

Im Rahmen der Wortbildung von Adjektiven ist das einzig relevante Derivationsmuster die Ableitung mittels des Suffixes -ig und bleibt dies auch bis zu einem Alter von 5 Jahren (vgl. Rainer 2010; Mattes 2018: 313).

In Bezug auf die Komplexität der Ableitungsbasen hat sich gezeigt, dass komplexe Basen ab einem Alter von 3 Jahren schnell zunehmen (Meibauer 2001; Meibauer et al. 2004).

2.4 Spracherwerbstheoretische Ansätze

Für die Untersuchung des (Erst)spracherwerbs gibt es vielfältige theoretische Erklärungsmodelle, mit zum Teil unterschiedlichen Grundsätzen. In dieser Arbeit folgt die Verfasserin dem konstruktivistischen Ansatz, der auf einem Modell selbstorganisierter Prozesse basiert (vgl. Karpf 1990; Dressler & Karpf 1995; Karmiloff-Smith 1996; Elman et al. 1996). Der Begriff der Selbstorganisation bezieht sich dabei auf die laufende Rekonstruktion, Umstrukturierung und Anpassung des im Kind vorhandenen Systems der angeborenen Veranlagung aufgrund des ständigen Austauschs bzw. der permanenten Interaktion mit der Umgebung bzw. den Reizen aus der Umwelt. Wenngleich der Ansatz von einer angeborenen Veranlagung ausgeht, ist er unbedingt von dem nativistischen Ansatz Chomskys (1965, 1986, 2002) und der damit verbundenen Theorie der Universalgrammatik (*universal grammar*, UG) mit ihren sprachrelevanten universellen Prinzipien und sprachspezifischen Parametern (siehe Chomsky & Lasnik 1993) zu unterscheiden. Der Theorie zufolge legt der linguistische Input zwar fest, welche Sprache erworben wird, ist jedoch zu ärmlich, um für den relativ schnellen und einfachen Erwerb der Erstsprache verantwortlich sein zu können. Nachdem der Prozess des Spracherwerbs eines Kindes aus generativ-grammatischer Sicht hauptsächlich biologisch gesteuert und nicht vom Input abhängig ist, benötigt es die zusätzliche Hilfe der Prinzipien und Parameter der UG, um diesen adäquat zu unterstützen.

Anders sehen es die Vertreter funktionalistischer, produktionsbasierter (*usage-based*) Modelle (vgl. Bates & MacWhinney 1987: 160; Dressler 1995; Elsen 1999; Tomasello 2003; Clark 2009; Bybee 2010; Lieven 2014), die in ihrem Grundsatz die kommunikative Funktion der Sprache als zentralen Aspekt ansehen. Sie gehen von einer direkten Abhängigkeit zwischen dem Input und dem Output der Kinder aus, liefern jedoch für die Bildung von Sackgassen (*blind alleys*) im frühen Erwerbsverlauf keine Erklärung (vgl. Dressler, Christofidou, et al. 2019: 108).

2.4.1 Morphologische Entwicklungsphasen im frühen Spracherwerb

Die morphologische Entwicklung im Erstspracherwerb von Kindern untergliedert die Autorin gemäß Dressler und Karpf (1995) und Dressler et al. (2010: 325 f.) in

drei tendenzielle Erwerbsphasen, nämlich die prämorphologische Phase, die protomorphologische Phase und die Phase der Morphologie. Dem Ansatz liegen drei Theorien zugrunde: die morphologische Natürlichkeitstheorie (siehe Kap. 2.4.2), die Theorie der selbstorganisierenden Prozesse und die Theorie des Funktionalismus (siehe Kap. 2.4). Die Merkmale der einzelnen Spracherwerbsphasen sollen im Folgenden kurz erläutert werden:

In der **prämorphologischen Phase** ist die Sprache der Kinder auf ganzheitlich gespeicherte Formen reduziert, und es werden noch keine morphologischen Regeln abgeleitet. Kennzeichnend für diese Periode sind die häufigen Wortkürzungen und -verschmelzungen (Amalgamierung), Reduplikationen, sowie Onomatopöien. So produzieren die Kinder lexikalisch gespeicherte Komposita und Derivationen, ohne dass die beiden Konstituenten eines Kompositums oder die Basis ihrerseits auch für sich bzw. in anderer Kombination in der Sprache der Kinder vorkommen. Da das Kind sich der morphologischen Grenzen eines Lexems zunächst nicht bewusst ist und überdies kürzere Wörter einfacher auszusprechen sind, kommt es zu Amalgamierungen komplexer Formen. Das heißt für Komposita, dass die beiden Konstituenten zu einem unteilbaren Ganzen verschmelzen, wobei nicht immer nur Anfang und Ende erhalten bleiben, z. B. Baunsne für Bau+steine⁴, Antu⁵ für Hand+tuch (2;2), sondern auch zusätzliche Umformungen möglich sind, z. B. Medwa für Mineral+wasser³.

In der **protomorphologischen Phase** kommt es zur Ableitung erster morphologischer Regeln auf Grundlage der bereits erworbenen Lexeme, und in weiterer Folge zu (Über)Generalisierungen und Analogiebildungen. Die Kinder beginnen die Regeln, bzw. identifizierten Muster, kreativ anzuwenden. Sie erwerben die Fähigkeit, Wörter zu zerlegen und zusammensetzen, und bilden Analogien und Neologismen (*bahn-ig gemäß [N + ig] stein-ig, durst-ig). So kommt es etwa bei der Bildung von Partizipien nach dem Vorbild von hüpfen → ge-hüpf-t oder tanzen → ge-tanz-t auf Basis der erkannten Strukturen zu übergeneralisierten Formen wie *ge-geh-t ← gehen oder *ge-ess-t ← essen. Diese regelgeleiteten Mechanismen führen unter anderem zur Entstehung von Mini-

⁴ Jans und Lenas Beispiele entnommen aus: Korecky-Kröll et al. (2017: 26).

⁵ Carmens Beispiel entnommen aus: Rainer (2010: 21).

Paradigmen. Dabei handelt es sich um eine minimale Anzahl von Einheiten, die im Output der Kinder auftreten und ihre Fähigkeit zur Segmentierung komplexer Lexeme widerspiegeln. Erfüllen die Mini-Paradigmen bestimmte Kriterien, lassen sie in weiterer Folge Rückschlüsse auf die potenzielle Produktivität eines Musters in der Kindersprache zu (siehe die Erläuterung zum Mini-Paradigma-Kriterium in Kap. 4.5). Dieser produktive Gebrauch wird als Indiz für den angehenden Erwerb eines Wortbildungsmusters gewertet (siehe für Diminutive: Korecky-Kröll & Dressler 2007; für Komposita: Korecky-Kröll et al. 2017; für Derivationsmorphologie: Sommer-Lolei et al. 2021).

Die dritte und letzte Phase ist gekennzeichnet durch das Auftreten einer Vielzahl morphologischer Regeln, um der steigenden Komplexität der Formen gerecht zu werden. Diese Phase wird als *Morphology Proper* bezeichnet oder nach Peters (2001: 236) auch „(Full) morphology“ genannt. Sämtliche grundlegenden morphologischen Muster der Zielsprache entwickeln sich und nähern sich der Erwachsenenmorphologie an, sodass mit der Zeit das gesamte morphologische System erworben wird. Dazu trägt insbesondere in den frühen Phasen des Spracherwerbs die an das Kind gerichtete Sprache (CDS) bei, da sie die wesentlichen morphologischen Merkmale der Zielsprache enthält. Jene zentralen Aspekte der Morphologie des elterlichen Inputs, die besonders einheitlich, frequent, regelmäßig und transparent sind, werden am leichtesten vom Kind aufgenommen. Sie beinhalten somit den entscheidenden Kern der Morphologie, weshalb sie als Kernmorphologie (*Core Morphology*) bezeichnet werden (vgl. Ravid et al. 2008: 26). Angenommen wird ein Erwerb der Basiselemente der Zielsprachenmorphologie vor einem Alter von 3 Jahren (vgl. Dressler et al. 2010: 326; Korecky-Kröll 2011: 580, 583). Basierend auf dieser Annahme, haben die in dieser Arbeit untersuchten Kinder des INPUT-Projekts (siehe Kap. 3.2) die Schwelle zur *Full Morphology* überschritten und alle wesentlichen morphologischen Hauptaspekte der Morphologie des Deutschen erworben. Ob diese Entwicklung auch für die Derivationsmorphologie bei Substantiven, Verben und Adjektiven beobachtet werden kann, bzw. welche Teilbereiche tatsächlich als potenziell produktiv oder gar als erworben gelten können, wird in Kap. 5.6 eingehend untersucht.

2.4.2 Morphologische Natürlichkeitstheorie

Das Konzept der Natürlichen Morphologie entwickelte sich aus Stampes (1969) Theorie zur Natürlichen Phonologie und deren weiterer Ausarbeitung in Donegan und Stampe (1979), sowie auf der Basis von Theorien zur Markiertheit des Prager Strukturalismus (siehe Trubetzkoy 1939; Jakobson 1971). Die morphologische Natürlichkeitstheorie gilt neben der Natürlichen Phonologie, sowie der Natürlichen Textlinguistik, der Natürlichen Syntax und ihrer jeweiligen Schnittstellen (Morphosyntax, Morphonologie, Morphonotaktik) als Teildisziplin der Natürlichen Linguistik (vgl. Dressler 2009: 33 f.). Die Natürliche Morphologie wurde von Dressler, Mayerthaler, Panagl und Wurzel (1987) basierend auf mehreren ihrer Vorarbeiten zur Natürlichkeit begründet. Für eine kurze Einführung mit Verweisen zu den einzelnen Arbeiten siehe Dressler (2006: 1 f.), für eine umfassendere Zusammenfassung siehe Kilani-Schoch und Dressler (2005) oder in jüngerer Vergangenheit Dressler und Kilani-Schoch (2016: 356 f.).

Der Terminus „natürlich“ ist nach Dressler (2005: 267) im Fall der Natürlichen Morphologie gleichzusetzen mit kognitiv simpel, einfach zugänglich (speziell für Kinder), elementar und daher universell in den Sprachen bevorzugt, wobei anzumerken ist, dass es sich bei der Natürlichkeit um ein graduelles, relatives Konzept handelt. Die Natürlichkeit ist dabei als Überbegriff zu verstehen, dem folgende drei Subtheorien hierarchisch untergeordnet sind (a) eine Theorie der universellen Markiertheit systemunabhängiger morphologischer Natürlichkeit, (b) eine Theorie der typologischen Adäquatheit und (c) eine Theorie der sprachspezifischen Systemadäquatheit bzw. der systemabhängigen Natürlichkeit. Die Theorie der Natürlichen Morphologie unterscheidet daher verschiedene Arten von Natürlichkeit und setzt sie zueinander in Beziehung (vgl. Dressler 2005: 281).

In den frühesten Phasen des Morphologieerwerbs (siehe auch Kap. 2.4.1 zur Prämorphologischen Phase) zeigt sich das Fehlen typologischer oder sprachspezifischer Adäquatheit daran, dass sich universelle Präferenzen durchsetzen, also von kleinen Kindern bevorzugt werden. Dem Erwerb morphologischer und grammatischer Regeln einer Sprache geht immer erst eine Phase extragrammatischer Bildungen voraus, was sich beispielsweise an Reduplikationen zeigt (vgl. Dressler & Karpf 1995; Dressler 2005: 281).

Die Theorie der universellen Markiertheit ist eine Präferenztheorie, die auf Basis einzelner Präferenzparameter ableitet, was universell bevorzugt und somit natürlich ist. Es handelt sich um die Präferenz für (a) Ikonizität, (b) Indexikalität, (c) morphosemantische und morphotaktische Transparenz, (d) Uniformität, (e) Figur-Grund, (f) Binarität und (g) die optimale Form der Einheiten (vgl. Dressler 2005: 268–276), die im Folgenden kurz beschrieben werden.

2.4.2.1 Ikonizität und Indexikalität

Wesentlich für die Wortbildung ist die konstruktionelle Ikonizität, die nach Mayerthaler (1981) als wichtigster Subparameter der morphologischen Ikonizität zu nennen ist und auf dem Konzept der Ikone und deren Klassifikation von Peirce (1965) basiert (siehe dazu auch Korecky-Kröll 2011: 12 f.). Je ikonischer ein bestimmtes Wortbildungsmuster ist, desto natürlicher ist es. Daraus ergibt sich für die Derivation die folgende Reihenfolge (mit abnehmender Ikonizität bzw. Natürlichkeit): Affigierung > (Stamm)Modifikation > Konversion > Subtraktion (vgl. Dressler 2005: 269). Die gleiche Reihenfolge ergibt sich ebenfalls aufgrund des Präferenzparameters der Indexikalität, wo Nähe gegenüber Distanz bevorzugt wird. Die offensichtliche Nähe eines Wortstamms oder Lexems zu seinem Affix weist daher eine höhere Indexikalität auf und ist dadurch natürlicher, als die Änderung des Stammvokals im Zuge der Modifikation oder gar eine gänzlich fehlende Veränderung wie bei der Konversion.

2.4.2.2 Morphosemantische und morphotaktische Transparenz

Auch die morphosemantische und morphotaktische Transparenz spielen als Präferenzparameter in der Wortbildung eine Rolle. Ein morphologisch komplexes Lexem ist dann morphosemantisch transparent, wenn die Bedeutung der einzelnen Morpheme in dem gebildeten komplexen Wort erhalten bleibt. Dies ist beispielsweise in Weck-er der Fall, wo sowohl die Bedeutung der Basis weck- erhalten bleibt, als auch das Ableitungssuffix -er seine Bedeutung ‚Instrument zum X-en‘ beiträgt. Kratz-er ist durch seine Polysemie, also die fehlende Uniformität (siehe Kap. 2.4.2.3), bereits weniger transparent, da es sich einerseits um ein Instrument zum Kratzen, als auch andererseits um den Akt oder das Resultat des Kratzens handeln kann. Im Falle von Komposita ist Eis+kratz-er in seiner Bedeutung transparent, Wolke-n+kratz-er hingegen opak, da es sich um ein sehr hohes Gebäude handelt, das nur im übertragenen Sinn an den Wolken kratzt. Für

gewöhnlich sind Komposita morphosemantisch transparenter als Derivationen, da sie durch ihre Konstruktion auch ein Mehr an Inhalt mitbringen, das für ein schnelles Auffassen der Bedeutung wesentlich ist. Dies ist mitunter einer der Gründe, weshalb Komposita im Erstspracherwerb des Deutschen vor Derivationen auftreten.

Morphotaktische Transparenz ist nach Dressler (2005: 272) dann gegeben, wenn die Einfachheit der Perzeption einer morphologisch komplexen Form nicht durch (mor)phonologische oder allomorphische Prozesse behindert wird. Rein phonologische Prozesse wie z. B. bei Resyllabifizierung im Zuge der Affigierung wie in Garten → Gärt-ner erschweren die Wahrnehmung kaum, während morphologische Bildungsregeln den Grad der morphotaktischen Transparenz deutlich reduzieren, wie bei explodieren → Explosion. Morphotaktische Opazität ist im Fall von Suppletion gegeben, z. B. drei → dritt.

2.4.2.3 Uniformität

Im Rahmen des Präferenzkriteriums der Uniformität ist eine Einheit dann am natürlichsten, wenn sie eineindeutig ist. Ein Morphem, das nur in einer Funktion vorkommt, ist demnach natürlicher als ein Morphem, mit dem sich viele verschiedene Bedeutungen verbinden lassen. Polysemie kommt in den flektierenden Sprachen häufig vor, was zwar sprachökonomisch, allerdings nicht eindeutig ist (vgl. Elsen 1999: 18), wie z. B. mit dem Suffix -er deverbal abgeleitete Substantive können Nomina Agentis wie Fahr-er, Nomina Instrumenti wie Schläg-er oder auch Nomina Acti wie Kratz-er sein. Daher ist die Uniformität im Deutschen und anderen flektierenden Sprachen generell geringer als in agglutinierenden Sprachen.

2.4.2.4 Figur-Grund-Präferenz, Binarität und optimale Form der Einheit

Bei der Figur-Grund-Präferenz werden Figur und Grund, die sich klar unterscheiden lassen, bevorzugt. In der Wortbildung stellt der Kopf die Figur und der Nicht-Kopf den Grund dar, weshalb die morphosemantische Transparenz des Kopfes wichtiger als die des Nicht-Kopfes ist. Die Figur-Grund-Unterscheidung ist festgelegt durch die morphosemantische und -taktische Nachrangigkeit des Grundes (Nicht-Kopf) gegenüber der Figur (Kopf). Bei der Suffigierung und der Komposition ist dies die bevorzugte Gegebenheit (vgl. Dressler 2005: 274 f.).

Zweigliedrige Komposita sind aufgrund der Präferenz für Binarität ebenfalls natürlicher als mehrgliedrige oder Phrasenkomposita, dasselbe gilt für Suffixe, die nicht nur aufgrund ihrer binären Struktur natürlicher sind als Präfixe und Zirkumfixe und daher bevorzugt werden. Außerdem begünstigt der *Recency Effect* die bessere Wahrnehmung des Endes eines Wortes, was speziell im frühen Spracherwerb eine Rolle spielt (siehe dazu Kap. 2.6.2). Ebenso erfüllen Prä- und Suffixe die optimale Form der Einheiten, da sie meist nicht länger als eine Silbe (z. B. -er, -ung, -ig, be-, ent-, un-) und zusammen mit der Basis häufig nicht länger als drei Silben sind (vgl. Dressler 2005: 276).

2.4.2.5 Zusammenfassung

Als morphologisch natürlich gelten Suffixe, da sie an einer gut wahrnehmbaren Position stehen, meist nicht länger als eine Silbe sind, in unmittelbarer Nähe zur Basis stehen und morphotaktisch transparent sind. Präfixe und auch Zirkumfixe sind dagegen weniger natürlich bzw. markierter, da Präfixe unbetont und daher im Spracherwerb schwer wahrnehmbar sind, fallen sie oft der Aphärese zum Opfer. Eine geringere Natürlichkeit weisen Modifikationen des Wortstamms auf, da die Veränderung allein wortintern stattfindet. Konversionen und Subtraktionen gelten, aufgrund der nicht vorhandenen Veränderung oder sogar Kürzung, als am wenigsten natürlich.

Der Theorie zufolge sollten Kinder jene morphologischen Regeln und Muster einer Sprache früher erkennen und auch erwerben, die nach den oben genannten universellen Präferenzen und Parametern natürlicher sind (vgl. Dressler 2005: 278). Diese semiotischen Parameter sind gegenüber anderen Einflussfaktoren, wie etwa der Frequenz, zweitrangig (siehe dazu Kap. 2.6). Die Geschwindigkeit der morphologischen Entwicklung in Kindern ist eng an die Sprachtypologie gebunden. Wie eine sprachübergreifende Untersuchung in Laaha und Gillis (2007) zeigt, vollzieht sich die Entwicklung in Kindern, die agglutinierende Sprachen erwerben, schneller im Vergleich zu Kindern, die eine stark oder schwach flektierende Sprache erwerben (vgl. Dressler et al. 2007: 68).

2.5 Grundlegende Definitionen der Spracherwerbsforschung

Um ein Verständnis für die in der (Psycho)linguistik verwendeten Begriffe der Emergenz, der morphologischen Produktivität und der potenziellen Produktivität in CS zu schaffen, sollen die Termini, wie sie für diese Arbeit Anwendung finden, an dieser Stelle definiert und veranschaulicht werden.

2.5.1 Emergenz

Unter Emergenz (*emergence*) wird in der Spracherwerbsforschung das erste Auftreten eines Wortbildungsmusters in der Sprache eines Kindes verstanden. Wann ein bestimmtes Wortbildungsmuster im Laufe der Entwicklung auftritt, ist von vielen Faktoren abhängig, wie etwa ob ein Kind eher ein Früh- oder Spätentwickler ist, bereits ältere Geschwister hat, mehr oder weniger sprachlichen Input durch seine Hauptbezugsperson(en) und Zugang zu Medien erhält. Da der Spracherwerb also in hohem Maße individuell verläuft, können bei einem Vergleich von mehreren Kindern ausschließlich Tendenzen festgestellt werden.

2.5.2 Produktivität

Der Terminus der Produktivität (*productivity*) bezieht sich im Bereich der Wortbildung auf die Verwendung in einer Sprache bereits vorhandener Wortbildungsprozesse zur Bildung neuer Lexeme. Anders ausgedrückt gilt ein Muster als morphologisch produktiv, wenn die Sprachbenutzer:innen auf ein ihnen durch andere Wortformen bereits bekanntes Muster zurückgreifen, um damit neue Wörter zu bilden. Dies kann in der Kindersprache beispielsweise dann der Fall sein, wenn das Kind ein bestimmtes Wort noch nicht kennt und aus der Notwendigkeit heraus spontan ein Neues kreiert, wie „Tanker“⁶ für eine ‚Person, die tankt‘, oder *bahnig für ‚zur Bahn gehörig; etwas, das zum Zug/zur Bahn gehört‘⁷. Derartige Neubildungen weisen auf eine produktive Verwendung (*productive use/usage*) eines dem Kind bereits zugänglichen bzw. bekannten Wortbildungsmusters hin. Wortneuschöpfungen sind daher ein Zeichen für die Produktivität eines Musters, jedoch nicht notwendigerweise gleichermaßen ein Beleg für die potenzielle Produktivität in CS. Für gewöhnlich wird die morphologische Produktivität in CDS

⁶ Beispiel entnommen aus: Elsen (1999: 174).

⁷ Beispiel aus den longitudinalen Daten von Jan im Alter von 3 Jahren.

von einer hohen Lemma- bzw. Typenfrequenz begleitet (vgl. Lieven 2010). Nicht zu verwechseln ist die morphologische Produktivität jedoch mit Lemma- und Tokenfrequenz, da (un)produktive Wortbildungsmuster in einem Korpus mehr oder weniger häufig sein können.

2.5.3 Potenzielle Produktivität in CS

Die potenzielle Produktivität in CS wird häufig anhand des Auftretens erster Neologismen festgemacht, wie dies in Tagebuchstudien der Fall ist (vgl. Neugebauer-Kostenblut 1914; Stern & Stern 1928; Elsen 1999; Rainer 2010). In dieser Arbeit stellt die Verfasserin die potenzielle Produktivität eines Derivationsmusters in CS anhand der Erfüllung des Mini-Paradigma-Kriteriums fest. Das Kriterium wurde von Dressler et al. (2003) kreiert, als Möglichkeit der Feststellung potenziell produktiver Muster in der Kindersprache. Darauf folgte die Anwendung des Kriteriums in Studien zur Verbalflexion in Bittner et al. (2003) und Kilani-Schoch und Dressler (2005), zur Komposition von Dressler et al. (2010) und Korecky-Kröll et al. (2017) und zur Derivationsmorphologie von Sommer-Lolei et al. (2021: 116 f.). Demzufolge gilt ein Wortbildungsmuster als potenziell produktiv, wenn ein Kind dieses in mindestens drei Lexemen produktiv verwendet sowie die Basis des gebildeten Wortes in drei weiteren Worttypen Anwendung findet (für eine genaue Erläuterung mit Beispiel siehe Kap. 4.5). Das alleinige Auftreten eines komplexen Wortes oder dessen Häufigkeit im Output reichen daher nicht aus, um den Produktivitätsstatus eines Musters im kindlichen Sprachsystem beurteilen zu können.

2.6 Einflussfaktoren auf den Morphologieerwerb

Es gibt zahlreiche Faktoren, die im Zuge des Sprachlernprozesses eines Kindes eine Rolle spielen bzw. dessen Verlauf beeinflussen können. Dazu zählen einerseits sämtliche Fähigkeiten, die das Kind mitbringt, wie Aufmerksamkeitsspanne, Intelligenz und individuelle Eignung bzw. Fertigkeiten, andererseits Faktoren, die von außen auf das System einwirken. Dazu zählen der sozioökonomische Status der Familie des Kindes, der morphologische Reichtum der zu erlernenden Sprache, sowie der zur Verfügung stehende Input durch die Hauptbezugsperson. Generell hat sich in verschiedenen Studien dazu gezeigt, dass die Quantität und Qualität des

Inputs den lexikalischen Erwerb maßgeblich beeinflusst (siehe Hart & Risley 1995; Hoff 2002). In Kauschke und Klann-Delius (2007) wurden die charakteristischen Merkmale der CDS (Frequenz, Diversität, unterstützende Funktionen) im Deutschen mit einem Fokus auf Wortklassen untersucht, mit dem Ergebnis, dass Kinder auf Basis des zur Verfügung stehenden Inputs ausreichende Informationen erhalten, um ihr sprachliches Wissen in einem dynamischen Prozess immer weiter auszubauen. Ein Vergleich von Substantiven und Verben in schwach und stark flektierenden, sowie agglutinierenden Sprachen konnte einen positiven Einfluss der Sprachtypologie auf die Entwicklungsgeschwindigkeit nachweisen, da ein großer morphologischer Reichtum im Input zu einer schnelleren Entwicklung von Paradigmen bei den Kindern führte (Dressler et al. 2007: 67; Laaha et al. 2007: 31 f.). In Bezug auf den Erwerb und die Erweiterung des Wortschatzes, also auf die lexikalische Entwicklung, sind daher auf Wort- und Morphemebene folgende Faktoren zu nennen: Frequenz bzw. Häufigkeit im Input (Kap. 2.6.1), positionelle und prosodische Salienz (Kap. 2.6.2, 2.6.3), morphosemantische und -taktische Transparenz (Kap. 2.6.4), morphologische Komplexität (Kap. 2.6.5) und morphologische Produktivität (siehe Kap. 2.5.2).

2.6.1 Frequenz

Die Frequenz bezeichnet die Verwendungshäufigkeit eines Lexems oder Wortbildungsmusters, also dessen Vorkommen in einem Korpus. Dabei ist zwischen Lemmafrequenz, Typenfrequenz und Tokenfrequenz zu unterscheiden. Wie bereits erwähnt, kann ein Wortbildungsmuster unabhängig von seiner morphologischen Produktivität (siehe Kap. 2.5.2) mehr oder weniger oft in einem Korpus vorkommen.

Die Lemmafrequenz, die in dieser Arbeit neben der Tokenfrequenz herangezogen wird, ist die Gesamtanzahl der Lexikoneinträge (= Lemmas), z. B. KIND. Dabei wird jedes Lemma einmal gezählt. Die Typenfrequenz beinhaltet hingegen die verschiedenen Wortformen eines Lemmas. Als Wortform wird die flektierte Form eines Wortes bezeichnet, „die für bestimmte grammat. Merkmale gekennzeichnet ist“⁸, wie z. B. für die Pluralform in Kind-er ‚Kind-PL‘ oder für die Form des Genitivs Singular in Kind-es ‚Kind-GEN.SG‘. Die Tokenfrequenz gibt die Summe

⁸ Glück (2010, s. v. Wortform).

des Auftretens eines Lemmas wieder. In der vorliegenden Untersuchung analysiert die Verfasserin die Lemma- und Tokenfrequenzen (vgl. dazu Kap. 4.4.4, S. 83).

Die Frequenz im Input ist ein zentraler Aspekt, der Einfluss nimmt auf die lexikalische Entwicklung eines Kindes. In den meisten Untersuchungen wurde dies anhand der Lemmafrequenz⁹ festgestellt.

In Bezug auf die Entwicklung des Wortschatzes im Rahmen der Sprachproduktion konnten Pan et al. (2005) für das Englische zeigen, dass die lexikalischen Wachstumsmuster in Kindern mit der Lemmafrequenz des Inputs, also ihrer jeweiligen CDS, korrelierten. Ebenso zeigte sich ein Zusammenhang zwischen der Lese- und Schreibkompetenz der Hauptbezugsperson zu ihrer Lemmafrequenz. Für das Deutsche konnte dies in der Untersuchung von Kauschke und Klann-Delius (2007: 199) bestätigt werden, da sich auch hier ein Zusammenhang zwischen der Sprachproduktion der Kinder und dem lexikalischen Reichtum des Inputs zeigte. Die Autorinnen schließen daraus, dass ein variiertes Input im Sinne einer höheren Lemmafrequenz für die Abweichungen in CS verantwortlich ist.

Demnach führt ein vielfältiges und abwechslungsreiches lexikalisches Angebot im Input in Form unterschiedlicher Lemmas bzw. ihrer Verwendung in verschiedenen Wortformen (= Typen) durch die Hauptbezugsperson, zu einer höheren Verwendung in CS. Ebenso führt jede zusätzliche Information, die der Kontext einem Lexem beisteuert, zu einem genaueren Verständnis seiner Bedeutung(en) (siehe Roeper 2007: 30). Höhere Variation im Input trägt damit zu einer Erweiterung des Wortschatzes bei und begünstigt den Erwerb zusätzlicher (abstrakter) Bedeutungen.

Häufigkeiten im Input allein sind jedoch, wie sich gezeigt hat, kein verlässlicher Prädiktor für das erste Auftreten und die Reihenfolge der Emergenz bestimmter Wortbildungsmuster in CS (vgl. Sommer-Lolei et al. 2021: 131). Allerdings spielt die Frequenz im Spracherwerb eine Rolle in Bezug auf eine kritische Masse in CDS (siehe Marchman & Bates 1994; Van der Schuit et al. 2011; Serra i Raventós 2014).

⁹ In englischsprachigen Publikationen wird häufig der Terminus *type frequency* (Typenfrequenz) synonym zu Lemmafrequenz verwendet.

2.6.2 Positionelle Salienz

Unter positioneller Salienz ist das Hervorstechen oder die Auffälligkeit eines Elementes zu verstehen in Abhängigkeit von seiner Position. Anders ausgedrückt werden Einheiten, die am Beginn oder Ende eines Wortes oder einer Äußerung stehen, leichter bzw. besser wahrgenommen und sind daher salient. Dies ist für den Spracherwerb im Bereich der Wortbildung von entscheidender Bedeutung, da die Wortbildungsmuster oder Affixe aufgrund ihrer Position im Wort mehr oder weniger auffällig sind und daher früher oder später vom Kind erkannt und verwendet werden. Insgesamt beeinflusst die Anordnung von Wörtern oder Wortbildungsmustern in den Äußerungen im Input des Kindes (CDS) dessen Spracherwerb (vgl. Longobardi et al. 2015: 97). Steht die sozusagen vorteilbringende Einheit an erster Stelle innerhalb einer Äußerung oder eines Wortes wird vom *Primacy Effect* gesprochen, steht sie an letzter Stelle, handelt es sich um den *Recency Effect* (vgl. Cornell & Bergstrom 1983: 494). Beide Effekte sollen in der Folge kurz erläutert werden:

Die ursprünglich aus der Psychologie stammende Bezeichnung *Recency Effect* bezieht sich im Allgemeinen auf Untersuchungen zu jenem Effekt, der sich beim auswendigen, freien Rezitieren einer zuvor präsentierten Liste zeigt. Die letzten Einträge einer Liste, unabhängig davon, ob es sich um Zahlen oder Wörter handelt, sind für gewöhnlich besser abrufbar als die ersten. Die in der Mitte befindlichen Einträge sind am wenigsten zugänglich (vgl. Eysenck & Keane 2010: 207 f.). Was also unmittelbar zuvor präsentiert wurde, bleibt besser im Gedächtnis. Der *Recency Effect* fördert die Wahrnehmung der letzten Einheit in einer Äußerung oder einem Lexem. Dies bedeutet für die Wortbildung einen begünstigenden Effekt für den Erwerb von Suffixen im Vergleich zu anderen Wortbildungsmustern an weniger salienten Stellen. Aus syntaktischer Perspektive zeigt sich der *Recency Effect* beispielsweise daran, dass die am Ende einer Äußerung, also in satzfinaler Position in CDS vorkommenden trennbaren Partikel von Partikelverben in CS zu Beginn des Spracherwerbs isoliert auftreten. Die Kinder produzieren „rauf!“ statt „Heb mich rauf“ oder „zu!“ statt „Ich mache das zu“. Neben ihrer günstigen Position spielt jedoch bei trennbaren Partikeln auch die prosodische Salienz eine Rolle, da sie betont sind (siehe Kap. 2.6.3).

Der *Primacy Effect* ist ein ähnliches Phänomen, das in Bezug auf die lexikalische Speicherung und den Abruf von Wörtern im mentalen Lexikon bei Erwachsenen festgestellt wurde. Der sogenannte Badewanneneffekt (*bathtub effect*) besagt, dass sich erwachsene Menschen im Schnitt wesentlich besser an den Beginn als an das Ende eines Wortes bzw. einer Äußerung erinnern, und oft weniger gut bzw. kaum an dessen Mittelteil (vgl. Aitchison 1994: 134 ff.). Der Effekt zeigt sich insbesondere in der Sprachverarbeitung bei Erwachsenen an der besseren Wahrnehmbarkeit von Wortanfängen, beispielsweise Präfixen, obwohl diese häufig unbetont sind (siehe Kap. 2.6.3).

2.6.3 Prosodische Salienz

Neben der positionellen Salienz spielt auch die prosodische (auch perzeptuelle) Salienz im Spracherwerb eines Kindes eine Rolle. Wie bereits in Kap. 2.6.2 erwähnt, sind sämtliche Merkmale salient, die sich abheben bzw. hervorstechen. Daher sind betonte Elemente (Silben, Morpheme) prosodisch salient und werden gegenüber unbetonten, nichtsalienten Einheiten bevorzugt, das heißt schneller und akkurater aufgefasst (vgl. Gillis 2003: 196 f.; Dressler 2007: 8), was auf alle Altersgruppen gleichermaßen zutrifft. Für gewöhnlich betont werden im Deutschen trennbare Präfixe, z. B. *wég#gehen*, *án#sehen*; Präfixe zur Bildung von Substantiven, z. B. *Úr-opa*, *Érz-feind*; und einige Suffixe wie *-ier-*, *-ei/-er-ei*. Untrennbare Präfixe zur Bildung von Verben sind hingegen unbetont, z. B. *be-*, *ent-*, *er-*, *ver-*, *zer-*. In Komposita trägt das Erstglied eines Kompositums meist den Hauptakzent, das Letztglied den Nebenakzent, z. B. *Háus+tier*, *Sónnen+stráhl*.

2.6.4 Morphosemantische und morphotaktische Transparenz

Die morphosemantische und morphotaktische Transparenz, wie bereits in Kap. 2.4.2.2 im Rahmen der morphologischen Natürlichkeitstheorie eingehend behandelt, sind auch im Erstspracherwerb relevant. Setzt sich die Bedeutung eines komplexen Lexems aus den jeweiligen Bedeutungen der wortbildungsrelevanten Morpheme (Lexeme, Stamm, Affix) zusammen, ist das Wortbildungsprodukt demnach morphosemantisch transparent, so sind diese Bildungen für ein Kind einfacher zu erfassen und zu verarbeiten. Dies ist beispielsweise bei zweigliedrigen Komposita meist der Fall. Bei komplexen Wörtern, die morphotaktisch transparent sind, handelt es sich um Wortbildungen, die einfach wahrgenommen werden

können und deren Perzeption nicht durch (mor)phonologische oder allomorphische Regeln eingeschränkt wird. Dabei sind Affigierungen, z. B. Les-er, Heiter-keit, witz-ig, morphotaktisch transparenter als Modifikationen, wie purer Ab- oder Umlaut des Stammvokals bei impliziten Derivationen, z. B. Flug, fäll-en.

2.6.5 Morphologische Komplexität

Neben den bereits behandelten Faktoren ist die morphologische Komplexität eines Lexems im Zuge des Erstspracherwerbs wesentlich. In dieser Arbeit legt die Autorin die Komplexität im In- und Output anhand der Anzahl der zur Bildung eines Kompositums oder Derivats nötigen Wortbildungsprozesse fest. Die morphologische Komplexität erhöht sich demnach mit der Anzahl der benötigten Wortbildungselemente. Dabei wird zwischen der Komplexität der Basis und des gebildeten Lexems differenziert. Zur Komplexität in der Wortbildung, sowie einer detaillierten Beschreibung der morphologischen Komplexitätsgrade, wie sie in dieser Untersuchung Anwendung finden, siehe Kapitel 7 und 7.1, sowie Tabelle 5.

Kinder produzieren zunächst mehr einfache als derivierte oder zusammengesetzte Lexeme. Derivationen, die mehr als ein Affix aufweisen, treten bis zu einem Alter von 3 Jahren nur vereinzelt auf (vgl. Sommer-Lolei et al. 2021: 132), ebenso wie mehrgliedrige Komposita (vgl. Korecky-Kröll et al. 2017: 32) und die besonders komplexen synthetischen Komposita, die Ableitung und Komposition in sich vereinen (siehe Dressler, Sommer-Lolei, et al. 2019). Die Verwendung komplexer Lexeme, insbesondere die Häufigkeit abgeleiteter Verben, nimmt im Erwerbsverlauf im Zeitraum zwischen 1;9 und 3;0 stetig zu, was Studien von Schipke und Kauschke (2011) und Sommer-Lolei et al. (2021) belegen. Für den Erwerbsverlauf derivierter Substantive bei Jan (1;3–4;11) und den INPUT-Kindern siehe Kap. 5.3.2, für den Verlauf derivierter Verben siehe Kap. 5.4.2. Für die zunehmende morphologische Komplexität im Erwerbsverlauf siehe Kap. 7.2.

2.6.6 Erwartete Zusammenhänge

Basierend auf den in den vorangehenden Kapiteln 2.6.1–2.6.5 ausgeführten Annahmen sollten sich für die Wortbildung in CS die folgenden Präferenzen und Zusammenhänge zeigen: (a) die lexikalische Varietät im Input (hohe Lemmafrequenz in CDS) führt zu einer verbesserten Aufnahme und Abrufbarkeit

in CS, was eine höhere Verwendungsrate und eine frühere Erweiterung des Wortschatzes mit sich bringt, (b) hohe Tokenfrequenzen im Input sind nicht so wesentlich für den Spracherwerb wie hohe Lemmafrequenzen, (c) wortfinale Elemente (Suffixe) werden aufgrund des *Recency Effect* bevorzugt, (d) betonte Affixe werden bevorzugt aufgrund der prosodischen Salienz, (e) die morphosemantische und -taktische Transparenz eines Lexems begünstigt den Erwerb, (f) morphologisch komplexe Lexeme und Basen treten später in CS auf als einfache.

2.7 Fragestellungen und Annahmen

Im Folgenden sollen die Forschungsfragen und Hypothesen, denen im Zuge der quantitativen und qualitativen Auswertungen und Untersuchungen nachgegangen wird, beschrieben werden.

2.7.1 Forschungsfragen

Vorangegangene Untersuchungen (siehe Korecky-Kröll et al. 2017) haben gezeigt, dass Komposita im Erstspracherwerb aufgrund des großen Reichtums an zusammengesetzten Lexemen im Deutschen generell früh in der Kindersprache auftreten, erworben und kreativ verwendet werden. Wie sieht es jedoch im Alter von 3–5 Jahren aus, wenn vermehrt andere Bildungen hinzukommen? Daher stellt sich zunächst die Frage nach der Verteilung der Derivationen und Komposita in den einzelnen Korpora und wie die einzelnen Wortbildungsmuster in CS und CDS verteilt sind. Welche Unterschiede zeigen sich zwischen jüngeren (bis 3 Jahre) und älteren Kindern (bis fast 5 Jahre) im Hinblick auf die Verwendung abgeleiteter und zusammengesetzter Substantive, Verben und Adjektive? Und wie verhält es sich dabei mit Unterschieden zwischen Kindern mit höherem oder niedrigerem sozioökonomischen Status? Welche Fehler finden sich in derivierten Lexemen und Komposita in der Spontansprache und in dem durchgeführten Elizitationsexperiment?

Welche Einflussfaktoren auf den Morphologieerwerb lassen sich als maßgeblich feststellen? Beschleunigen Diversität, Frequenz, positionelle und prosodische Salienz, morphosemantische und morphotaktische Transparenz in CDS den Erstspracherwerb und die Sprachverarbeitung?

Eine weitere Frage, die es zu klären gilt, bezieht sich auf den Wortartwechsel im Zuge der Wortbildung. Sind die Derivationsprodukte in CS und CDS eher wortartbewahrend oder wortartverändernd? Treten wortartverändernde Ableitungen früher in CS auf als wortartbewahrende?

Der häufig prognostizierte stete Anstieg der morphologischen Komplexität im Erwerbsverlauf soll anhand der Daten der Kinder bis knapp 5 Jahre untersucht und festgestellt werden. Es wird der Frage nachgegangen, wie komplex die gebildeten Lexeme und deren Basen sind und wann die ersten komplexen Formen in CS auftreten. Überdies soll geklärt werden, welche Wortarten bevorzugt als Basen dienen. Dabei wird die Komplexität nicht durch die Anzahl der Morpheme bestimmt, sondern durch die Anzahl der Wortbildungsprozesse, die zur Bildung der Derivate nötig sind (siehe dazu Kap. 7).

In Bezug auf das erste Auftreten einer morphologisch komplexen Form und der Feststellung der potenziellen Produktivität eines Wortbildungsmusters wird die Erfüllung des Mini-Paradigma-Kriteriums (MPK), wie in Sommer-Lolei et al. (2021: 127–128, 130) bis 3 Jahre erfasst, mit dem Kind Jan bis zu einem Alter von 4;11 fortgeführt, um bestehende Lücken zu schließen. Daher soll die Frage beantwortet werden, ob bzw. wann Jan das MPK für suffigierte Verben und Adjektive erfüllt.

Eine weitere wesentliche Forschungsfrage ist, ob sich Korrelationen zwischen den Daten der Elterninterviews, also zwischen der Inputmenge der Hauptbezugspersonen oder dem Medienkonsum (YouTube, Internet, Smartphone, Fernsehen; Hörbuch und Radio; Büchervorlesen) zu den Ergebnissen des Derivationsexperiments oder der Spontansprachanalysen zeigen.

2.7.2 Hypothesen

H1 Der sozioökonomische Status der Familie beeinflusst die Entwicklungsgeschwindigkeit der Derivationsmorphologie des Kindes

Es werden sich Unterschiede im sprachlichen Output der Kinder mit höherem sozioökonomischen Status (HSES) und mit niedrigerem sozioökonomischen Status (LSES) als auch der HSES- und LSES-Hauptbezugspersonen zeigen. Kinder und

Hauptbezugspersonen mit höherem sozioökonomischen Status verfügen über mehr Variation in der Sprache und kommunizieren in einem größeren Ausmaß als dies bei LSES-Kindern und -Hauptbezugspersonen der Fall ist. Das heißt sie verwenden sowohl mehr Lemmas als auch Tokens abgeleiteter Lexeme (Derivationen) in allen drei untersuchten Wortarten, verwenden morphologisch komplexere Lexeme und Basen und zeigen insgesamt eine größere Vielfalt an Ableitungsmustern. Ebenso werden die HSES-Kinder im Rahmen des Derivationstests mehr korrekte Derivationen bilden.

H2 Wortbildungsmuster, die morphologisch natürlicher sind, werden in CS bevorzugt verwendet und entwickeln sich früher und schneller.

Aufgrund des begünstigenden Effekts der positionellen Salienz, insbesondere des *Recency Effect*, der prosodischen Salienz und verschiedener Natürlichkeitsparameter (Ikonizität, Binarität, morphotaktische Transparenz, optimale Form der Einheit) werden bestimmte Wortbildungsmuster (insbesondere Suffixe) von Kindern bevorzugt und scheinen daher früher in CS auf und werden auch häufiger verwendet. Wortbildungsmuster, die hingegen weniger natürlich sind, wie Konversion oder Modifikation, kommen weniger häufig vor und treten später in CS auf.

H3 Die Menge des Inputs durch die Hauptbezugsperson und die Menge des Medienkonsums wirken sich auf die Sprachproduktion der Kinder aus.

Der Input durch die Hauptbezugsperson, unabhängig von ihrem sozioökonomischen Status, der stets größer werdende Einfluss der Medien in der heutigen Zeit und der zunehmende Medienkonsum, gemessen in Minuten pro Woche, insbesondere Neuer Medien, wie YouTube, Internetnutzung und Smartphone (YIS), der Massenmedien Fernsehen (F) und Radio (R) und digitaler Medien wie Hörbücher (HB), sowie von Printmedien durch Büchervorlesen, wirkt sich auf die Spontansprache der Kinder aus, und zeigt sich ebenfalls an den Ergebnissen des 4 Jahre später durchgeführten Derivationsexperiments.

3 Spontansprachdaten

In dieser Arbeit untersucht die Autorin neben experimentellen Daten (wie in Kap. 9.3 beschrieben) vier Spontansprachkorpora (siehe Kap. 3.1 und 3.2), deren Ergebnisse in Relation zu den Daten der Elterninterviews (siehe Kap. 4.2.3) gesetzt werden, um zu sehen, ob sich kurz- und längerfristige Einflüsse des Inputs und Medienkonsums auf die Kindersprache feststellen lassen (siehe dazu Kap. 10).

Die Spontansprachkorpora, die sich in Langzeitdaten und Paneldaten unterteilen, enthalten Sprachdaten von Kindern und ihrer Hauptbezugsperson, die in der Mehrheit der Fälle die Mutter ist. Alle 32 Kinder, die in dieser Arbeit systematisch ausgewertet werden, wachsen monolingual in Wien mit Erstsprache österreichischem Standarddeutsch auf. Sie sind in ihrer geistigen, körperlichen und sprachlichen Entwicklung unauffällig und mit höherem oder niedrigerem sozioökonomischen Status (für die Operationalisierung des SES siehe Kap. 4.3).

Während es sich bei der longitudinalen Datenerhebung um eine Methode handelt, bei der die Kinder über einen längeren Zeitraum hinweg durchgehend aufgenommen werden (1–4-mal/Monat über ca. 1,5 bis 4 Jahre), werden die Paneldaten in regelmäßigen Abständen erhoben, etwa wie jene in dieser Arbeit, an vier aufeinanderfolgenden Zeitpunkten.

3.1 Longitudinale Spontansprachdaten

Für die Analyse werden drei longitudinale Spontansprachkorpora von drei Wiener Kindern herangezogen. Es sind dies der Bub Jan (JAN) und die beiden Mädchen Katharina (KAT) und Lena (LEN). Dabei handelt es sich um die sprachliche Interaktion der Kinder in Verbindung mit ihren Müttern (alle mit höherem sozioökonomischen Status = HSES), deren Spontansprache ab einem Alter von ca. eineinhalb Jahren über einen längeren Zeitraum 1- bis 4-mal pro Monat in unterschiedlichen Alltagssituationen, wie etwa beim Essen, Spielen, Anziehen, Malen, Vorlesen etc., in ihrem vertrauten Umfeld zu Hause von der Mutter mit einem Audioaufnahmegerät aufgenommen wurden. Das analysierte Datenmaterial besteht aus 317.052 Worttokens aus 90,9 Stunden transkribierten Audiomaterials von 188 Aufnahmen (siehe Tabelle 2). Für monatliche Gesamtübersichten, sowie

mittlere Äußerungslängen und Standardabweichungen in CS und CDS siehe für Jan Anhang A3.1 Tabelle 45, für Kathi Anhang A3.2 Tabelle 46 und für Lena Anhang A3.3 Tabelle 47. Für monatliche Übersichten derivierter Substantive, Verben und Adjektive in Lemmas und Tokens in CS und CDS siehe für Jan Anhang A5.1 Tabelle 53, für Kathi Anhang A5.2 Tabelle 54 und für Lena Anhang A5.3 Tabelle 55. Für monatliche Lemma- und Tokenfrequenzen in CS und CDS synthetischer Komposita sowie Nominal- und Adjektivkomposita siehe für Jan Anhang A6.1 Tabelle 59, für Kathi Anhang A6.2 Tabelle 60 und für Lena Anhang A6.3 Tabelle 61.

Tabelle 2. Longitudinale Spontansprachdaten (CS und CDS)

Korpus	Alter ^a	Anz. Äußerungen		Anz. Worttokens		Anz. Aufnahme	Stunden (gesamt)
		CS	CDS	CS	CDS		
JAN	(m/H) 1;3-4;11	19.166	40.849	47.701	187.178	114	57,3
KAT	(w/H) 1;6-3;0	3.417	6.917	6.366	25.660	34	13,6
LEN	(w/H) 1;7-3;0	6.064	11.877	12.087	38.060	40	20,0
Gesamt		28.647	59.643	66.154	250.898	188	90,9

a. Altersangabe im Format Jahr;Monat (z. B. 1;6 = 1 Jahr und 6 Monate).

Die in sämtlichen Tabellen angeführten Zahlen spiegeln nicht die Gesamtkorpora wider, sondern ausschließlich jene Daten, die für die weitere Analyse verwendet werden. Daher handelt es sich bei den in dieser und allen folgenden Tabellen enthaltenen Zahlen um alle spontanen Äußerungen des Kindes (CS) und seiner Mutter (CDS). Zitate, direkte Imitationen und an Erwachsene oder Tiere gerichtete Sprache fließen nicht in die Analyse mit ein. Überdies enthalten die Lemma- und Tokenfrequenzen entweder zusammengefasst oder in Wortklassen einzeln angegeben, keine Eigennamen, und hauptsächlich Vollverben (keine Auxiliar- und Modalverben). Für eine detaillierte Beschreibung der Datenauswahl siehe Kapitel 4.4.2.

Wie in Tabelle 2 ersichtlich, enden die Aufnahmen der beiden Kinder Lena und Katharina im Alter von 3 Jahren, während Jans Aufnahmen bis knapp vor seinen 5. Geburtstag reichen. Insgesamt existieren von Jan und seiner Mutter in manchen Monaten mehr Aufnahmen, was zwangsläufig zu einer Mehrzahl an Daten führt. Dieses Ungleichgewicht zwischen den Korpora stellt für die statistische Auswertung kein Problem dar, da in der Analyse die prozentualen Anteile und Mittelwerte verglichen werden.

3.2 Paneldaten

Zusätzlich zu dem longitudinalen Datenmaterial analysiert die Verfasserin die Spontansprache von 29 typisch entwickelten, mit österreichischem Standarddeutsch aufwachsenden, einsprachigen Wiener Kindergartenkindern und ihrer jeweiligen Hauptbezugsperson. Unter Hauptbezugsperson ist jene erwachsene Person zu verstehen, die mit dem Kind die meiste Zeit verbringt. Die Paneldaten sind das Ergebnis des INPUT-Projekts¹⁰, die aus 58 Stunden transkribiertem Audio- und Videomaterial bestehen, das zu Hause im gewohnten Umfeld der Kinder von der Hauptbezugsperson mit einem Audioaufnahmegerät sowie zusätzlich mit einer im Raum platzierten Videokamera in Anwesenheit einer Experimentatorin auf Audio und Video aufgenommen wurde. Diese Aufnahmen wurden über einen Zeitraum von beinahe zwei Jahren an vier Zeitpunkten erhoben. Zu den jeweiligen Erhebungszeitpunkten wiesen die Kinder ein Durchschnittsalter von 3;1, 3;4, 4;4 und 4;8 auf. Das Sample aus dem INPUT-Korpus besteht aus 8 Mädchen und 7 Buben mit höherem sozioökonomischen Status (HSES/H) und 8 Buben und 6 Mädchen mit niedrigerem Status (LSES/L).

In Summe beläuft sich das spontansprachliche Material auf 286.623 transkribierte Worttokens. In Tabelle 3 werden für jedes Kind dessen Alter zum jeweiligen Erhebungszeitpunkt, seine SES-Zugehörigkeit, sowie Anzahl der Äußerungen und der Worttokens im Input und Output detailliert dargestellt.¹¹ Für eine genaue Übersicht zu Lemma- und Tokenfrequenzen in CS und CDS, mittleren Äußerungslängen und Standardabweichungen siehe in Anhang A3.4 Tabellen 48 und 49. Für eine Übersicht derivierter Substantive, Verben und Adjektive in Lemmas und Tokens je Datenzeitpunkt in CS und CDS siehe in Anhang A5.4 Tabelle 56 (Gesamt) und je Kind in den Tabellen 57 und 58. Für eine

¹⁰ INPUT steht für “Investigating Parental and Other Caretakers’ Utterances to Kindergarten Children” (SSH11-027) das vom Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF) von 2012 bis 2016 gefördert, und an der Universität Wien von Wolfgang U. Dressler geleitet wurde. In dem Projekt wurden monolinguale Kinder mit L1 Österreichisches Standarddeutsch, sowie bilinguale Kinder mit L1 Türkisch und L2 Österr. Deutsch untersucht. Für weiterführende Informationen zum INPUT-Projekt siehe:

URL: <https://comparative-psycholinguistics.univie.ac.at/projects/input> [08.03.2021].

¹¹ Die in Tabelle 3 angegebenen Zahlen wurden mit dem Programm CLAN ermittelt, daher unterscheiden sich diese von den Lemma- und Tokenfrequenzen in den Tabellen 48 und 49 in Anhang A3.4, die auf den Excel-Daten basieren.

Gesamtübersicht der Nominalkomposita, synthetischer Komposita und Adjektivkomposita in Lemmas und Tokens in CS und CDS siehe Anhang A6.4 Tabelle 62.

Tabelle 3. Paneldaten zu Hause CS und CDS (INPUT-Korpus)

Korpus: INPUT		Erhebungszeitpunkt				Anz. Äußerungen		Anz. Worttokens	
		1	2	3	4	CS	CDS	CS	CDS
D02-SIJ	(m/H)	3;2	3;5	4;5	4;11	1.172	1.994	3.638	7.145
D03-NIP	(m/H)	3;1	3;3	4;5	4;8	1.015	1.895	2.600	6.186
D04-SOS	(w/H)	3;2	3;5	4;5	4;9	1.195	2.005	4.599	7.573
D07-NIG	(w/H)	2;11	3;2	4;2	4;5	1.314	2.201	3.693	8.654
D09-KAB	(w/H)	3;1	3;4	4;4	4;7	1.063	2.001	2.811	6.281
D11-DAG	(m/H)	3;1	3;4	4;6	4;9	953	1.733	3.284	7.192
D13-JOP	(m/H)	3;0	3;3	4;4	4;8	1.467	1.482	4.271	5.461
D15-LUN	(m/H)	3;0	3;3	4;3	4;7	1.271	2.211	3.442	7.801
D16-LEP	(w/H)	3;1	3;5	4;6	4;11	1.120	2.811	3.309	9.062
D20-LUD	(w/H)	3;2	3;6	4;6	4;9	1.637	2.480	4.827	8.710
D23-SOK	(w/H)	2;11	3;3	4;3	4;5	1.459	1.501	4.157	4.610
D27-JOZ	(m/H)	3;0	3;3	4;3	4;7	1.307	1.936	3.861	6.783
D28-MOH	(m/H)	3;0	3;3	4;5	4;8	1.863	1.529	6.146	6.495
D29-SAK	(w/H)	3;3	3;5	4;4	4;9	1.072	2.221	3.390	9.083
D31-AND	(w/H)	3;3	3;7	4;6	4;10	1.197	2.109	4.144	7.954
Summe HSES						19.105	30.109	58.172	108.990
D05-FLB	(m/L)	3;0	3;3	4;4	4;8	1.066	2.395	2.864	8.138
D06-LAB	(w/L)	3;1	3;5	4;5	4;8	1.220	2.218	3.590	7.852
D08-FAJ	(m/L)	3;0	3;4	4;4	4;7	1.548	636	4.255	2.541
D10-LER	(w/L)	3;1	3;4	4;4	4;7	1.129	1.972	2.845	6.553
D12-JAS	(w/L)	3;1	3;4	4;6	4;9	1.155	1.588	2.633	5.069
D14-KEW	(m/L)	3;1	3;4	4;4	4;7	1.151	1.482	2.447	5.225
D17-TOB	(m/L)	3;0	3;4	4;4	4;8	1.462	948	4.023	3.349
D18-ALM	(w/L)	3;0	3;3	4;4	4;7	1.936	1.941	6.170	6.821
D19-VAS	(w/L)	3;1	3;4	4;5	4;7	1.249	1.536	2.896	4.856
D21-PAS	(m/L)	3;3	3;6	4;6	4;9	1.521	1.409	3.784	5.072
D22-LAS	(w/L)	2;11	3;3	4;3	4;5	1.022	1.572	2.946	4.564
D24-RAG	(m/L)	3;0	3;3	4;4	4;7	628	1.388	1.916	5.809
D25-PHG	(m/L)	3;0	3;3	4;4	4;7	844	1.318	2.567	5.389
D26-JUB	(m/L)	3;2	3;6	4;5	4;9	656	1.224	1.667	3.620
Summe LSES						16.587	21.627	44.603	74.858
Gesamt (H+LSES - summiert)						35.691	51.736	102.775	183.848

Für Tabelle 3 wurden alle vier Erhebungszeitpunkte insgesamt ausgewertet und zur Berechnung des Inputs (CDS) ausschließlich die Daten der Hauptbezugsperson herangezogen. In der überwiegenden Mehrheit der Teilnehmer ist dies die Mutter.

Im Fall des zweieiigen Zwillingspärchens (D24-RAG, D25-PHG) ist die Hauptbezugsperson identisch. Bei zwei der 14 LSES-Kinder (D17-TOB, D12-JAS) sind die Väter die jeweilige Hauptbezugsperson, hiervon allerdings ist bei einem Kind (D12-JAS) der LSES-Vater nur in den ersten drei Aufnahmen die Hauptbezugsperson, während in der vierten Aufnahme, aufgrund vorübergehender Trennung, die HSES-Mutter zur Hauptbezugsperson wird. In diesem speziellen Fall wurden daher die Daten des Vaters aus den Aufnahmen 1–3 und jene der Mutter aus Aufnahme 4 entsprechend zusammengefügt, um den tatsächlichen Input abzubilden.

Bei dem Kind D08-FAJ (LSES) zeigt sich zwar ein quantitativ größerer sprachlicher Input durch den Vater, allerdings ist dieser lediglich zu den Erhebungszeitpunkten anwesend. Im Sinne dieser Arbeit, den durch die Hauptbezugsperson an das Kind gerichteten Input zu messen und zu analysieren, und ebenso, um die Daten in dieser Hinsicht nicht zu verfälschen, hat die Verfasserin die Daten der alleinerziehenden Mutter angeführt.

Wie in Tabelle 3 ersichtlich, die die Gesamtsummen aller vier Aufnahmen darstellt, haben die Kinder D28-MOH (HSES), D08-FAJ (LSES), D21-PAS (LSES) und D17-TOB (LSES) mehr Äußerungen als ihre Hauptbezugsperson. In zwei der vier Fälle zeigt sich dies ebenso an einer höheren Worttokenzahl (D08-FAJ, D21-PAS). Dies liegt zum einen an einer möglichen Introvertiertheit der Hauptbezugsperson, wie dies etwa bei FAJs Mutter der Fall ist, zum anderen daran, dass die Mütter ihre Kinder zum Reden ermuntern und in weiterer Folge auch reden lassen, wie dies bei MOHs Mutter der Fall ist, und sich daher etwas mehr zurücknehmen, um dem Kind mehr Raum für Kommunikation zu gewähren.

Von den insgesamt 31 einsprachigen Kindern mit Erstsprache Deutsch (7 m/HSES, 9 w/HSES, 9 m/LSES, 6 w/LSES), die am INPUT-Projekt teilgenommen haben, wertet die Verfasserin 29 Kinder systematisch aus. Diese verteilen sich nach Geschlecht und SES in 8 w/HSES & 6 w/LSES, sowie 7m/HSES & 8 m/LSES, oder 15 HSES & 14 LSES-Kinder (vgl. Tabelle 3), und sind daher nahezu ausgeglichen. Nachdem geschlechterspezifische Unterschiede nicht das vorrangige Thema dieser Arbeit darstellen, sondern eine umfassende Auswertung aller vorhandenen Daten, und in den Analysen Mittelwerte verglichen werden, stellt

dieses leichte Ungleichgewicht kein Problem dar. Das HSES-Mädchen (D01-ODK), das zwar in Wien lebt, aber nicht mit der Wiener Umgangssprache, sondern mit dem Alemannischen aufwächst, habe ich, aufgrund mangelnder Vergleichbarkeit, aus der Analyse ausgeschlossen. Ebenso wie den LSES-Buben (D30-JAR), dessen Hörprobleme in den ersten Lebensjahren, zu einer Verminderung der rezeptiven Wahrnehmung geführt haben, weshalb sich seine Sprachentwicklung verzögert zeigt. Da er deshalb nicht als unauffällig bzw. typisch entwickelt angesehen werden kann, schließt die Autorin ihn ebenfalls aus, auch wenn er an den letzten beiden Erhebungszeitpunkten stetig aufholt.

4 Methoden

Für meine Arbeit kommen eine Reihe von verschiedenen Methoden zur Anwendung, die ineinandergreifen und es ermöglichen, die spontansprachlichen Daten nicht nur detailliert auszuwerten, sondern diese auch den Test- und Interviewdaten gegenüberzustellen. Dadurch eröffnet sich eine völlig neue Perspektive auf die Zusammenhänge von spontanen Bildungen im Output der Kinder zu deren Input durch die Hauptbezugsperson, auch unter Berücksichtigung des Einflusses der Medien in diesem Zeitraum, sowie auf Rückschlüsse auf spätere sprachliche Fertigkeiten. Die folgenden Kapitel geben einen Überblick über die Methoden der Datenerhebungen (Kap. 4.1), der Erstellung, Durchführung und Auswertung der Elterninterviews (Kap. 4.2), der Operationalisierung des sozioökonomischen Status (Kap. 4.3), der Transkription und morphologischen Codierung der Daten (Kap. 4.4), und der Feststellung des Mini-Paradigma-Kriteriums (Kap. 4.5).

4.1 Datenerhebung Spontansprachaufnahmen

4.1.1 Erhebung Longitudinale Spontansprachaufnahmen

Die longitudinalen Spontansprachaufnahmen der drei Kinder entstanden in den Jahren 1993–1995 (KAT), 1999–2001 (LEN) und 1999–2002 (JAN). Die Audioaufnahmen erfolgten mit verschiedenen Kassettenrekordern (z. B. Sony Walkman Professional Stereo Cassette-Corder WM-D6C mit Mikrofon, Dictaphone Voice Processor 2223) und wurden einige Zeit später im mp3- und wav-Format digitalisiert, um das Hintergrundrauschen der originalen

Kassettenaufnahmen zu minimieren und damit ein besseres Sprachverständnis für die anschließende Transkription zu gewährleisten.

Die Aufnahmen der longitudinalen Spontansprachuntersuchung erfolgten 1–4-mal pro Monat und variierten in ihrer Länge hauptsächlich zwischen 30–45 Minuten (siehe Anhang 3 Tabellen 45, 46 und 47 für monatliche Übersichten der Korpora). Das gesamte vorhandene Audiomaterial wurde transkribiert. Für umfassende Hintergrundinformationen zu den Kindern Jan und Katharina siehe Korecky-Kröll (2011: 30 f.), zum Kind Lena siehe Lettner (2008: 125–129).

Alle drei Kinder wurden in ihrem gewohnten Umfeld mit einem Kassettenrekorder aufgenommen und waren noch sehr klein, als mit den Aufnahmen begonnen wurde (Jan 1;3, Katharina 1;6, Lena 1;7, vgl. Tabelle 2). Daher waren sie zwar an dem Aufnahmegerät interessiert, verhielten sich in ihrer Kommunikation mit ihrer Mutter jedoch absolut natürlich. Jan und Lena wurden allein in Anwesenheit ihrer Mutter, eventuell ihrer Geschwister und anderer vertrauter Personen (Vater, Großeltern) aufgenommen, während bei den Spontansprachaufnahmen von Katharina eine befreundete Experimentatorin der Freundin der Mutter zugegen war, um die Aufnahmen durchzuführen, die sich jedoch im Hintergrund hielt. Nachdem Katharina, wie bereits erwähnt, zu Beginn der Aufnahmen 1,5 Jahre alt war, war sie der Anwesenheit der Experimentatorin gegenüber völlig unbefangen und interagierte und kommunizierte mit ihrer Mutter auf natürliche Weise. Keine der Mütter erhielt gezielte Anweisungen oder Vorgaben, sondern sollte sich ganz so verhalten wie gewöhnlich.

4.1.2 Paneldatenerhebung Spontansprachaufnahmen

Die einstündigen Audio- und Videoaufnahmen der Kinder des INPUT-Projekts erfolgten mit digitalen Aufnahmegeräten zwischen Juni 2012 und März 2015. Anfangs mit einem Roland R-09HR, später mit zwei Olympus LS-3 Digitalen PCM Rekordern, die zusätzlich mit dem Olympus ME-52W Rauschunterdrückungsmikrofon ausgerüstet waren. Das Audioaufnahmegerät wurde von der Hauptbezugsperson während der Aufnahme bei sich getragen, um eine möglichst einwandfreie und gut verständliche Aufnahme zu erhalten. Sämtliche Audioaufnahmen erfolgten im mp3-Format, alle Videoaufnahmen im mp4-Format. Die

Videoaufnahmen wurden mit einer Sony Videokamera erstellt, wobei die Kamera statisch im Raum platziert wurde.

Die Hauptbezugspersonen erhielten keine speziellen Instruktionen, sondern wurden darum gebeten, zum Zeitpunkt der Aufnahme das Fernsehgerät ausgeschaltet zu lassen und sich so zu verhalten, wie sie dies immer im Umgang mit ihren Kindern tun, um eine möglichst unbeeinflusste Aufnahme sicherzustellen. Die Aufnahme fand im vertrauten Zuhause statt, es stand den Eltern und Kindern frei, die Aufnahmestunde nach Belieben zu gestalten (Unterhalten, Spielen, Vorlesen, Singen etc.), zusätzlich wurde durch das Elterninterview mit der Experimentatorin (vgl. Kap. 4.2) ein gewisses Maß an Vertrautheit geschaffen. Durch diese Rahmenbedingungen wurde eine nahezu natürliche Umgebung für die Aufnahme geschaffen, die einer Alltagssituation annähernd gleichkommt.

Die Freiheit, die den Eltern und Kindern bei der Auswahl der Interaktion eingeräumt wurde, führt zu erheblicher Variation zwischen den verschiedenen Aufnahmen, aber auch zu wechselnden Situationen innerhalb einer Aufnahme. Die gemeinsamen Aktionen umfassten dabei jede Art von Spiel (Brett-, Karten-, Rollen-, Geschicklichkeits-, Gedulds-, Gedächtnis-, Bau-, und Konstruktionsspiele etc.), Buchanschauen oder -vorlesen, Basteln, Zeichnen und Malen, Herumtoben, spontane Unterhaltungen während oder abseits des Spielens, sowie in wenigen Ausnahmefällen, Situationen, in denen Essen zubereitet oder gegessen wurde. Dabei dienten die ersten Minuten der Aufnahme meist dazu, sich in der neuen Situation einzufinden und mit dem Setting (Kamera, Aufnahmegerät, Experimentatorin) vertraut zu machen.

Von der einstündigen Spontansprachaufnahme (Audio- und Videoaufnahmen), wurden die sprachlich interaktivsten 30 Minuten für die Transkription ausgewählt. Häufig kam es vor, dass zwischen einzelnen Situationswechseln, wenn das Kind zur Toilette musste oder die Hauptbezugsperson kurz den Raum verließ, keine oder kaum Konversation stattfand, weshalb sich die Transkripte meist aus zwei Abschnitten unterschiedlicher Länge zusammensetzen, die in Summe eine halbe Stunde Interaktion ergeben. Von jeder Aufnahme zu jedem Erhebungszeitpunkt wurden jeweils exakt 30 Minuten transkribiert, unabhängig davon, ob das Transkript mitten im Satz endete. Der quantitative Datenumfang im Sinne von

transkribierten Minuten ist daher für jedes Kind exakt gleich, wodurch die Daten sehr gut miteinander vergleichbar sind (siehe Anhang A3.4 Tabellen 48 und 49 für detaillierte Übersichten des INPUT-Korpus).

Bei der Errechnung des Durchschnittsalters je SES-Gruppe, ergibt sich in allen Datenpunkten, mit Ausnahme des vierten, eine gleiche Altersverteilung (vgl. Tabelle 3). Zum vierten Erhebungszeitpunkt sind die LSES-Kinder im Durchschnitt um einen Monat jünger als die HSES-Kinder (4;7 vs. 4;8). Die Verfasserin erwartet von dieser äußerst geringen Differenz keine Auswirkungen auf die spontansprachlichen Ergebnisse zu Gunsten der älteren Kinder, vor allem auch deshalb, weil es sich um den letzten Erhebungszeitpunkt kurz vor dem fünften Geburtstag handelt und eine Tendenz bereits nach der Auswertung der ersten drei Zeitpunkte erkennbar ist.

4.2 Elterninterviews

Im INPUT-Projekt wurden mit allen Hauptbezugspersonen an jedem der vier Erhebungszeitpunkte leitfadengestützte Elterninterviews durchgeführt.

4.2.1 Erstellung und Durchführung der Interviews

Die vier Interviewleitfäden wurden von Key Researcher Christine Czinglar (für das Deutsche) und Projektmitarbeiterin Kumru Uzunkaya-Sharma (für das Türkische) entwickelt. Sie orientierten sich dabei hauptsächlich an Brizić (2007),¹² Gathercole und Thomas (2007), Scheele et al. (2010) und einer Vorversion von Unsworth (2013). Dabei hatte jedes Interview einen bestimmten Schwerpunkt. In Interview 1 ging es primär um die Erhebung der relevanten SES-Variablen zur Feststellung des sozioökonomischen Status (siehe dazu Kap. 4.3). In Interview 2, das rund 3 Monate nach dem ersten durchgeführt wurde, lag das Hauptaugenmerk auf dem sprachlichen Input des Kindes (zu Hause und im Kindergarten, weiterer Input durch Fernsehen, Hörbücher, Vorlesen, Konversationen etc.), sowie auf Bildungsaspirationen. Interview 3, das ca. ein Jahr nach dem zweiten stattfand, war ähnlich aufgebaut wie Interview 2, und diente dazu etwaige Veränderungen der Inputfaktoren festzustellen. Das letzte Interview, das zum vierten

¹² Die überdies konstruktives Feedback zu Erstversionen der Interviewleitfäden abgegeben hat, wofür ich mich im Namen aller Beteiligten an dieser Stelle noch einmal bedanken möchte.

Erhebungszeitpunkt, rund 3 Monate nach dem dritten durchgeführt wurde, war ein Kurzinterview, das etwaige Änderungen der Familiensituation bzw. im (sprachlichen) Umfeld des Kindes dokumentierte. Durchgeführt wurden die Interviews mit den deutschsprachigen Familien von Christine Czinglar und Katharina Korecky-Kröll bei den Familien zuhause zwischen einem Test und der einstündigen Spontansprachaufnahme. Die Dauer eines Interviews wurde mit rund 30 Minuten anberaumt, hing jedoch davon ab, wie ausführlich die Fragen von den Hauptbezugspersonen beantwortet wurden. Außerdem wurden die Eltern auf die Freiwilligkeit der Beantwortung der Fragen hingewiesen. Wenn die Hauptbezugsperson eine Frage nicht beantworten wollte, wurde diese übersprungen.

Da die Elterninterviews im Rahmen der Spontanspracherhebung und Testungen stattfanden, wurden diese ebenfalls auf Audio und Video aufgenommen (siehe Kap. 4.1). Während des gesamten Interviews notierte sich die Interviewerin alle Antworten und Kommentare der Hauptbezugsperson auf dem ausgedruckten Fragebogen. Den Eltern wurde dabei die Gelegenheit geboten, ihre Antworten auch zu einem späteren Zeitpunkt noch einmal zu ergänzen oder zu verändern. Neben der handschriftlichen Dokumentation des Interviews wurden die Audio- und Videoaufnahmen zur Erfassung und späteren Auswertung der Antworten herangezogen (siehe Kap. 4.1.2, Kap. 4.2.3).

4.2.2 Interviewdesign

Jedes Interview zielt darauf ab, bestimmte Informationen zu erheben bzw. zu sammeln. Im Fall der Elterninterviews wurde eine Vielzahl an Variablen abgefragt, die bei einer späteren Datenanalyse entscheidend sein können.

Im Wesentlichen sollten all jene Bereiche ausführlich erfasst werden, die es ermöglichen, einen Einblick in die Familiensituation und -geschichte, sowie alle Informationen zum sprachlichen Hintergrund der Familie zu erhalten, um mögliche Einflussfaktoren auf den Spracherwerb des Kindes zu identifizieren.

Im ersten Interview sollte die Hauptbezugsperson des jeweiligen Kindes festgestellt werden, da dies entscheidend für die spätere Transkription der Spontansprachaufnahmen war, auch sollten möglichst viele Daten gesammelt werden, das

Aufwachsen und das Umfeld des Kindes betreffend (erstes Wort, erste Wortverbindungen, Alter zum Zeitpunkt des Sprachbeginns, weitere Sprachen, die in der Familie gesprochen werden, andere wichtige Bezugspersonen oder Geschwister, Eingewöhnung im Kindergarten). Wie im folgenden Kapitel (4.3) genauer dargestellt, wurden sämtliche Daten erhoben, die eine Zuordnung der Familie zu einem sozioökonomischen Status möglich machen (höchste abgeschlossene Schulbildung der Hauptbezugsperson, Berufe beider Eltern, Wohnsituation, Zufriedenheit mit dem Familieneinkommen), ebenso wurden in diesem Zusammenhang auch Bildungsaspirationen erfasst.

Insbesondere wurde in allen Interviews auf die Erhebung von Informationen Wert gelegt, die über die altersgemäße Entwicklung des Kindes Aufschluss geben, sowie überdies einen individuellen Einblick gewähren in die jeweilige Inputmenge in der Erstsprache und eventuell weiteren Sprachen und die dazugehörigen Bezugspersonen. Besonders detailliert wurden in den ersten drei Elterninterviews sämtliche Daten zu einem typischen Tagesablauf der Kinder an einem Kindergartentag sowie am Wochenende erhoben, und zu allen Aktivitäten, die für die Sprachentwicklung relevant sein können, wie etwa Anzahl und sprachliche Interaktion mit Geschwistern, spontane Konversationen, Kommunikation beim Spielen, Zeit, die mit Erzählen, Vorlesen, Singen, Reimen, Dichten, Musizieren, Hörbüchern, Radio, Fernsehen oder sozialen Medien (Spielen am Tablet oder Computer, Nutzung von Smartphone, Internet oder YouTube) verbracht wird. Das vierte Interview erhob ebenso Daten zu der Zeit, die das Kind im Kindergarten verbringt, diente jedoch primär dazu, etwaige Veränderungen (Beruf, zusätzliche Ausbildungen und Wohnsituation) festzustellen.

Die Interviews dienten außerdem dem Aufbau einer Vertrauensbasis zwischen der Interviewerin, die auch die Experimentatorin war, und der Hauptbezugsperson, und auch dazu, die Eltern in ihrer Funktion als kompetenteste Experten für ihr Kind zu bestärken. Im Zuge des Interviews, das möglichst wie ein freies Gespräch geführt wurde, erhielt die Hauptbezugsperson auch die Möglichkeit Problembereiche anzusprechen oder etwaige Unsicherheiten zu thematisieren. Als positiver Nebeneffekt ermöglicht das Interview dem Kind eine Pause zwischen der Spontansprachaufnahme und einer etwaigen Testung.

4.2.3 Interviewdaten – Erfassung und Auswertung

Die Interviewdaten aller vier Erhebungszeitpunkte, lagen in Form handschriftlich ausgefüllter Erhebungsbögen, sowie als Audio- und Videoaufnahme vor. Für die Auswertung der Daten hat die Verfasserin zunächst, auf Basis eines beratenden Gesprächs mit der Statistikerin Isabella Hager, mit dem Apple®-Programm Numbers® ein sehr umfangreiches Tabellenblatt erstellt, das später nach Microsoft® Excel exportiert wurde. In diesem Arbeitsblatt wurde jeder einzelne Fragebogen so detailliert wie möglich erfasst. Jedem Kind wurde dabei eine Fallnummer, sowie ein anonymisiertes Namenskürzel zugewiesen. Nachdem sich die Interviewschwerpunkte abhängig vom Erhebungszeitpunkt etwas voneinander unterscheiden (vgl. Kap. 4.2.1), gibt es in der Auswertung der vier Zeitpunkte sowohl gleiche als auch unterschiedliche Variablen. In Summe gibt es in den vier Elterninterviews 314 Variablen. In diese Arbeit werden, neben soziodemographischen Variablen wie Alter, SES und Geschlecht, auch Verhaltensvariablen (z. B. Mediennutzung), sowie sämtliche Kriterien in die Analyse miteinbezogen, die für den Spracherwerb wesentlich sind bzw. sein können (siehe Anhang A2.3 Tabelle 44).

Sämtliche grundlegende Faktoren wurden bereits zum ersten Zeitpunkt erhoben. Diese beinhalten unter anderem Angaben zur Hauptbezugsperson des Kindes, deren Sprachen (L1, L2 und weitere), höchster Bildungsabschluss und Beruf, und zum Kind selbst, etwa dessen Alter zum Zeitpunkt der Aufnahme, Geschlecht, sozioökonomischer Status, Anzahl der Geschwister und Geburtsreihung des Fokuskindes, Alter bei Eintritt in den Kindergarten, Erstsprache (sowie weitere Sprachen), Sprachbeginn (erstes Wort, erste Wortverbindungen), produktives und rezeptives Sprachverhalten bzw. -niveau, Anzahl der Bezugspersonen je Sprache, Tagesablauf an einem Wochentag und am Wochenende, sowie generell über Zeit, die das Kind in Interaktion, Konversation oder gemeinsamer Aktivität mit der Hauptbezugsperson verbringt. Für einen umfassenden Überblick über die sprachrelevanten Variablen, die in den Elterninterviews erhoben und in dem Datenblatt erfasst wurden, siehe Anhang A2.3 Tabelle 44.

Die zum ersten, zweiten und dritten Erhebungszeitpunkt erhobenen Angaben zum sprachlichen Input der Kinder wurden in dem Datenblatt jeweils pro Tag und pro Woche erfasst und gliedern sich einerseits in den Input, den das Kind zu Hause

(durch die Hauptbezugsperson, den Partner, andere Erwachsene, Geschwister) erhält, andererseits den Input im Kindergarten (durch die primäre und sekundäre Kindergartenpädagogin, andere Kinder). Zum vierten Erhebungszeitpunkt liegen ausschließlich Daten zum Input im KG vor. Ebenfalls erfasst wurde die zum zweiten und dritten Zeitpunkt erhobene zeitliche Dauer, die das Kind mit der Hauptbezugsperson zusammen bei spontanen Interaktionen (Spielen, Erzählen, Singen, Reimen, Vorgelesen bekommen) und Konversationen verbringt, sowie der zeitliche Umfang des Medienkonsums bzw. der Mediennutzung des Kindes pro Woche. Hauptaugenmerk lag dabei auf der Zugänglichkeit und Nutzung von Printmedien (Büchern), die das Kind vorgelesen bekommt, auf dem Konsum von Massenmedien wie Fernsehen und Radio, und überdies auf Neuen Medien wie Hörbüchern, Smartphone, und Internet (YouTube), da diese in der Auswertung der Daten berücksichtigt werden (siehe Kap. 10.1 und 10.2). Zu allen in dieser Arbeit untersuchten Kindern des INPUT-Korpus liegen sämtliche sprachrelevanten Daten zur Inputmenge zu Hause und im Kindergarten über den gesamten Zeitraum vor. Für die statistische Auswertung der Spontansprachdaten und experimentellen Daten werden dabei jene Daten herangezogen, die den L1-Input zu Hause durch die Hauptbezugsperson abbilden.

Der sprachliche Input der Kinder setzt sich aus Sprachmaterial zusammen, das primär aus zwei verschiedenen Kanälen stammt. Zum einen ist dies Spontansprache, die direkt an das Kind gerichtet wird, zum anderen handelt es sich um sprachlichen Input, den das Kind durch den Filter der Medien wahrnimmt. Diese Äußerungen sind im Gegensatz zur Spontansprache an einen universellen Empfänger, und nicht direkt an das Kind, gerichtet. Für die vorliegende Arbeit am wesentlichsten ist der sprachliche Input durch die Hauptbezugsperson des Kindes. Aufgrund der in den letzten Jahren ansteigenden Nutzung der Medien (Fernsehen, Internet, Smartphone, YouTube, Hörbücher etc.), möchte die Verfasserin diese nicht unbeachtet lassen, sondern in die Auswertung der Daten miteinbeziehen, um zu sehen, zu welchem Grad, bzw. in welchem Umfang dieser mediale Input Einfluss auf den Spracherwerb nimmt.

4.3 Operationalisierung des sozioökonomischen Status

Für eine umfassende Analyse der Spontansprache wird in dieser Arbeit der sozioökonomische Status (SES) als externer Einflussfaktor miteinbezogen. Dabei soll gezeigt werden, ob bzw. in welchem Umfang ein Zusammenhang zwischen dem sozioökonomischen Status einer Familie und dem Sprachverhalten hergestellt werden kann. Die Zuordnung des sozioökonomischen Status beruht dabei auf den Angaben der Eltern der Kinder, die im Zuge umfangreicher Elterninterviews (siehe dazu Kap. 4.2) erhoben wurden. Es wurden mehrere Variablen, wie etwa die höchsten Bildungsabschlüsse und die aktuellen Berufe der Eltern, die Wohnsituation und die Zufriedenheit mit dem Familieneinkommen zur Einschätzung des sozioökonomischen Status herangezogen. Nachdem die dezidierte Frage nach der Höhe des Einkommens in Österreich ein überaus sensibles Thema darstellt, wurde in den Interviews nach der Zufriedenheit mit dem Einkommen gefragt.

In Czinglar et al. (2015: 212–215) wird die Operationalisierung des SES, wie sie im INPUT-Projekt und damit auf die Daten in Tabelle 3 Anwendung fand, ausführlich dargestellt. Die wichtigsten Faktoren für die Zuordnung sind im Folgenden zusammengefasst.

Als wesentlichste Variable für die Zuordnung der Familien zu einem sozioökonomischen Status diente der höchste Bildungsabschluss der Hauptbezugsperson. Die Bildungsabschlüsse wurden nach der *International Standard Classification of Education* ‚Internationale Standardklassifikation des Bildungswesens‘ (ISCED-97) der UNESCO klassifiziert (OECD 1999). Die Grenze, ob eine Familie dem niedrigeren oder höheren SES zugeordnet wurde, wurde bei ISCED-97-Level 3a gezogen (vgl. Abb. 1 in Czinglar et al. 2015: 212). Bildungsabschlüsse ab dem Niveau von 3a wurden dabei als höherer SES (HSES¹³), Werte zwischen 1 und 3b als niedrigerer SES (LSES¹⁴) eingestuft. Die zweite

¹³ In diese Kategorie fallen alle Bildungsabschlüsse ab Maturaniveau, d.h. Abschluss einer (berufsbildenden oder allgemeinbildenden) höheren Schule mit Matura, abgeschlossene längere Weiterbildung nach dem Abschluss einer Lehre, Kollegabschluss bzw. Abschluss eines Aufbaulehrgangs, Meisterprüfung, Fachhochschulabschluss oder Universitätsabschluss.

¹⁴ Alle Bildungsabschlüsse unter Maturaniveau, wie etwa Lehrabschluss, Pflichtschulabschluss oder der Abschluss einer berufsbildenden mittleren Schule fallen in diese Kategorie.

wichtige Variable war das Berufsprestige, des aktuell von der Hauptbezugsperson ausgeübten Berufes. Dieses wurde anhand des *International Socioeconomic Index of Occupational Status*, 'Internationaler Sozioökonomischer Index des beruflichen Status' (ISEI) ermittelt (nach Ganzeboom & Treiman 1996). Die so festgestellten Prestigewerte der unterschiedlichen Berufe wurden in jenen Fällen für die Zuordnung zu einem sozioökonomischen Status herangezogen, in denen der höchste Bildungsabschluss der Hauptbezugsperson eklatant unter dem Berufsprestige der ausgeübten Tätigkeit angesiedelt war. Auf diese Weise konnte eine Hauptbezugsperson, die ihrer Ausbildung nach als LSES eingestuft worden wäre, aufgrund der weitaus höheren beruflichen Position in die HSES-Gruppe aufsteigen. Dies traf für die deutschsprachigen Hauptbezugspersonen auf eine Mutter (D27-JOZ) zu, die trotz ihres geringeren Bildungsabschlusses zu HSES hochgestuft werden konnte. Abstufungen nach unten, aufgrund einer der Ausbildung nicht entsprechenden Berufsposition gab es keine.

4.4 Transkription und morphologische Codierung

Die Audio- und Videoaufnahmen wurden mittels einer für das Deutsche adaptierten Version des CHILDES¹⁵ Systems (vgl. MacWhinney 2000), zunächst im CHAT¹⁶-Format transkribiert und in einem weiteren Schritt mit dem MOR-Programm, sowie dem von Laaha (2004) entwickelten, und sich ständig erweiternden, Lexikon semiautomatisch morphologisch codiert (vgl. dazu auch Korecky-Kröll 2011: 26; Laaha & Korecky-Kröll 2009). Transkribiert wurden sämtliche verbalen und nonverbalen Zeichen des Kindes, seiner Hauptbezugsperson und weiterer anwesender Personen. Also alle spontanen Äußerungen, rezitieren von Inhalten durch Singen und Vorlesen, Laute und Geräusche (z. B. Husten, Schnäuzen, Weinen, Seufzen, unverständliches Geplapper, Brabbeln), sowie für den Kontext wesentliche Bewegungen, Gesten und Zusatzinformationen (z. B. Hinzeigen, Zumachen, Aufstehen). Für die Analyse der transliterierten Konversationen im

¹⁵ CHILDES = *Child Language Data Exchange System*, vgl. URL: <https://childes.talkbank.org/> [23.03.2021].

¹⁶ CHAT = *Codes for the Human Analysis of Transcripts*, ist ein Transkriptionsformat, das die automatische Datensuche und -analyse vereinfacht, vgl. URL: <https://talkbank.org/> [23.03.2021].

CHAT-Format wurde in weiterer Folge das Programm CLAN¹⁷ installiert und für einfache Abfragen wie etwa die mittlere Äußerungslänge (*Mean length of utterance*, MLU), Standardabweichung (*Standard deviation*, SD) und Lemma- und Tokenfrequenzen herangezogen. Für die genaue Analyse der synthetischen Komposita, aber hauptsächlich der Derivationen, war es außerdem nötig, die Daten umfassend zu überarbeiten und wesentlich detaillierter zu kategorisieren, als dies nach der semiautomatischen Codierung der Fall ist. Für Beispiele zu Transkription und morphologischer Codierung siehe Anhang 1.

4.4.1 Erweiterte Datencodierung

Sämtliche Transkripte wurden mittels des Javascript-Programms CLANTOCSV (Korecky 2015) vom CHAT-Format in CSV konvertiert und in einem weiteren Schritt in Microsoft[®] Excel importiert.

Dabei muss für jedes einzelne Transkript zuerst eine longtier-Version des morphologisch codierten Files erstellt werden, damit es danach mit dem Programm CLANTOCSV ins CSV-Format konvertiert werden kann. Die so entstandenen CSV-Dateien können dann einzeln nach Excel importiert werden. Nachdem die Konvertierung der Daten in dem File erst ab dem ersten Sprecher, markiert mit einem Asterisk (*), z. B. *KAT, *CHI, *PAR, erfolgt, müssen die Metadaten nach dem Import in den ersten fünf Spalten des Excel-Dokuments ergänzt werden. Dabei handelt es sich um wesentliche Identifikations- und Zuordnungsmerkmale wie das anonymisierte Kürzel des Kindes, dessen sozioökonomischen Status, sein Geschlecht und Alter, sowie den Zeitpunkt der Datenerhebung. Nur durch diese Metadaten ist es danach in der Auswertung möglich, die einzelnen Korpora nach Sprecher, Erhebungszeitpunkt, Geschlecht, Alter oder sozioökonomischen Status getrennt auszuwerten. Für eine genaue Übersicht der nötigen Basiscodierungen siehe Anhang A2.1 Tabelle 42.

¹⁷ CLAN = *Computerized Language Analysis*. Das Programmpaket steht zum kostenlosen Download für diverse Betriebssysteme zur Verfügung unter URL: <https://dali.talkbank.org/clan/> [23.03.2021].

Im Zuge des Imports wird für jedes einzelne Wort einer Äußerung eine eigene Zeile angelegt (siehe Beispiel in Tabelle 4). Dadurch kann jedes Wort separat codiert werden, was bei der Analyse enorme Vorteile mit sich bringt.

Tabelle 4. Beispiel einer Äußerung nach dem Import in Excel

Äußerung D28-MOH (4;5)	Lemma	Wortform	Wortklasse
eine heizung die die luft reinblast von drauss(e)n.	ein	ein-e	DET:art:indef
eine heizung die die luft reinblast von drauss(e)n.	heizung	heizung	N:02:f
eine heizung die die luft reinblast von drauss(e)n.	die	d-ie	PRO:rel
eine heizung die die luft reinblast von drauss(e)n.	die	d-ie	DET:art:def
eine heizung die die luft reinblast von drauss(e)n.	luft	luft	N:04:f
eine heizung die die luft reinblast von drauss(e)n.	rein#blas	rein#blas-3S**a	V:07
eine heizung die die luft reinblast von drauss(e)n.	von	von	PREP
eine heizung die die luft reinblast von drauss(e)n.	draussen	draussen	ADV

a. Dialektale Formen, die ohne Umlaut gebildet werden, wie in diesem Beispiel (*reinblast* statt *reinbläst*), werden zwar mit ** markiert, nicht jedoch als Fehler gewertet.

Wie aus Tabelle 4 ersichtlich, sind die morphologischen Codierungen für eine Auswertung der Wortbildungsmuster nicht ausreichend, da lediglich grundlegende Kategorien für Substantive, Verben und Adjektive wie etwa Wortklasse, Genus und Flexionsklasse angegeben sind. Deshalb wurden in den Excel-Listen zahlreiche Spalten für Zusatzcodierungen hinzugefügt. Die wesentlichsten sind hier kurz erklärt, für eine detaillierte Übersicht der Zusatzcodierungen siehe Anhang A2.2 Tabelle 43.

Zunächst wurden die Listen um die Spalte *Word_category* (Wortkategorie) erweitert, in der jedes einzelne Substantiv, Verb und Adjektiv einer Kategorie zugeordnet wurde: 0 Simplex, 1 Simplex (pseudo-deriviert), 2 Derivat, 3 Kompositum, 4 synthetisches Kompositum. Kategorie 1 wurde eingeführt, um zu sehen, wie viele Lemmata, die die Kinder im Input hören, oder auch selbst produzieren, ein Muster enthalten, das rein formal eine Derivation sein könnte (z. B. Pseudo-Suffigierung in *Haube* und *fertig*, Pseudo-Präfigierung in *gewinnen*). Große Häufigkeiten könnten zu einem schnelleren Erwerb bestimmter Kategorien begünstigend beitragen. Diese Zuordnungen ermöglichen in der Auswertung das Filtern bestimmter Kategorien auf sehr einfache Weise.

Verben wurden in folgende Kategorien unterteilt (*V_type*): Simplex-Verben, Partikelverben (Verben mit trennbarem Präfix), Präfixverben (Verben mit

untrennbarem Präfix), Zirkumfixverben, suffigierte Verben, und Verben, die durch Konversion gebildet werden (morphologisch, sowie durch Stammveränderung). Außerdem wurden die verwendeten Präfixe (trennbar und untrennbar) in einer eigenen Spalte (*Prefixtype*) zusätzlich codiert als Präfixe (untrennbar, z. B. be-, verzer-), Partikel (trennbar, z. B. auf#, rein#, weg#), Doppelpartikel (trennbar, z. B. herum#, hinein#) und Zirkumfixe (untrennbar, z. B. be-...-ig, ent-...-ig).

Zusammengesetzte Adjektive und Substantive, sind zwar in der morphologischen Basiscodierung durch ein + gekennzeichnet, dies reicht in der Regel jedoch für die Analyse der Nominalkomposita nicht aus. Daher wurden diese in einer Spalte (N_Comp) zunächst zugeordnet, ob sie (a) ein Wortbildungsmuster im Erstglied enthalten, (b) synthetische Komposita sind, oder (c) kein zusätzliches Wortbildungsmuster oder Affix enthalten. Überdies wurde in einer zusätzlichen Spalte (X_Comp_type) der genaue Typ der Konstituenten von Nominal- und Adjektivkomposita angegeben (z. B. AA dunkel+blau, VN Hüpf+burg, NN Tennis+schläg-er), vgl. dazu in Anhang A2.2 Tabelle 43.

Die wichtigsten Codierungen für die Auswertung der Daten sind jedoch die Spalten: *Derivation_type_txt*, *Affix_rel_txt*, und *Deriv2_txt*. Diese enthalten den Eintrag der gewerteten Derivation, die exakte Angabe des Affixes bzw. der relevanten Derivation, sowie allfällige weitere Derivationsmuster, wie etwa Umlaut. Dabei wird immer die zuletzt angewendete Derivation als relevant eingetragen, wie anhand der Beispiele (1), (2) und (3) demonstriert:

- (1) Auto+fahr-er: 1. fahren → Fahr-er, 2. Auto + Fahr-er
Derivation_type = synth. KOMP, Affix_rel: KOMP, Deriv2 = SUFF er
- (2) Ge-fäng-nis: 1. fangen → ge-fangen, 2. -nis + UML
Derivation_type = SUFF, Affix_rel: -nis, Deriv2 = Umlaut (ä), PREF ge
- (3) weg#bringen: 1. bringen → weg#bringen
Derivation_type = PTL, Affix_rel: PTL weg

Durch die Aufteilung in verschiedene Spalten ist es in der Auswertung möglich, die Auswahl mehr oder weniger auf bestimmte Muster einzuengen und damit generelle oder spezifische Auswertungen durchzuführen.

In engem Zusammenhang zu dem Ableitungsmuster steht die Wortart der Ableitungsbasis. Diese wird in einer eigenen Spalte (Base_txt, vgl. in Anhang A2.2 Tabelle 43) so genau wie möglich angegeben. Ist die Derivationsbasis ein Verb, so wird zwischen Partikel-, Präfix- und Zirkumfixverben, sowie suffigierten Verben unterschieden.

Zur Auswertung des morphologischen Komplexitätsgrades eines Derivats oder Kompositums wird dieser in der Spalte *Complexity* angegeben. Um die Datenanalyse bis 4;11 in dieser Arbeit mit jener von Mattes (2018), die bis 6;0 (JAN) reicht, abzustimmen, hat sich die Autorin bei der Zuordnung der Komplexitätsgrade an ihrer Arbeit orientiert:

Tabelle 5. Morphologische Komplexitätsgrade der Derivate und Komposita
(modifiziert übernommen aus: Mattes (2018: 123))

Grad	Anz. Derivation Wortbildung	Muster	Beispiele
0	keine	Simplizia und Verbstämme, die formal ein Affix enthalten, jedoch keine entsprechende Basis (pseudo-deriviert)	Käfer, Hund, gewinn-, beginn-, radier-, fertig, gesund
1	eine	Affigierung, Partikel, Um-/Ablaut, Konversion, komplexe Verbstämme, einfache Komposita	probier-, leg-, witzig, besuch-, Stoß, Wiege, Rettung, Apfel+baum
2	zwei	Affigierung, Partikel, Um-/Ablaut, Konversion, synthetische Komposita, Komposita mit Ableitung im Erstglied, komplexe Komposita	Läufer, Absperrung, Beton+mischer, ableg-, Besuch, gefährlich, Eis+lauf+schuh
3	drei	Affigierung, Partikel, Um-/Ablaut, Konversion, synthetische Komposita, Komposita mit Ableitung im Erstglied, komplexe Komposita	Einzäunung, erkältet, Fern+bedienung, Verkäuferin, abbestell-, zweifärbig,
4	vier (oder mehr)	mehrere Affigierungen, Komposition, Um- /Ablaut, Konversion	Undurchschaubarkeit, Befüll+stelle, bewerk-stellig-, Welt+verbesserung

Für eine genaue Beschreibung des Komplexitätsbegriffs, wie er in dieser Arbeit Anwendung findet, sowie der Kategorisierung der Lexeme siehe Kap. 7.

Nach Beendigung der Codierungen weisen die einzelnen Datenfiles eine Größe von 18,2 MB (KAT), 28,2 MB (LEN), 57,4 MB (JAN) und 64,3 MB (INPUT) auf.

4.4.2 Datenauswahl

Nachdem in dieser Arbeit der Fokus auf der Spontansprache des Kindes (CS) und seiner Hauptbezugsperson (CDS) liegt, fließen in die Auswertung ausschließlich diese Daten ein. Also sämtliche spontan produzierten, oder neu kreierten Äußerungen. Alle anderen Sprecher,¹⁸ die auf den Aufnahmen zu hören sind, wurden zwar ebenfalls transkribiert und morphologisch codiert, werden jedoch nicht untersucht. Außerdem werden all jene Äußerungen ausgeschlossen, die sich an andere im Raum befindliche Erwachsene (ADS) oder an anwesende Haustiere richten (*pet-directed speech*, PDS), da es sich um Äußerungen handelt, die nicht direkt an das Kind gerichtet sind, sondern vom Kind nebenbei mit angehört werden.

Alle rezitierten Reime, Lieder (Singen), Buchtexte (Vorlesen) und Zitate werden ebenso nicht analysiert, da es sich hier um wiedergegebene Schriftsprache oder auswendig gelernte Inhalte und nicht um Spontansprache handelt. Überdies werden alle direkten Imitationen des Kindes aus der Analyse ausgeschlossen. Als direkte Imitationen gelten all jene spontanen Imitationen, die unmittelbar auf die Äußerung der Hauptbezugsperson folgen und diese exakt wiedergeben.

Zitate, Imitationen, ADS und PDS werden bereits beim Transkribieren am Ende einer Äußerung speziell markiert: [+ cit], [+ imi], [+ ads], [+ pds] und erhalten dadurch keine MOR-Zeile, sondern eine CIT-, IMI-, ADS- oder PDS-Zeile (siehe im Anhang Codierungsbeispiele für CIT in A1.3 und für IMI in A1.4), was es im Zuge der Analyse einfach macht, diese ein- oder auszuschließen. Mit dieser Vorgehensweise ist sichergestellt, dass allein die Spontansprache im Input und Output gemessen werden kann.

In den einzelnen Wortklassen der Substantive, Verben und Adjektive werden marginale und extragrammatische Wortbildungen wie etwa Abkürzungen, Verschmelzungen und Rückbildungen explizit von der Analyse ausgeschlossen, sowie folgende Kategorien: Onomastische Bildungen, also Eigennamen, wie etwa Bezeichnungen oder Namen von Personen, Lebewesen oder Objekten (z. B. Daniel, Bernhardiner, Scrabble, Österreich, Donau), da sich die Morphologie der

¹⁸ Es kommt in den Aufnahmen immer wieder vor, dass Geschwister, Großeltern, Besucher oder der andere Elternteil, der nicht die Hauptbezugsperson ist, anwesend und zu hören sind.

Eigennamen von jener gewöhnlicher Substantive häufig unterscheidet, sowie von Eigennamen abgeleitete Adjektive (z. B. Wien-er). Ebenso werden Diminutive im Rahmen der nominalen Ableitungen gesondert behandelt, und nur der Vollständigkeit halber in den Tabellen angeführt, da diese hauptsächlich auf pragmatischer Ebene eine evaluierende Funktion einnehmen. Für eine umfassende Analyse von Diminutiven im Österreichischen Deutsch siehe Korecky-Kröll und Dressler (2007), Lettner (2008), sowie Dressler und Korecky-Kröll (2015).

Die in dieser Arbeit untersuchten Verben umfassen sämtliche Verben, mit Ausnahme von Auxiliärverben, da diese nur in Verbindung mit einem Vollverb das Prädikat eines Satzes bilden können, und Modalverben, sofern diese nicht als Vollverb vorkommen (z. B. Er mag Hunde vs. Er mag den Hund streicheln). Aufgrund der eindeutigen morphologischen Codierungen der Eigennamen (als n:prop), Hilfs- und Modalverben (V:aux, V:mod) können diese ausgeschlossen werden, während Kopulaverben in die Analyse miteinfließen, da eine manuelle Durchsicht zu zeitaufwändig und daher nicht zu bewerkstelligen gewesen wäre.

4.4.3 Datenbereinigung

Vor allem in Lenas und Kathis Daten ist die Verfasserin immer wieder auf fehlerhafte morphologische Codierungen oder Rechtschreibfehler (Tippfehler) im Lexikon gestoßen. Die häufigsten und für die Analyse relevantesten Fehler sind in Tabelle 6 anhand von Beispielen in fehlerhafter und korrekter Codierung aufgelistet.

Tabelle 6. Beispiele für fehlerhafte morphologische Codierungen in den Korpora

Korpus	Fehler	fehlerhafte Codierung	korrekte Codierung
LEN	Flexionsendung	ab#deckt-3S	ab#deck-3S
LEN	Flexionsendung	rein#sage-1S	rein#sag-1S
JAN	Rechtschreibfehler	Lieblnigs+fahrer	Lieblings+fahrer
JAN	Rechtschreibfehler	zusammen#lleg-PP	zusammen#leg-PP
LEN	Rechtschreibfehler	herun#spring-2S	herum#spring-2S
KAT	Rechtschreibfehler	zureuck	zurueck ^a
KAT	Rechtschreibfehler	raem-INF	raeum-INF
KAT	Rechtschreibfehler	vesteck-PP	versteck-PP
LEN	Partikelmarkierung	zer#geh-3S (Partikel)	zergeh-3S (Präfix)
D28-MOH	fehlende Partikel	V:07 blas-3S**	rein#V:07 blas-3S**
KAT	Präfix	uebrigbleib- (Präfix)	uebrig#bleib- (Partikel)

LEN	Stammfehler	ab#sitz-INF ^b	ab#setz-INF
KAT	Stammfehler/Partikel	herunter#faell-3S ^c	runter#fall-3S
LEN	Flexionsklasse	V:01 back-INF	V:08 back-INF
LEN	Sprecher	*LEN: ^d	*MUT:

a. Die Beispiele dieser Tabelle bilden die Einträge im Lexikon ab. Das Lexikon, auf dem die morphologische Codierung basiert, wurde vor vielen Jahren von Sabine Laaha (2004) entwickelt, als das Programm CLAN noch keine Umlaute verarbeiten konnte, weshalb sämtliche Daten ohne Umlaute und scharfes ß transliteriert wurden und werden. Das Lexikon umfasst aktuell ca. 55.000 Einträge, daher wäre eine Änderung auf Umlaute nur manuell möglich und zu zeitintensiv bzw. fehleranfällig.

b. Die von Lena im Alter von 2;5 produzierte Äußerung: *LEN: ich kann nicht haeue [: haube] absitz(e)n [*], führte im Zuge der morphologischen Codierung dazu, dass das existente Verb **absitzen** (eine Strafe, oder von einem Pferd) ins Lexikon aufgenommen wurde, statt es korrekterweise als **absetzen** zu codieren (mit fehlender Umlautung).

c. In diesem Lemma, geäußert von Kathis Mutter im Altersmonat 2;4 finden sich gleich zwei Codierungsfehler: *MUT: ja umdrehen sons(t) faellts [: faellt es] (he)runter. Das Verb wurde morphologisch als V:01|herunter#faell-3S codiert. Tatsächlich hat die Mutter jedoch nur *runter* gesagt. *Runter* und *herunter* sind allerdings insofern als zwei unterschiedliche Partikel anzusehen, als *runter* die umgangssprachliche Form darstellt, während die Doppelpartikel *herunter* dem Standard entspricht und sich auch bezüglich der Sprecherperspektive unterscheidet (vgl. Rainer 2010: 84). Ebenso fehlerhaft ist der Stamm, der nicht *fäll-* ist, sondern eindeutig *fall-*, da die Umlautung erst im Zuge der Flexion geschieht (vgl. fall-en ‚fall-INF‘ → fäll-t ‚fall-3SG‘) und ist daher von *fällen* (einen Baum) jedenfalls zu unterscheiden.

d. In Lenas Aufnahme im Alter von 2;3 kommt die Äußerung: *LEN: Lena bitte jetzt lass sie mal in ruhe kakao trink(e)n. Es handelt sich jedoch eindeutig um die Äußerung der Mutter, was von der Autorin in den Daten korrigiert wurde.

Die in Tabelle 6 aufgelisteten Fehler führen dazu, dass Lemmas doppelt gezählt werden, wie dies etwa bei Rechtschreibfehlern oder der Angabe der Flexionsendungen der Fall wäre. Ebenso würde V:01/08|back- zweimal als Lemma vorkommen, aufgrund der unterschiedlichen Flexionsklassen. Die inkorrekte Zuordnung des Sprechers wiederum hätte zur Folge, dass eine Äußerung wie in unserem Beispiel (siehe Tabellenfußnote d) dem Kind statt der Mutter zugeordnet würde und dadurch mitunter komplexe Wortbildungsmuster früher in der Sprachproduktion auftreten, als dies eigentlich der Fall ist. Das Rautezeichen als Partikelmarkierung dient bei der morphologischen Codierung der Unterscheidung zu Präfixverben. Dadurch können die sehr frequenten Partikelverben auf einfache Weise herausgefiltert und analysiert werden. Wird ein Lemma versehentlich mit einer Raute versehen, wie in *zer#geh-, obwohl es sich um ein untrennbares Präfix handelt, würde es zu den Partikelverben statt den Präfixverben gezählt, oder vice versa zu den Präfix- statt den Partikelverben, wie in *übrigbleib-, und somit die tatsächliche Anzahl der jeweiligen Wortklasse verfälschen. Eine gänzlich fehlende

Partikel wie in dem Beispiel des Kindes D28-MOH würde das Lemma als Simplex ausweisen, obwohl es sich um ein Partikelverb handelt. Dieser Fehler liegt darin begründet, dass die Partikel nach der morphologischen Codierung und Disambiguierung im File zu suchen sind und vom Transkribierenden vor die Wortklasse zu verschieben sind (vgl. Beispiel (8)). Es kann vorkommen, dass eine Partikel übersehen wird, oder ein Software- oder Speicherfehler vorliegt. Im Allgemeinen bietet die doppelte Kontrolle, wie sie bei allen Files Anwendung fand, eine relativ sichere Methode zur Fehlervermeidung.

Neben den fehlerhaften Codierungen wurden in Lenas Korpus Zitate (CIT) und Imitationen (IMI) erst nach der morphologischen Codierung am Ende einer Äußerung ergänzt. Das ursprüngliche Fehlen der Markierungen [+ imi] und [+ cit] führte daher dazu, dass bei der halb automatischen Codierung ausschließlich MOR-Zeilen, statt MOR-, IMI- und CIT-Zeilen, angelegt wurden. Die korrekten Zeilenbeschriftungen für IMI und CIT mussten deshalb manuell in den einzelnen Files nachgetragen werden. Unglücklicherweise wurde das Fehlen der Markierungen von Imitationen und Zitaten erst bei der qualitativen Analyse bemerkt, was darin resultierte, dass Lenas Daten und die ihrer Mutter nochmals genau überprüft und überarbeitet werden mussten. Dieser Schritt war aufgrund der Datenauswahl (siehe Kap. 4.4.2) äußerst wichtig, da es sonst unmöglich gewesen wäre, spontane Imitationen und Zitate in diesem Korpus auszuschließen.

Die genaue Durchsicht und Bearbeitung der Daten ist insgesamt betrachtet ein sehr zeitintensiver, wenngleich notwendiger, Prozess. Für die Vollständigkeit und Richtigkeit der Ergebnisse der Datenauswertungen sind sämtliche Korrekturen allerdings absolut notwendig und wesentlich.

4.4.4 Zählung

In jedem Korpus wurde jedes Verb, Substantiv und Adjektiv in CS und CDS nach dem Import und der Datenbereinigung und sofern sie den in Kap. 4.4.2 genannten Kriterien entsprechen, entweder als Simplex, Kompositum oder Derivat kategorisiert und gegebenenfalls in verschiedene Subkategorien eingeteilt (vgl. in Anhang A2.2 Tabelle 43 *Derivation_type_txt*). Bei komplexen Formen, die mehr als ein Wortbildungsmuster enthalten, wurde das zuletzt angebrachte Muster gezählt (vgl. Beispiele (1), (2) und (3)).

Die Komposita in den Wortklassen ADJ und N bilden alle zusammengesetzten Adjektive und Substantive ab. Als Komposita gelten alle Einträge, die auf der morphologischen Codierungszeile ein + enthalten (z. B. Kaffee+bohne, hell+rosa). Dabei wird bei den Substantivkomposita zwischen folgenden vier Formen unterschieden: (a) KOMP – Komposita, die keine Ableitung enthalten, z. B. Haus+tier, (b) KOMP1 – Komposita mit einer Ableitung im Erstglied, z. B. Flug+hafen, (c) synth. KOMP – synthetische Komposita, z. B. Staub+saug-er, (d) KOMPd – Diminutivkomposita, z. B. Blume-n+körb-chen. Nachdem im Rahmen dieser Arbeit Diminutive nicht im Speziellen ausgewertet werden, werden sämtliche Komposita, die einen Diminutiv (&DIM) enthalten, zu Komposita der Kategorie KOMP gezählt. Auch fließt die Subkategorie KOMP1 in KOMP ein, wenn nicht anders angegeben.

Aufgrund ihrer Sonderstellung zwischen Derivation und Komposition wird die Kategorie der synthetischen Komposita, also Substantivkomposita, die eine Ableitung im Kopf enthalten, einzeln gezählt, um deren Verwendung und Entwicklungsverlauf verfolgen zu können.

Bei den Derivationen wird hauptsächlich zwischen den Kategorien Suffigierung, Präfigierung, Zirkumfigierung und Konversion unterschieden. Obwohl auch Diminutive mittels Suffixen gebildet werden, werden sie in dieser Arbeit gesondert gezählt und von den anderen suffigierten Derivationen separiert betrachtet (vgl. Kap. 4.4.2 und in Anhang A2.2 Tabelle 43). Die Kategorie DIM enthält alle Bildungen, die nicht in Komposita vorkommen (z. B. Vog-erl, Schiff-lein).

Die einzelnen Wortbildungsmuster wurden für CS und CDS separat ausgewertet, und ihre Verteilung in den vier Korpora berechnet und untersucht. Dabei wurden die Häufigkeiten in Lemmas und Tokens gemessen, und nicht in Types, da die Zählung flektierter Formen für die Wortbildung nicht maßgeblich ist. Viel wesentlicher ist die Anzahl der unterschiedlichen Lexeme, unabhängig von ihrer Flexion.

Die Unterscheidung, ob ein Lexem tatsächlich mittels eines Affixes abgeleitet ist oder ob es sich um ein pseudo-deriviertes Lexem handelt, das allein aufgrund seiner Struktur abgeleitet sein könnte, jedoch über keine entsprechende Basis verfügt

(z. B. **wenig**, **gesund**, **beginnen**), wurde basierend auf verschiedenen, für die deutsche Sprache wesentlichen, Quellen getroffen. Diese umfassen wichtige elektronische Wörterbücher und Nachschlagewerke, wie [duden.de](https://www.duden.de)¹⁹, [canoonet.eu](http://www.canoonet.eu)²⁰, sowie das etymologische Wörterbuch von Kluge (2011). Ebenso wurde die Entscheidung, ob es sich bei einem Lexem um ein existierendes Wort des Deutschen handelt oder um eine Neubildung, nach dessen Vorkommen im Wörterbuch entschieden, wenn es der Autorin ebenfalls unbekannt war.

Im Gegensatz zu der Datenauswertung in Sommer-Lolei et al. (2021) werden alle abgeleiteten Formen gezählt. Dies schließt Bildungen ein, die sehr opak, oder lexikalisiert sind (z. B. Schaff-ner, Schling-el), sowie auch Bildungen, die sich mit nichtlexikalisierten Minimaleinheiten²¹ verbinden (vgl. Definitionen zu Konfix in Fleischer & Barz 2012: 63 f.; Elsen 2014: 294), die im Unterschied zu Ableitungsbasen nicht für sich alleine wortfähig sind (z. B. unikale Morpheme als Basis für Bildungen auf -ik und -isch, wie in Kom-ik - kom-isch, Log-ik - log-isch).

4.5 Mini-Paradigma-Kriterium

Das Mini-Paradigma-Kriterium (MPK) wurde zunächst von Dressler et al. (2003) entwickelt und auf den Erwerb der verbalen Flexionsmorphologie von Bittner et al. (2003) und Kilani-Schoch und Dressler (2005) weiter angewandt. Später wurde das Kriterium auf Komposita in Bezug auf die Emergenz von Mini-Paradigmen in Dressler et al. (2010), und zur Feststellung potenzieller Produktivität im frühen Spracherwerb von Komposita in Korecky-Kröll et al. (2017) adaptiert. Sommer-Lolei et al. (2021: 116 f.) passten das Mini-Paradigma-Kriterium an die Derivationsmorphologie an und untersuchten die potenzielle Produktivität der häufigsten Derivationsmuster bei Jan, Kathi und Lena bis zu einem Alter von 3 Jahren.

Demzufolge gilt das Mini-Paradigma-Kriterium als erfüllt, wenn ein morphologischer Prozess (Affigierung, Konversion, Stammänderung) auf

¹⁹ URL: <https://www.duden.de> [04.05.2021].

²⁰ URL: <http://www.canoonet.eu> [03.01.2019]. Das zu diesem Zeitpunkt existente Online-Tool zur Wortbildung steht seit der Übernahme am 01. Juni 2019 durch die LEO GmbH URL: https://dict.leo.org/pages/about/ende/canoonet_de.html [04.05.2021] nicht länger zur Verfügung.

²¹ Vgl. zur Wortbildung: Glück (2010, s. v. Formativ).

zumindest drei verschiedene Basen angewendet wird und wenn eine Basis ihrerseits in mindestens drei verschiedenen Wortformen auftritt, wie flektierten oder unflektierten Wörtern, Komposita oder anderen Derivationen (vgl. Sommer-Lolei et al. (2021: 116).

Zur Feststellung eines Mini-Paradigmas werden die folgenden Schritte durchgeführt (vgl. für Derivationen a.a.O.: Beispiel (13)), die exemplarisch am Adjektivsuffix -ig demonstriert werden, siehe Beispiel (4):

(4) MPK-Feststellung bei Affixen:

- a. Ein Affix wird mit mindestens drei verschiedenen Basen kombiniert:

Jan: (1;9) durst-ig, eck-ig, (1;8) kitsch-ig, (1;10) schmutz-ig
= partielle Erfüllung des MPK für -ig im Alter von 1;9

- b. Eine Basis kommt in mindestens drei verschiedenen Typen vor:

(1;8) rutsch-en, (1;9) Rutsch-e, (1;10) runter#ge-rutsch-t
= partielle Erfüllung des MPK für -ig im Alter von 1;10

- c. Die Basis wird mit dem Affix abgeleitet:

(1;11) rutsch-ig
= vollständige Erfüllung des MPK für -ig im Alter von 1;11

Damit das Mini-Paradigma-Kriterium als erfüllt gilt, müssen alle drei Punkte erfüllt sein.

Es ist wichtig, an dieser Stelle zu erwähnen, dass zur Feststellung der potenziellen Produktivität eines Wortbildungsmusters nicht zwischen verschiedenen semantischen Funktionen eines Affixes unterschieden, sondern diese auf Basis formaler Kriterien identifiziert wird. Überdies wird in Bezug auf die Ableitungsbasen nicht zwischen Basen, die verschiedenen Wortklassen angehören, unterschieden. Beispielsweise wird nicht zwischen dem Suffix -ung als Nomen Actionis (Rett-ung), Nomen Loci (Wohn-ung) oder Nomen Instrumenti (Heiz-ung) differenziert und ebenso wenig, ob die Ableitungsbasis ein Verb oder Substantiv ist, z. B. -er in Nomina Agentis: Tat → Tät-er vs. fahr-en → Fahr-er.

5 Wortbildung in CS und CDS

5.1 Verteilung der Derivationen je Wortklasse

In Tabelle 7 ist die Gesamtanzahl der Substantive, Verben und Adjektive, sowie die Zahl derivierter Substantive (ohne Diminutive), Verben und Adjektive in CS und CDS für alle Korpora gegeben. Ebenso wird der prozentuale Anteil der Derivationen an der Gesamtheit innerhalb der jeweiligen Wortklasse berechnet.

Tabelle 7. Frequenzen (Lemmas, Tokens und Prozente) derivierter Substantive (ohne DIM), Verben und Adjektive in Relation zur Gesamtzahl an Substantiven, Verben und Adjektiven in CS und CDS in allen Korpora im gesamten Zeitraum

Korpus	JAN	KAT	LEN	INPUT HSES	INPUT LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N (gesamt)	1.048/8.134	317/1.133	302/1.803	1.309/5.926	801/4.422
N DERIV	155/1.108	37/68	36/183	183/829	106/600
% N DERIV/N	14,79/13,62	11,67/6,00	11,92/10,15	13,98/13,99	13,23/13,57
V (gesamt)	583/8.333	174/848	221/2.003	786/8.503	582/6.650
V DERIV	388/1.715	99/217	132/433	556/1.737	409/1.342
% V DERIV/V	66,55/20,58	56,90/25,59	59,73/21,62	70,74/20,43	70,27/20,18
ADJ (gesamt)	171/1.416	33/106	47/189	193/1.519	141/1.024
ADJ DERIV	60/240	5/6	4/7	67/189	44/114
% ADJ DERIV/ADJ	35,09/16,95	15,15/5,66	8,51/3,70	34,72/12,44	31,21/11,13
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N (gesamt)	2.382/18.696	729/2.763	775/4.086	2.034/10.240	1.144/6.453
N DERIV	370/2.326	102/292	122/528	297/1.363	172/901
% N DERIV/N	15,53/12,44	13,99/10,57	15,74/12,92	14,60/13,31	15,03/13,96
V (gesamt)	1.423/32.821	468/5.249	627/7.176	1.228/19.185	957/13.974
V DERIV	1.098/6.546	334/1.270	449/1.975	937/4.319	731/3.540
% V DERIV/V	77,16/19,94	71,37/24,23	71,61/27,52	76,30/22,51	76,38/25,33
ADJ (gesamt)	417/5.808	121/738	142/924	325/2.830	220/1.810
ADJ DERIV	216/1.080	38/108	60/123	154/483	89/246
% ADJ DERIV/ADJ	51,80/18,60	31,40/14,63	42,25/13,31	47,38/17,07	40,45/13,59

Ein direkter Vergleich der prozentualen Häufigkeiten in CS und CDS in Lemmas zeigt, dass die Kinder über den gesamten Zeitraum berechnet in allen drei Wortklassen weniger abgeleitete Lexeme produzieren als ihre Hauptbezugspersonen (vgl. Tabelle 7). Die Korpora von Kathi und Lena umfassen sämtliche Daten bis zu einem Alter von 3 Jahren, während alle anderen Daten bis knapp vor

den 5. Geburtstag reichen. Daraus ergeben sich bei den jüngeren Kindern im Vergleich teilweise höhere Differenzen zwischen CS und CDS.

Insgesamt zeigen sich die größten Unterschiede bei derivierten Adjektiven. Hier produzieren Jan, Kathi und Lena zwischen 16,25 % und 33,74 % weniger ADJ-Lemmas als ihre Mütter, die Kinder des INPUT-Korpus zwischen 9,24 % (LSES) und 12,66 % (HSES). Auch die Häufigkeit der Verwendung ist zwischen 1,65 % (JAN) und 9,61 % (LEN) geringer als bei den Hauptbezugspersonen.

Die Kinder produzieren demnach nicht nur weniger unterschiedliche Adjektivlemmas, sondern verwenden diese auch seltener.

Die geringsten Differenzen lassen sich bei derivierten Substantiven finden. In dieser Wortklasse liegt der Unterschied zu CDS zwischen 0,62 % (INPUT HSES) und 3,82 % (LEN) in Lemmas. Bei den Tokenfrequenzen gibt es eine prozentual geringere Verwendung bei den INPUT LSES-Kindern mit 0,39 %, bei LEN mit 2,77 % und bei KAT mit 4,57 %. Die INPUT HSES-Kinder und Jan zeigen hingegen eine um rund 1 % höhere Verwendungsrate als ihre Hauptbezugspersonen.

Derivierte Verben nehmen eine Position dazwischen ein. Wie auch bei den derivierten Adjektiven und Substantiven zu beobachten, produzieren die Kinder zwischen 5,56 % (INPUT HSES) und 14,47 % (KAT) weniger unterschiedliche abgeleitete Verben (Lemmas) als ihre Hauptbezugsperson. In Tokens zeigen sich geringere Frequenzen bei den INPUT-Kindern HSES (2,08 %), LSES (5,15 %) und bei Lena (5,90 %). Anders jedoch bei Jan (0,64 %) und Kathi (1,36 %), die höhere Frequenzen im Vergleich zu CDS zeigen.

5.2 Verteilung der Komposita je Wortklasse

In Tabelle 8 ist die Gesamtanzahl der Substantive und Adjektive, sowie die Zahl der Nominal- und Adjektivkomposita in CS und CDS für jedes Korpus gegeben. Bei den Substantiven wird zwischen Nominalkomposita (N KOMP)²², sowie

²² Die Kategorie der Nominalkomposita umfasst sämtliche Komposita mit Ausnahme synthetischer Komposita. Darin enthalten sind demnach Komposita ohne Ableitung, Komposita mit einer Ableitung im Erstglied (z. B. Tauch-er+brille) und Diminutivkomposita (z. B. Geld+börs-erl).

synthetischen Komposita (synth. KOMP) unterschieden und der prozentuale Anteil der Komposita an der Gesamtheit innerhalb der jeweiligen Wortklasse berechnet.

Tabelle 8. Frequenzen (Lemmas, Tokens und Prozente) nominaler Komposita (N KOMP), synthetischer Komposita (synth. KOMP) und Adjektivkomposita (ADJ KOMP) in Relation zur Gesamtzahl an Substantiven und Adjektiven in CS und CDS in allen Korpora

Korpus	JAN	KAT	LEN	INPUT HSES	INPUT LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N (gesamt)	1.048/8.134	317/1.133	302/1.803	1.309/5.926	801/4.422
N KOMP	295/1.654	66/119	48/105	407/811	183/379
% N KOMP/N	28,15/20,33	20,82/10,50	15,89/5,82	31,09/13,69	22,85/8,57
synth. KOMP	80/305	11/19	1/1	72/139	36/53
% synth. KOMP/N	7,63/3,75	3,47/1,68	0,33/0,06	5,50/2,35	4,49/1,20
ADJ (gesamt)	171/1.416	33/106	47/189	193/1.519	141/1.024
ADJ KOMP	14/16	1/1	0/0	13/22	5/8
% ADJ KOMP/ADJ	8,19/1,13	3,03/0,94	0,00/0,00	6,74/1,45	4,00/0,78
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N (gesamt)	2.382/18.696	729/2.763	775/4.086	2.034/10.240	1.144/6.453
N KOMP	832/3.528	196/372	206/592	686/1.565	295/686
% N KOMP/N	34,93/18,87	26,89/13,46	26,58/14,49	33,73/15,28	25,79/10,63
synth. KOMP	241/677	30/56	38/58	148/266	75/125
% synth. KOMP/N	10,12/3,62	4,12/2,03	4,90/1,42	7,28/2,60	6,56/1,94
ADJ (gesamt)	417/5.808	121/738	142/924	325/2.830	220/1.810
ADJ KOMP	25/41	5/6	1/1	24/36	11/20
% ADJ KOMP/ADJ	6,00/0,71	4,13/0,81	0,70/0,11	7,38/1,27	5,00/1,10

Die Kinder produzieren im direkten Vergleich der prozentualen Häufigkeiten zwischen CS und CDS im gesamten Zeitraum nicht nur weniger abgeleitete Lexeme als ihre Hauptbezugspersonen (vgl. Kap. 5.1), sondern ebenfalls weniger Nominal- und Adjektivkomposita (vgl. Tabelle 8).

Die größten Häufigkeiten in Lemmas und Tokens weist in CS und CDS die Gruppe der Nominalkomposita (N KOMP) auf. Daher ergeben sich auch hier die größten Differenzen in der Verwendung zwischen CS und CDS. Hier produzieren Jan, Kathi und Lena zwischen 6,07 % und 10,69 % weniger N KOMP-Lemmas als ihre Mütter, während die INPUT-Kinder eine Differenz von rund 3 % aufweisen. Auch die Tokenfrequenzen sind zwischen 1,59 % (INPUT HSES) und 8,67 % (LEN) geringer als bei den Hauptbezugspersonen, mit Ausnahme von Jan, der eine um 1,46 % höhere Verwendungshäufigkeit aufweist als seine Mutter.

Der Unterschied zwischen CS und CDS in der Produktion synthetischer Komposita (synth. KOMP) ist in Lemmas zwischen 0,65 % (KAT) und 4,57 % (LEN) niedriger als in CDS. Geringer fällt die Differenz mit 0,25 % (INPUT HSES) bis 1,36 % (LEN) bei den Tokenfrequenzen aus. Jan zeigt eine minimal höhere Verwendungsrate als seine Mutter.

Adjektivkomposita werden, in dem untersuchten Zeitraum, sowohl von den Hauptbezugspersonen als auch von den Kindern in äußerst geringem Ausmaß produziert. Mit Ausnahme von Lena, die keine Adjektivkomposita bis zu einem Alter von 3;0 produziert, und Jan, der eine um 2,19 % höhere Frequenz in Lemmas aufweist, zeigt sich bei allen anderen Korpora eine um rund 1 % niedrigere Verwendungsrate. In Tokens zeigt sich eine geringere Verwendungshäufigkeit nur bei den INPUT-LSES-Kindern (0,32 %), während sich bei Jan (0,42 %), Kathi (0,13 %), und den INPUT-HSES-Kindern (0,18 %) geringfügig höhere Frequenzen im Vergleich zu CDS zeigen.

Generell weist Jan, wenn auch nur minimal, im Vergleich zu seiner Mutter höhere Tokenfrequenzen in allen Kategorien der Komposita, sowie bei derivierten Verben und Substantiven auf, was aufzeigt, dass er die wenigen Bildungen häufiger verwendet.

Ein Vergleich der prozentualen Häufigkeiten der INPUT-HSES- und -LSES-Kinder zeigt, dass die HSES-Kinder sowohl mehr Komposita als auch Derivationen in allen Wortklassen in Lemmas und Tokens produzieren (vgl. Tabellen 7 und 8). In den CDS-Daten weisen die HSES-Hauptbezugspersonen im Vergleich zu den LSES-Hauptbezugspersonen höhere Lemma- und Tokenfrequenzen, in allen Kategorien der Komposita, sowie bei derivierten Adjektiven auf. Ein umgekehrtes Bild zeigt sich bei abgeleiteten Substantiven und Verben, wo die LSES-Hauptbezugspersonen prozentual betrachtet, mehr unterschiedliche Lemmas produzieren, und diese auch häufiger verwenden.

5.3 Wortbildung der Substantive in CS und CDS

Die Kalkulation der Durchschnittswerte der Lemmafrequenzen in CDS der beiden Korpora JAN (3;0–4;11) und INPUT (gesamter Korpus), ergibt einen prozentualen Anteil aller Substantive, die auf Wortbildung basieren, von 57,70 %. Dieser Wert ist auf folgende Kategorien verteilt: 28,94 % Nominalkomposita, 16,45 % suffigierte Substantive (inkl. Diminutivsuffixe), 7,64 % synthetische Komposita, 3,56 % Konversionen (inkl. implizite Ableitungen mit Stammmodifikation), 0,80 % präfigierte Substantive (inkl. 0,22 % zirkumfigierte Substantive), und 0,31 % nominalisierte Infinitive. Dies ergibt eine Steigerungsdifferenz im Vergleich zu der Auswertung der Daten bis 3;0 um 14,04 % die gesamte nominale Wortbildung betreffend (vgl. 50,60 % in Sommer-Lolei et al. 2021: 127). Daraus resultiert, dass 36,58 % aller Substantivlemmas in CDS durch Komposition gebildet werden (inkl. synthetische Komposita), und 21,12 % durch Derivation und Konversion (inkl. Diminutive, oder 15,98 % ohne Diminutive). Die Verwendung unterschiedlicher derivierter Substantivlemmas nimmt im Vergleich zu den untersuchten Daten bis zu einem Alter von 3;0 um 36,27 % zu (vgl. a.a.O. 15,50 %). Wird die Differenz ohne die Berücksichtigung der Diminutivsuffixe errechnet, so erhöht sich der Wert auf 52,16 %. Eine Berechnung der Vergleichswerte in Komposita (inkl. synth. KOMP) zeigt eine höhere Variation um 4,22 %. Gesamt betrachtet nehmen vor allem die Kategorien der suffigierten Substantive an Variation zu, während alle anderen Kategorien in ihrer Verwendung nur eine geringfügige Steigerung erfahren, oder annähernd gleichbleiben.

5.3.1 Verteilung der Wortbildungsmuster bei Substantiven

Die prozentuale Verteilung der einzelnen nominalen Wortbildungsmuster im Vergleich zu Simplex-Substantiven in CS und CDS ist für Jan, Kathi und Lena in Abbildung 1, für die INPUT-Kinder in Abbildung 2 genau dargestellt. Für eine detaillierte Übersicht der Verteilung in Lemmas, Tokens und Prozenten in allen Korpora im gesamten Zeitraum, siehe Anhang A4.1 Tabelle 50. Erwerbsverläufe derivierter Substantive sind in Kap. 5.3.2 dargestellt.



Abbildung 1. Prozentuale Verteilung der nominalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS bei JAN, KAT und LEN (gesamter Zeitraum)

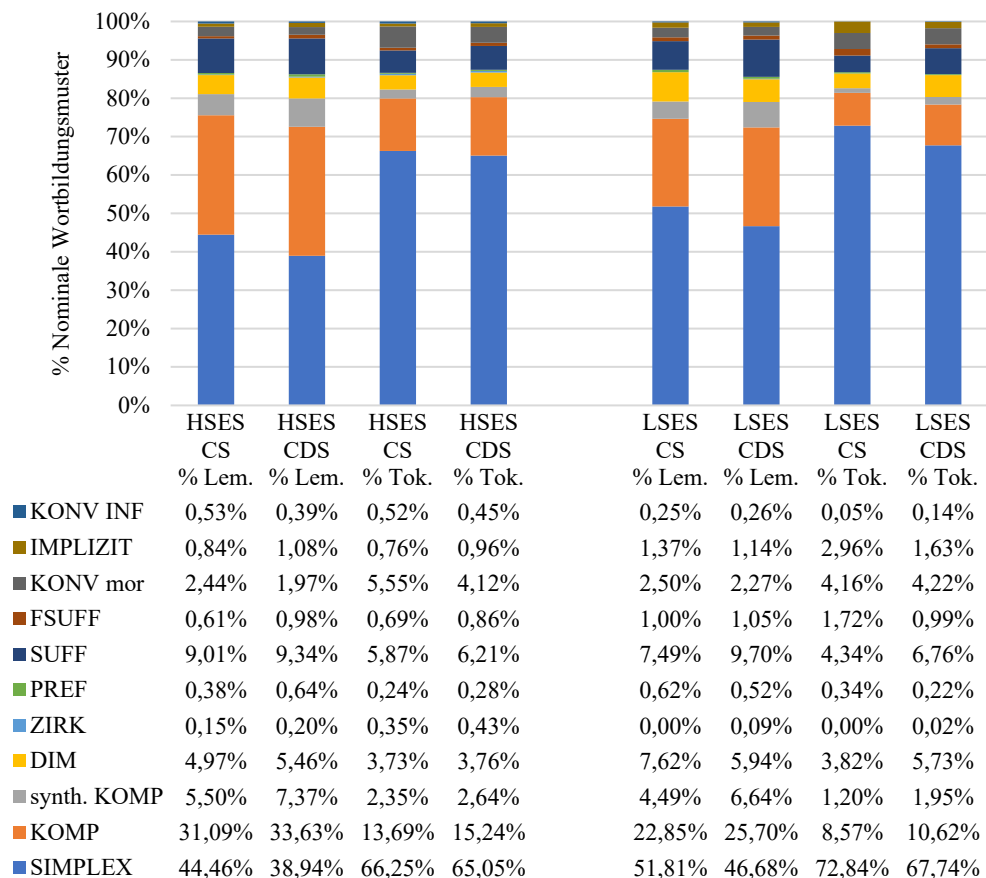


Abbildung 2. Prozentuale Verteilung der nominalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS im gesamten INPUT-Korpus (HSES und LSES)

Die Verteilung der Wortbildungsmuster, wie in Abbildungen 1 und 2 für die Korpora dargestellt, zeigt, welche Wortbildungsmuster für Kinder bis 3;0 bzw. bis 4;11 zur Bildung von Substantiven hauptsächlich relevant sind. Bei allen Kindern und ihren Hauptbezugspersonen sind Komposita von Anfang an in großer Variation (Lemmas) vorhanden. Diese bilden zusammen mit Substantiven, die mittels eines Suffixes abgeleitet werden, die wesentlichsten Kategorien. Dazu tragen auch synthetische Komposita trotz ihres hohen Grades an Komplexität bei. Diminutive zeigen sich bis 3;0 bei Kathi und Lena sehr produktiv und bleiben auch bis zur Vollendung des fünften Lebensjahres ein wesentlicher Bestandteil in CS und in CDS.

Zirkumfixe und Präfixe werden zur Ableitung von Substantiven bis 4;11 nur marginal verwendet. Überdies lässt sich feststellen, dass morphologische Konversionen häufiger gebildet werden als syntaktische Konversionen

(nominalisierte Infinitive) und auch als implizite Derivationen und Ableitungen mittels Fremdsuffixen (vgl. Abbildungen 1 und 2).

Die einzelnen Wortbildungsmuster zur Bildung von Substantiven sind für CS und CDS in absteigender Reihenfolge für das JAN-Korpus in Abbildung 3, für das INPUT-Korpus HSES in Abbildung 4 und für LSES in Abbildung 5 grafisch dargestellt. Dabei wurde ein Fokus auf die wesentlichsten Ableitungssuffixe und Derivationsmuster gelegt und auf die Darstellung nominaler Komposita und Diminutive verzichtet. Die einzelnen Wortbildungsmuster wurden an der Grundgesamtheit der Substantive im gesamten Zeitraum berechnet.

Im Vergleich der nominalen Bildungsmuster zeigt sich sowohl in CDS als auch in CS eine Präferenz für synthetische Komposita in Lemmas. Die Kinder und ihre Hauptbezugspersonen bilden zahlreiche unterschiedliche Lexeme, wiederholen diese jedoch im Fall der INPUT-Kinder und ihrer Hauptbezugspersonen weniger häufig als morphologische Konversionen, wo sich das umgekehrte Bild zeigt. Es werden demnach mehr unterschiedliche synthetische Komposita als morphologische Konversionen gebildet, zweite Kategorie jedoch wesentlich häufiger verwendet (Tokenfrequenz).

In Bezug auf die häufigsten Suffixe ist die Dominanz der Suffixe -er und -e in CDS und CS feststellbar, wenngleich auch diese Bildungen nur für durchschnittlich 5 % aller Substantivlemmas in CDS verantwortlich sind. Das Suffix -ung liegt mit knapp 2 % dahinter zurück. Andere heimische Suffixe, die in den Daten kaum vorkommen, sind -er-ei²³, -er-in, -heit, -keit, -nis, -schaft und spielen daher in dem untersuchten Zeitraum bis 4;11 eine untergeordnete Rolle in CS und CDS. Auch Fremdsuffixe, wie -ade, -age, -arium, -atur, -erie, -eur, -euse, -ie, -ik, -ist, -itor, -ment, -or, -ur) sind in CS und CDS bis 4;11 irrelevant. Einzig in Jans CS und CDS werden die wenigen fremdsuffigierten Substantivlemmas häufiger wiederholt (vgl. Abbildung 3).

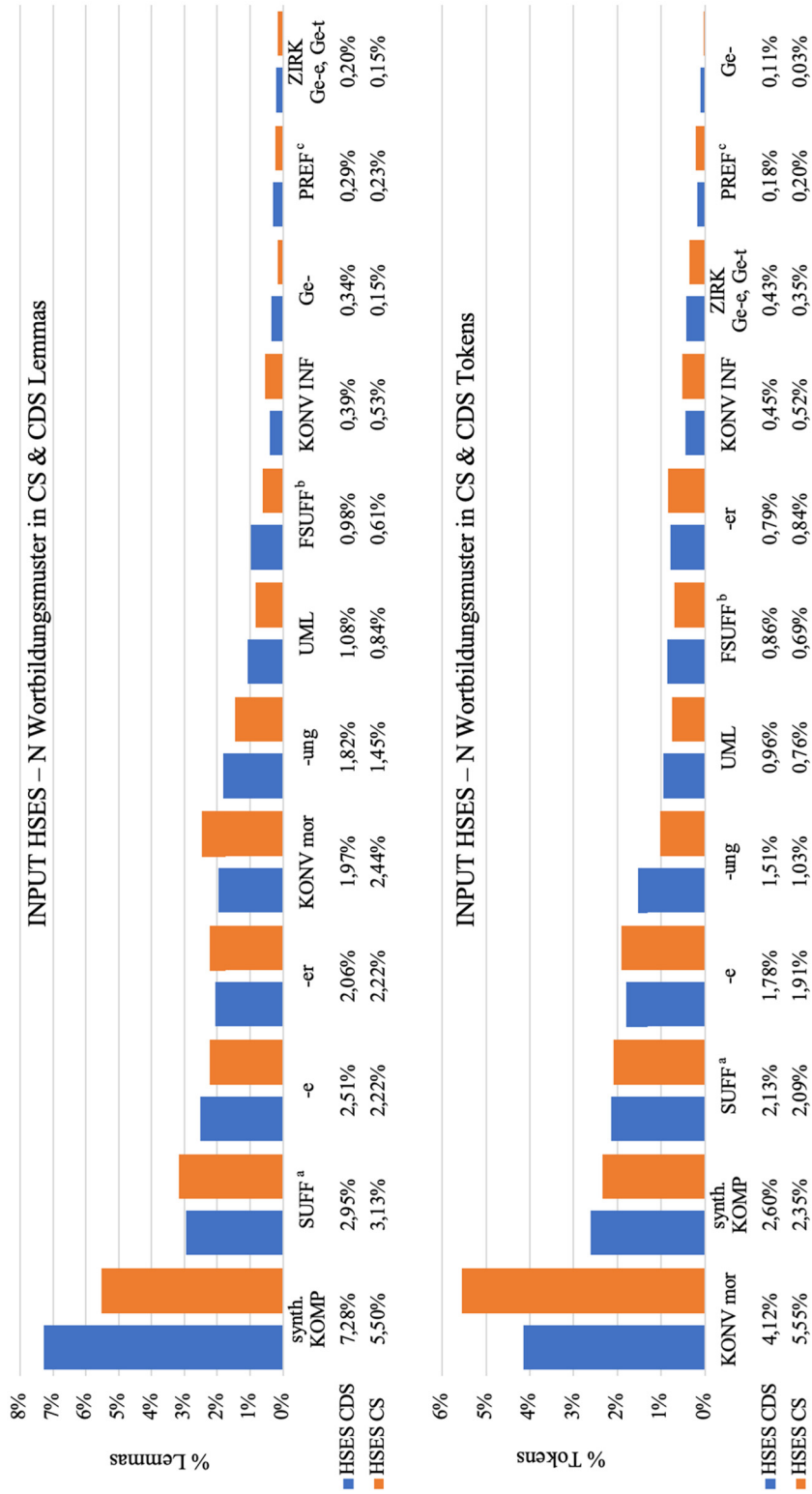
Zirkumfigierungen (Ge-e, Ge-t) zur Bildung von Substantiven, wie in Ge-red-e, Ge-flech-t, treten in CS und CDS in Lemmas und Tokens in geringem Ausmaß auf, ebenso wie Präfigierungen (Ge-, Ab-, Un-, Ur-) und nominalisierte Infinitive (KONV INF).

²³ Historisch gesehen handelt es sich bei dem Suffix -er-ei um ein Fremdsuffix.



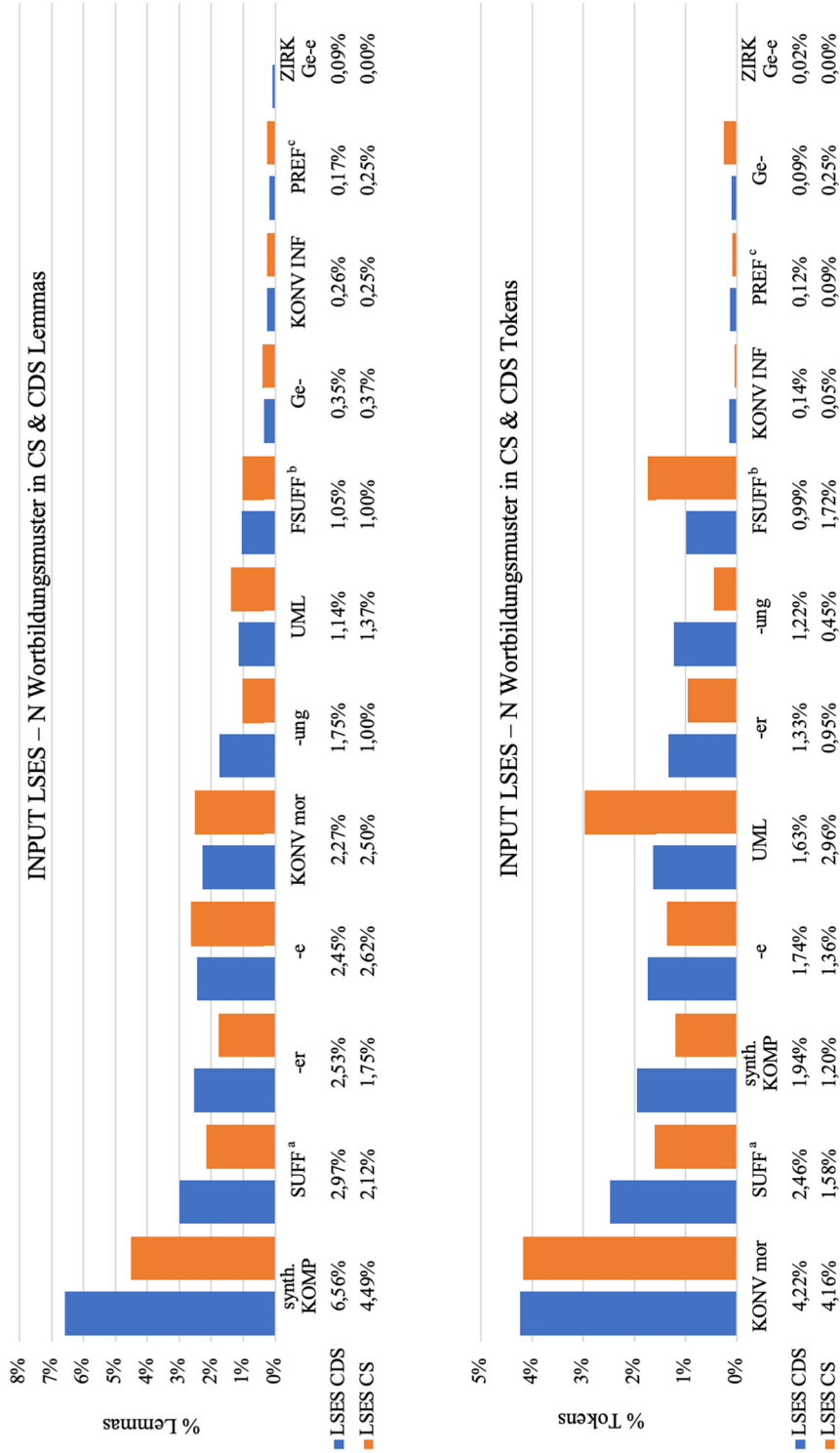
a. andere Suffixe in CS: -ei, -el, -er-ei, -er-in, -heit, -ine, -ling, -nis, -sel; in CDS: zusätzlich -de, -erlich, -iger, -igkeit, -iker, -keit, -schaft, -t, -tel, -tion
 b. Fremdsuffixe in CS: -age, -atur, -erie, -ik, -or; in CDS: zusätzlich -arium, -ie, -itor, -ur
 c. andere Präfixe in CS: Ab-, Un-

Abbildung 3. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum (1;3-4;11) nach absteigender Häufigkeit in CDS



a. andere Suffixe in CS: -ation, -ei, -el, -er-in, -heit, -igkeit, -in, -keit, -ler, -ling, -ner, -nis, -s, -schaft, -sel; in CDS: zusätzlich -er-ei, -ine, -ion, -ose, -t
 b. Fremdsuffixe in CS: -age, -erie, -ik, -ist, -or, -ur; in CDS: zusätzlich -ade, -arium, -atur, -eur, -ment
 c. andere Präfixe in CS: Un-, Ur-; in CDS: zusätzlich Ab-

Abbildung 4. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS



a. andere Suffixe in CS: -el, -ei, -er-ei, -heit, -in, -ling; in CDS: zusätzlich -er-in, -keit, -ler, -nis, -sel, -t

b. Frenedsuffixe in CS: -age, -erie, -ik, -ist, -or; in CDS: zusätzlich -arium, -euse

c. andere Präfixe in CS & CDS: Ab-, Un-

Abbildung 5. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) nominaler Wortbildungsmuster in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS

5.3.2 Erwerbsverlauf derivierter Substantive und synthetischer Komposita

Der Erwerbsverlauf derivierter Substantive im Vergleich zu Simplicia und synthetischen Komposita wird nachfolgend dargestellt. Dabei werden die Lemma- und Tokenfrequenzen des JAN- und INPUT-Korpus pro Erhebungszeitpunkt festgestellt, um den Entwicklungsverlauf der Kinder zu zeigen. Ebenso wird der Verlauf den jeweiligen Inputfrequenzen gegenübergestellt. Der monatliche Verlauf Jans ist in Abbildung 6 (Lem.) und in Abbildung 8 (Tok.) illustriert, Jans CDS-Verlauf in Abbildung 7 (Lem.) und in Abbildung 9 (Tok.). Die Verläufe nach Erhebungszeitpunkt (DP1–4) der INPUT-Kinder und ihrer Hauptbezugspersonen sind für CS in Abbildungen 10 (Lem.) und 12 (Tok.), für CDS in Abbildungen 11 (Lem.) und 13 (Tok.) ausgeführt.

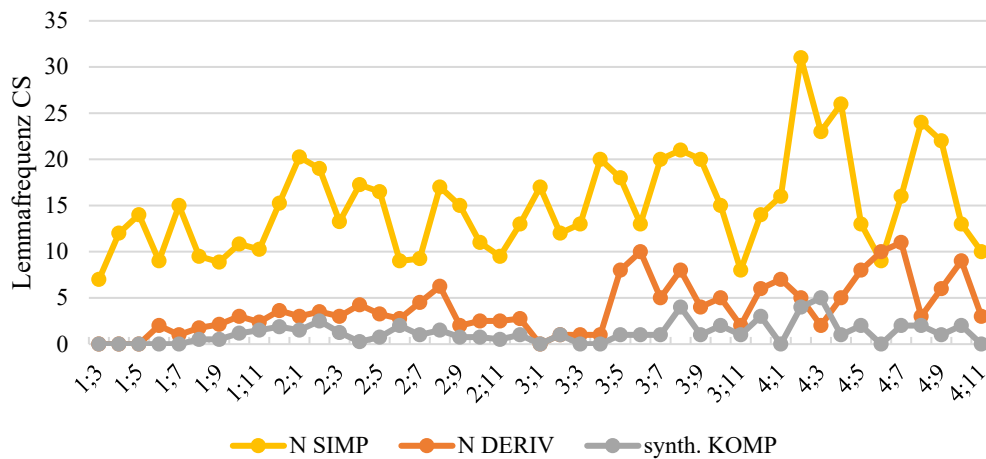


Abbildung 6. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita in Lemmas (1;3–4;11)

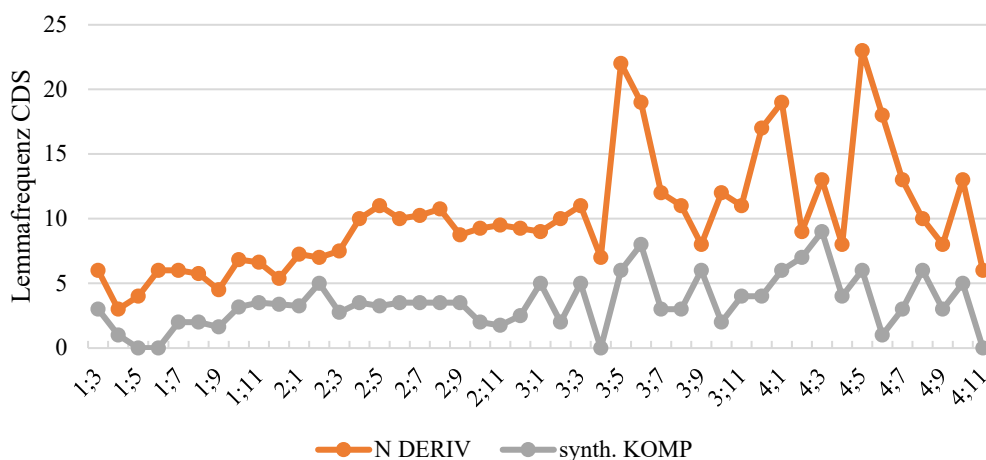


Abbildung 7. JAN-Korpus CDS: Monatlicher Verlauf derivierter Substantive und synthetischer Komposita in Lemmas (1;3–4;11)

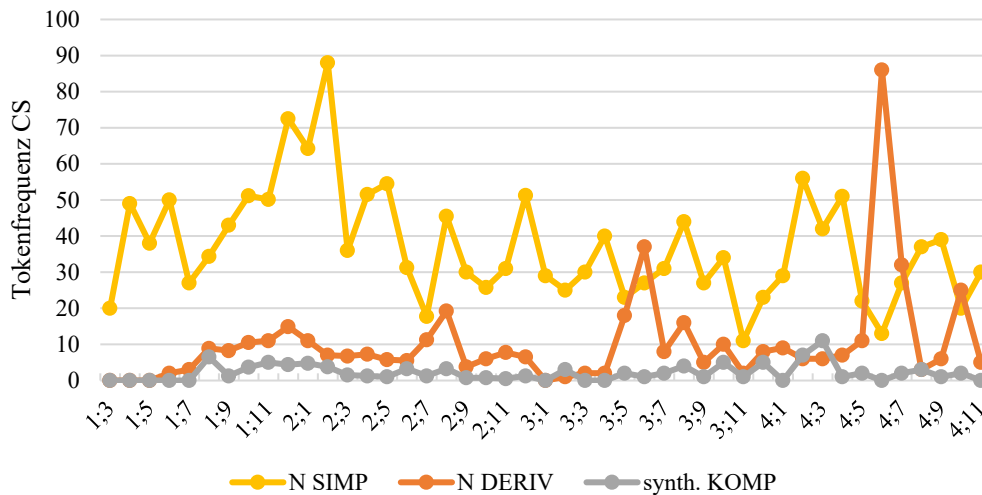


Abbildung 8. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita in Tokens (1;3–4;11)

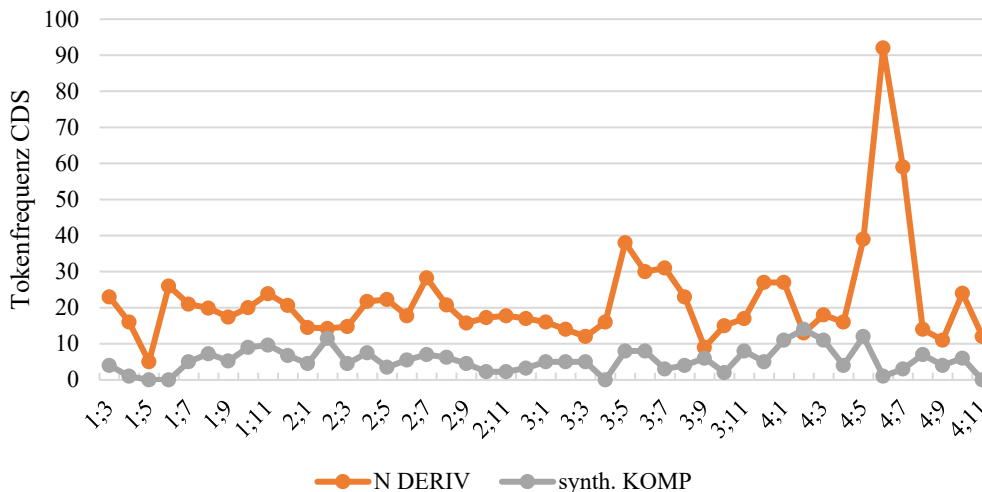


Abbildung 9. JAN-Korpus CDS: Monatlicher Verlauf derivierter Substantive und synthetischer Komposita in Tokens (1;3–4;11)

Simplex-Substantive (N SIMP) treten früher in CS auf als Simplex-Verben (vgl. bei Jan 1;3 vs. 1;4, in den Abbildungen 6 und 19) und sind im gesamten Untersuchungszeitraum in großer Variation (Lemmaanzahl) in der Kindersprache vorhanden (vgl. für Jans CS: Abbildung 6, für die INPUT-CS: Abbildung 10). Bis zu einem Alter von 4;11 dominieren einfache Substantive bei allen Kindern nicht nur in Lemmas, sondern auch in Tokens (vgl. Abbildungen 8 und 12). Die niedrigen Frequenzen synthetischer Komposita sind der höheren Komplexität geschuldet, da diese Komposition und Derivation in sich vereinen.

Derivierte Substantive (N DERIV) nehmen ab ihrem ersten Auftreten stetig an Lemma- und Tokenfrequenz zu, es zeigt sich jedoch ein stärkerer Anstieg von abgeleiteten Substantiven ab einem Alter von 3;5, von synthetischen Komposita (synth. KOMP) zwischen 3;8 und 4;3 bei Jan (vgl. für Lemmas Abbildung 6, für Tokens Abbildung 8), der ebenfalls in seiner CDS sichtbar ist (vgl. für Lemmas Abbildung 7, für Tokens Abbildung 9). Diese Tendenz ist auch in CS und CDS des INPUT-Korpus erkennbar, deren Lemmafrequenzen (N DERIV, synth. KOMP) vom ersten (DP1) zum vierten Zeitpunkt (DP4) zunehmen, mit etwas Variation dazwischen (vgl. Abbildungen 10 und 11). Einzige Ausnahme bilden die LSES-Kinder, deren unterschiedliche Lemmas in beiden Bildungsmustern annähernd gleich bleiben. Die Tokenfrequenzen nehmen nur bei den HSES-Kindern (vgl. Abbildung 12) im Verlauf zu, während diese bei den LSES-Kindern und den Hauptbezugspersonen beider SES-Gruppen annähernd gleichbleiben (vgl. Abbildungen 12 und 13).

Die Differenzanalyse der Lemmafrequenzen abgeleiteter Substantive bei den HSES- vs. LSES-Kindern ergibt eine Zunahme der Differenz im Erwerbsverlauf zwischen den beiden SES-Gruppen, die sich ebenfalls bei derivierten Verben zeigt (vgl. Kap.5.4.2). Die geringsten Differenzen weisen die LSES-Kinder im Vergleich zu den HSES-Kindern zu Datenzeitpunkt 3 auf (in einem Alter von 4;4). Hier unterscheiden sie sich vor allem in der geringeren Bildung -er-suffigierter Substantive. Die größten Differenzen treten bereits durchschnittlich 4 Monate später zu Datenzeitpunkt 4 auf (vgl. Abbildung 10). Die LSES-Kinder bilden dabei hauptsächlich weniger Konversionen als die HSES-Kinder, aber auch weniger abgeleitete Substantive auf -er und -ung. Die Daten belegen eine Vergrößerung der Differenz bei nominalen Derivationen zwischen den Kindern in dem Untersuchungszeitraum (3;1–4;8) und zeigen dadurch, dass die LSES-Kinder nicht bis 4;11 aufholen, sondern im Vergleich mit der Zeit zurückfallen.

Generell lässt sich feststellen, dass CS und CDS bei derivierten Substantiven und synthetischen Komposita ziemlich parallel verlaufen. Bei genauerer Betrachtung der Höchstwerte in Jans CDS zeigt sich eine direkte Auswirkung auf CS. Höhere Lemma- und Tokenfrequenzen in CDS führen ein bis drei Monate später zu höheren Lemma- und Tokenfrequenzen in CS (vgl. Lemmafrequenzen in Abbildungen 6 und 7: N DERIV CDS 2;5/3;5/4;5 vs. CS 2;8/3;6/4;7, synth. KOMP CDS

3;6/3;9/4;1 vs. CS 3;8/4;2/4;3, Tokenfrequenzen in Abbildungen 8 und 9: N DERIV CDS 2;7/3;5/4;5 vs. CS 2;8/3;6/4;6, synth. KOMP CDS 4;2 vs. CS 4;3). Der Ausreißer in derivierten Substantivtokens im Altersmonat 4;6 in CS und CDS beruht auf der großen Häufigkeit von substantivierten Farbadjektiven, wie z. B. Blau und Grün, die Mutter und Kind zu diesem Zeitpunkt oft wiederholen.

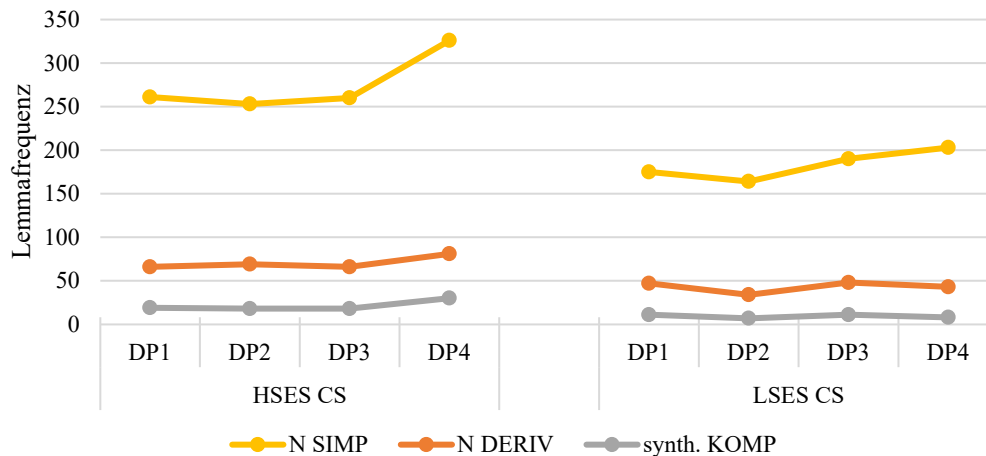


Abbildung 10. INPUT-Korpus CS: Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita nach Erhebungszeitpunkt (DP1–4) in Lemmas

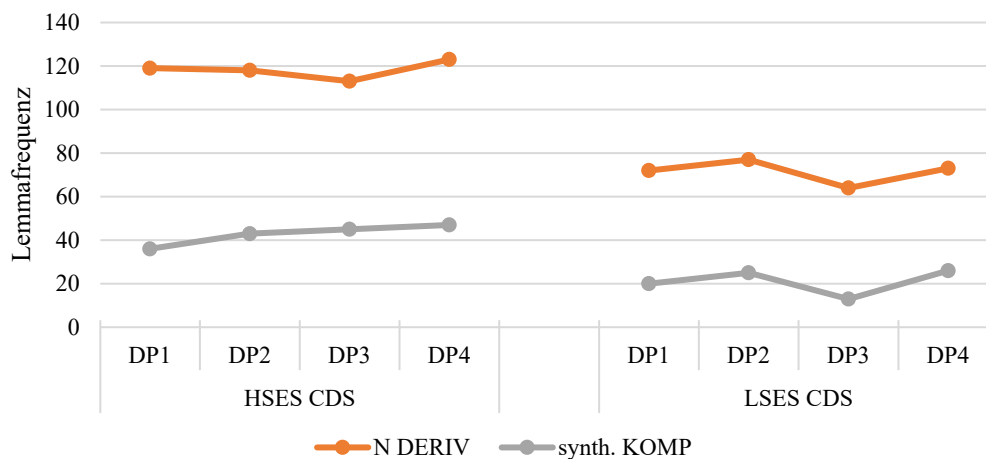


Abbildung 11. INPUT-Korpus CDS: Verlauf derivierter Substantive und synthetischer Komposita nach Erhebungszeitpunkt (DP1–4) in Lemmas

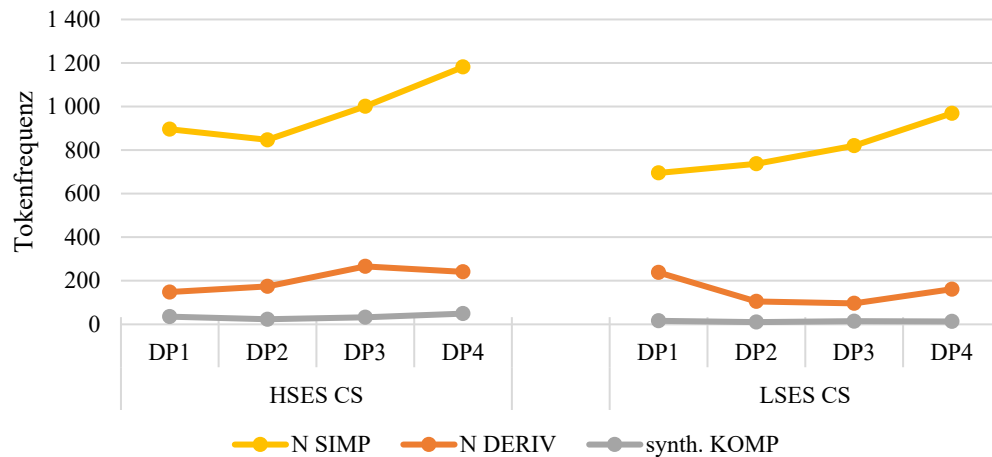


Abbildung 12. INPUT-Korpus CS: Erwerbsverlauf derivierter Substantive vs. Simplex-Substantive vs. synthetische Komposita nach Erhebungszeitpunkt (DP1–4) in Tokens

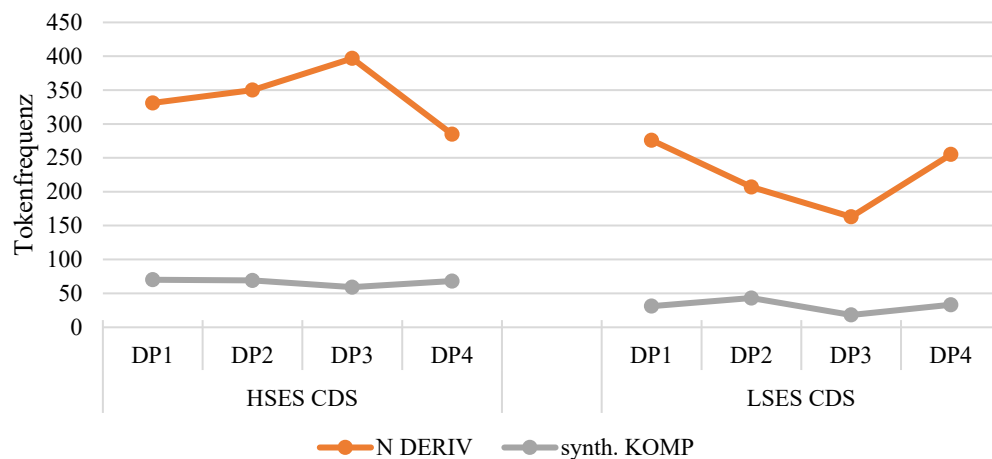


Abbildung 13. INPUT-Korpus CDS: Verlauf derivierter Substantive und synthetischer Komposita nach Erhebungszeitpunkt (DP1–4) in Tokens

Die statistischen Auswertungen der Unterschiede zwischen CS und CDS (JAN-Korpus) bei derivierten Substantiven im Vergleich zu synthetischen Komposita und Simplicia finden sich in Kap. 5.7.1.1. Die mittleren Differenzen derivierter Substantive und synthetischer Komposita, die sich zwischen den Kindern (CS vs. CS, S. 134) sowie den Hauptbezugspersonen (CDS vs. CDS, S. 136) ergeben, werden in Kap. 5.8.1 nach SES-Zugehörigkeit statistisch ausgewertet und auf Zusammenhänge untersucht. Für Gruppenunterschiede zwischen CS und CDS des INPUT-Korpus siehe Kap. 5.8.2.1, für die mittleren Unterschiede zwischen einem früheren und späteren Datenzeitpunkt (DP1+2 vs. DP3+4) siehe Kap. 5.8.2.2.

5.4 Wortbildung der Verben in CS und CDS

Die Berechnung der durchschnittlichen Anzahl derivierter Verben in Lemmas in CDS, ab 3;0 bis zu einem Alter von 4;11 (JAN- und INPUT-Korpus), beträgt 74,50 %. Diese Zahl beruht auf der hohen Lemmafrequenz von Partikelverben, die sich mit 59,38 % von allen anderen Kategorien deutlich abheben. Präfigierte Verben bilden mit 8,01 % (inkl. 0,16 % zirkumfigierte Verben) bereits die zweitgrößte Kategorie. Die Häufigkeiten von Konversionen belaufen sich auf 4,58 % (inkl. 1,86 % implizite Verben mit Stammmodifikation) und suffigierte Verben auf 2,53 %. Die abgeleiteten Verben nehmen um 4,35 % zu, verglichen mit dem Zeitraum bis 3;0 (vgl. 71,40 % in Sommer-Lolei et al. 2021: 128). Diese Steigerung betrifft Präfixverben, suffigierte Verben und Partikelverben.

5.4.1 Verteilung der Wortbildungsmuster bei Verben

Die prozentuale Verteilung der verbalen Wortbildungsmuster im Vergleich zu Simplex-Verben in CS und CDS ist für die longitudinal erhobenen Spontansprachdaten der Kinder Jan, Kathi und Lena in Abbildung 14, für die Kinder des INPUT-Korpus in Abbildung 15 detailliert dargestellt.

Die Frequenzen in Lemmas und Tokens, sowie der jeweilige prozentuale Anteil an der Gesamtheit der Verben in CS und CDS, das heißt die Verteilung der abgeleiteten Verben und Simplex-Verben in allen Korpora über den gesamten Zeitraum, sind in Anhang A4.2 Tabelle 51 umfassend dargestellt.

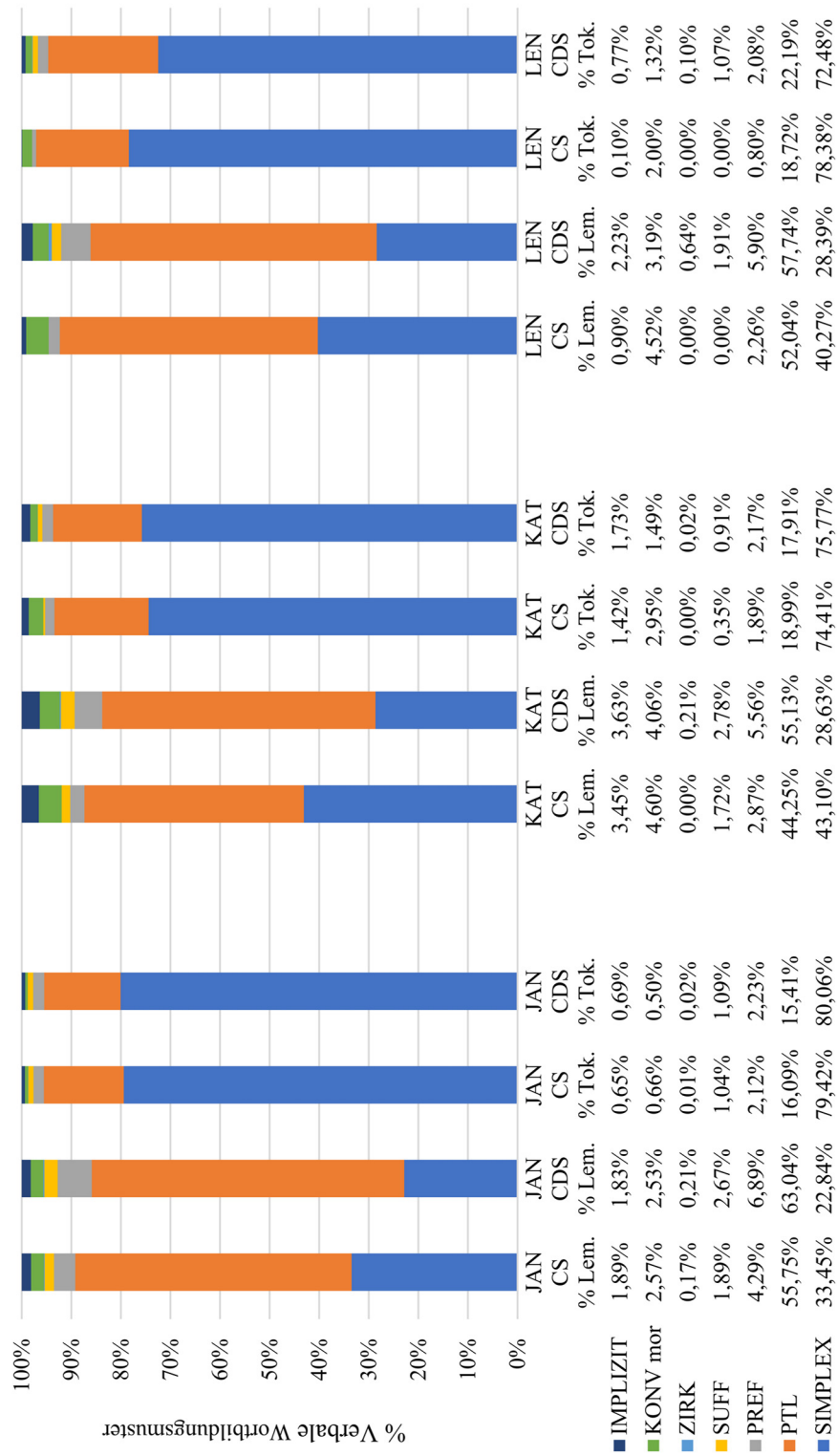


Abbildung 14. Prozentuale Verteilung verbaler Wortbildungsmuster (Lemmas und Tokens) in CS und CDS in den Spontansprachkorpora von JAN, KAT und LEN (gesamter Zeitraum)

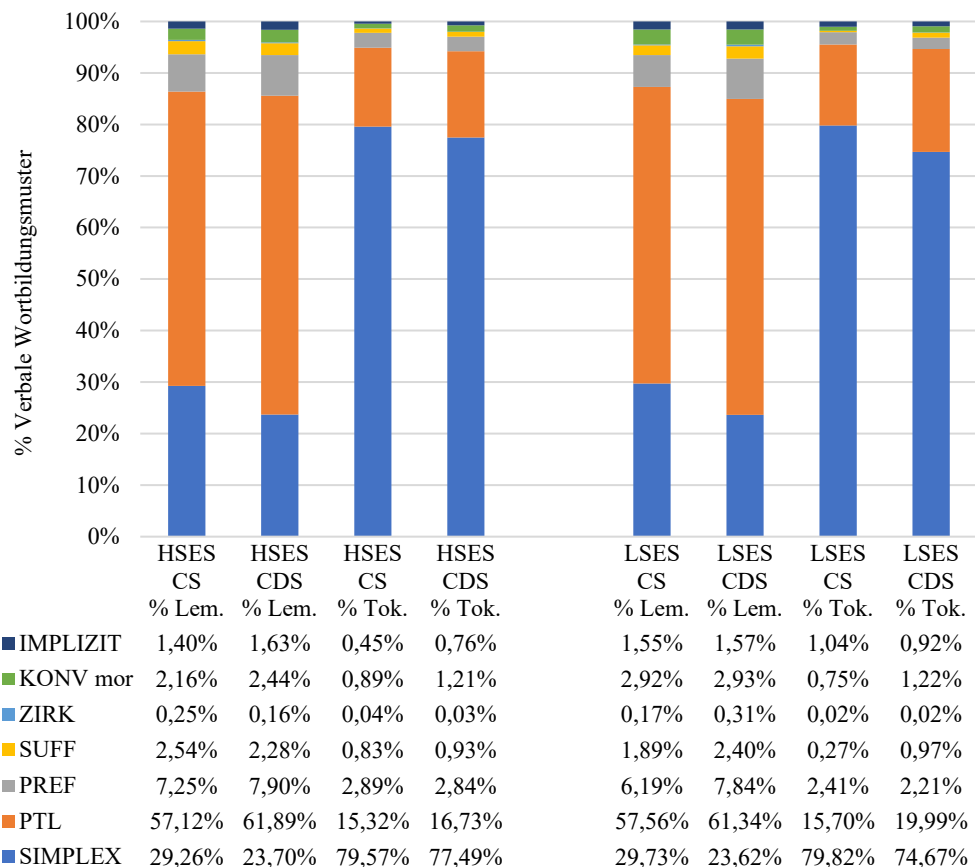


Abbildung 15. Prozentuale Verteilung der verbalen Wortbildungsmuster (Lemmas und Tokens) in CS und CDS im gesamten INPUT-Korpus (HSES und LSES)

Das für die Kinder und ihre Hauptbezugspersonen wesentlichste Wortbildungsmuster ist die Partikelverbbildung. In allen Korpora ist diese Kategorie die am häufigsten verwendete zur Bildung von Verben, bedingt durch die positionelle und prosodische Salienz der Partikeln, sowie deren größere Autonomie aufgrund ihrer Trennbarkeit. Dabei verwenden die Hauptbezugspersonen durchschnittlich 24 mehr verschiedene Partikel als die Kinder: 64 vs. 44 (Jan CDS vs. CS), 69 vs. 43 (LSES) und 85 vs. 60 (HSES), wobei die HSES-Kinder über das größte Repertoire an unterschiedlichen Partikeln verfügen. Auch Simplex-Verben kommen in ihrer Variation nicht an Partikelverben heran, was bedeutet, dass die untersuchten Kinder stärker mit Partikeln variieren als mit Basisverben. Einzig in ihrer Verwendungshäufigkeit (Tokens) dominieren die Simplex-Verben.

Die Partikelverben zeigen eine große Vielfalt sowohl im Input durch die Hauptbezugspersonen als auch im Output der Kinder, werden jedoch nicht so häufig wiederholt. In allen CDS-Daten bildet die Kategorie der Präfixverben mit

knapp 8 % (JAN, INPUT) das zweithäufigste Ableitungsmuster. Es wird deutlich, dass in CDS bei Kathi und Lena, deren Daten bis zu einem Alter von 3;0 reichen, der prozentuale Anteil der Präfixverben noch unter 6 % liegt. Präfixverben erhöhen sich demnach in CDS erst mit zunehmendem Alter der Kinder, wenn sie beginnen, diese selbst zu verwenden. Morphologische Konversionen, implizite Ableitungen und Suffigierungen werden zu einem geringeren Anteil verwendet, wobei erstere in CDS öfter höhere Frequenzen aufweisen. Zirkumfixe werden, wie bereits bei den Substantiven (vgl. Kap. 5.3), auch zur Derivation von Verben nur in geringem Ausmaß, oder im Fall von Kathis und Lenas CS bis 3;0 gar nicht verwendet.

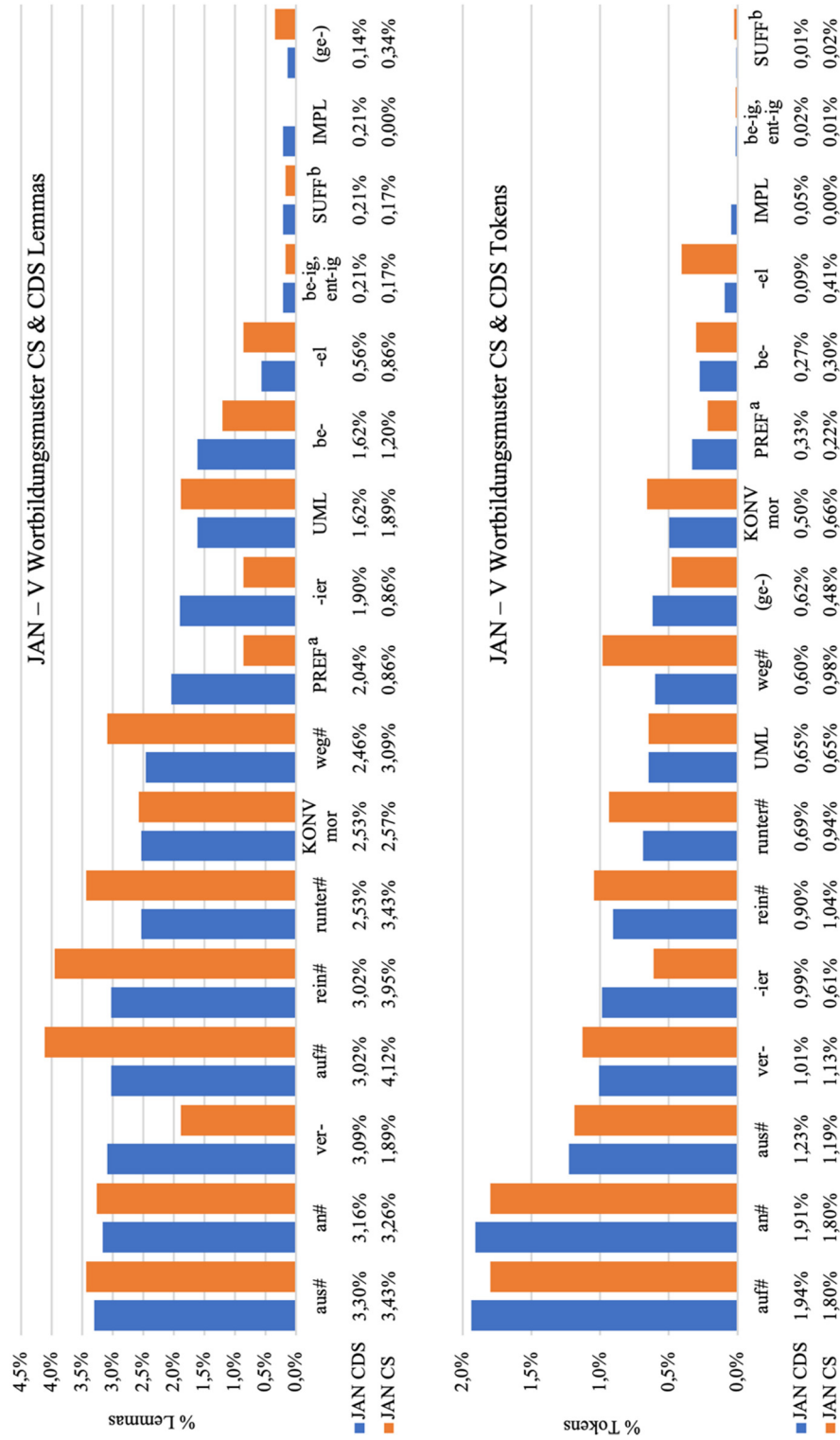
Die einzelnen Wortbildungsmuster und Affixe zur Bildung von Verben sind in absteigender Reihenfolge in CDS in Abbildung 16 für Jan, in Abbildung 17 für das HSES-Korpus und in Abbildung 18 für das LSES-Korpus übersichtlich dargestellt. Dabei wurden die einzelnen Wortbildungsmuster an der Grundgesamtheit der Verben berechnet. Die Abbildungen enthalten jeweils die häufigsten sechs trennbaren und drei untrennbaren Präfixe.

Die häufigsten Wortbildungsmuster in CS zur Ableitung von Verben sind bei allen Kindern die Partikel *auf#*, *rein#* und *aus#*, wo sich durchwegs höhere Lemmafrequenzen im Vergleich zu CDS zeigen. In Tokens sind neben den Partikeln *auf#* und *aus#*, letztere bei Jan und LSES, höhere Verwendungshäufigkeiten von *an#* bei Jan und von *hin#* bei den INPUT-Kindern feststellbar. Mehr Variation zeigt sich in CDS durch die Verwendung mehrerer unterschiedlicher Partikeln: *aus#*, *auf#*, *ab#*, *rein#*, *an#*, sowie des Präfixes *ver-*. In Tokens werden auch in CDS Bildungen mit der Partikel *auf#* am häufigsten verwendet, neben *aus#*, *an#*, *hin#* und morphologischen Konversionen (bei LSES).

Tendenziell kommt es in CS und CDS zu einer höheren Anzahl morphologischer Konversionen, während implizite Ableitungen mit Wortstammänderungen (IMPL) oder Stammvokaländerung (UML) weniger oft gebildet werden.

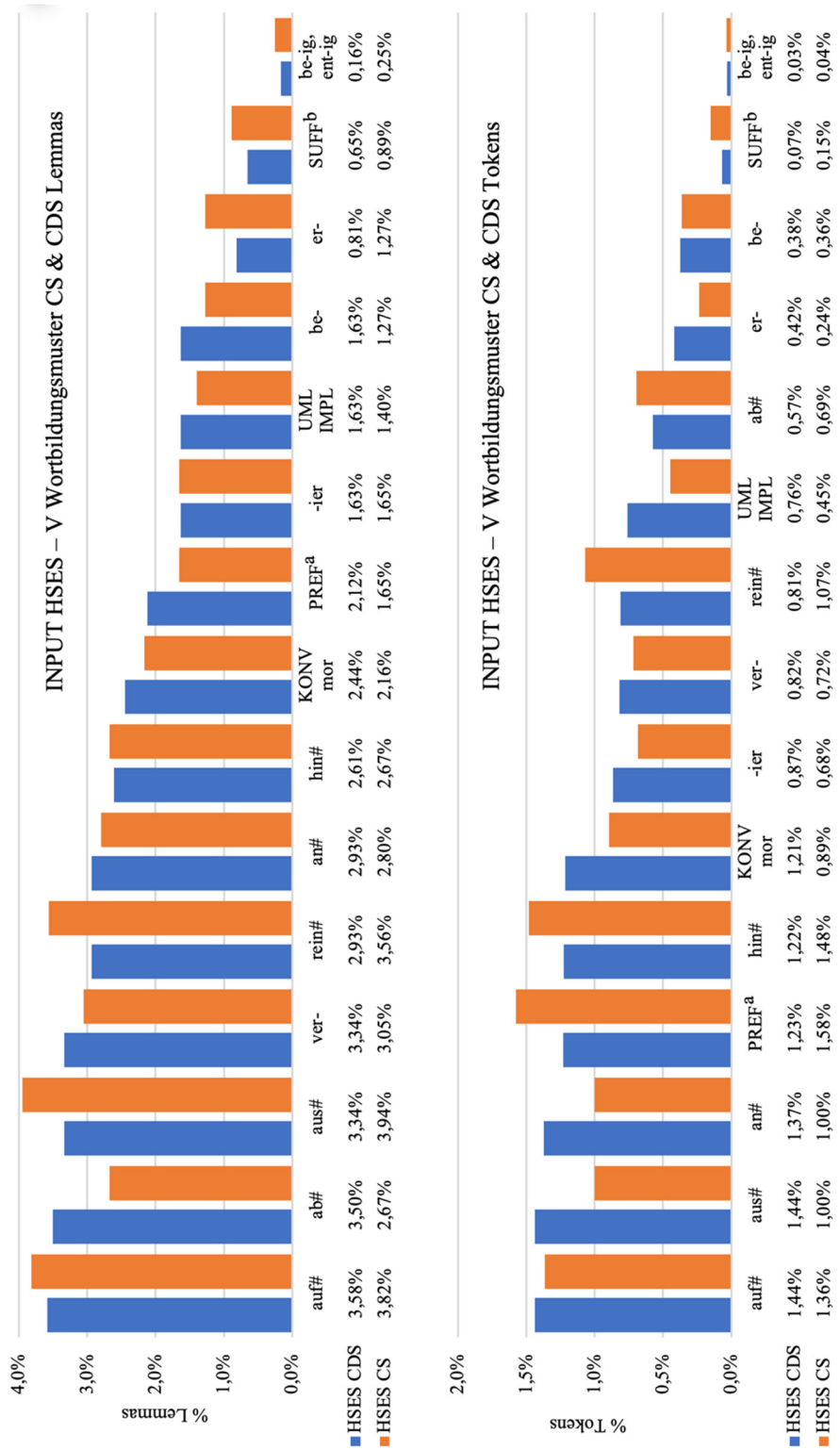
Zirkumfigierungen mit *be-...-ig* und *ent-...-ig* sind in CS und CDS bis 4;11 nur in geringem Ausmaß vorhanden, ebenso wie Suffigierungen mit *-er-n* und *-isier-en*. Bei den Suffixen ist nur im Fall von *-ier-en* in CDS und *-el-n* in Jans CS ein gewisses Maß an Produktivität in CDS feststellbar. Die meist verwendeten

untrennbaren Präfixe zur Ableitung von Verben sind in CS und CDS ver- und be-, sowie das Pseudopräfix ge- und im INPUT-Korpus er-. In geringerer Anzahl kommt es in CS zu Bildungen mit den untrennbaren Präfixen: ent-, über-, unter-, zer-, und zusätzlich in CDS mit wieder- (LSES) und de- (HSES).



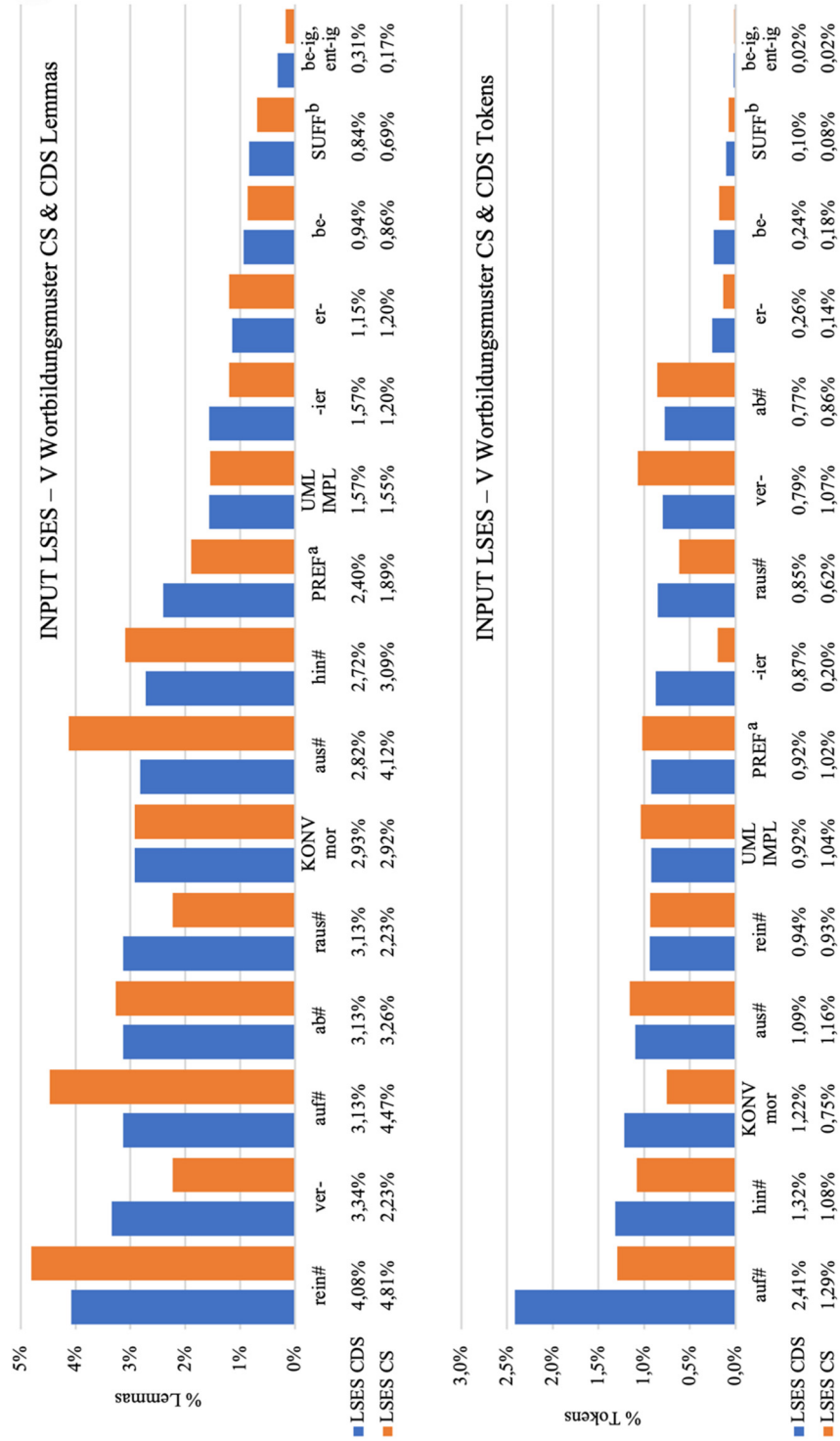
a. andere Präfixe in CS: ent-, über-, zer-; in CDS: zusätzlich unter-
 b. andere Suffixe in CS & CDS: -er, -isier

Abbildung 16. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum (1;3-4;11) nach absteigender Häufigkeit in CDS



a. andere Präfixe in CS: ent-, (ge-), über-, zer-, in CDS: zusätzlich de-, unter-
 b. andere Suffixe in CS & CDS: -el, -er, -isier

Abbildung 17. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS



a. andere Präfixe in CS: ent-, (ge-), über-, unter-, zer-; in CDS: zusätzlich wieder-
 b. andere Suffixe in CS & CDS: -el-, -isier

Abbildung 18. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Verben in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS

5.4.2 Erwerbsverlauf derivierter Verben

Die Entwicklungsverläufe derivierter Verben (V DERIV) im Vergleich zu Simplex-Verben (V SIMP) werden im Folgenden grafisch dargestellt und beschrieben. Aufgrund besserer Übersichtlichkeit werden bei Jans CS die monatlichen Erwerbsverläufe einerseits in Abbildungen 19 (Lem.) und 20 (Tok.) gesamt betrachtet (V DERIV vs. V SIMP), sowie andererseits die Verläufe präfigierter Verben (Präfixverben und Partikelverben) dem Verlauf bloßer Partikel (= die alleinige Verwendung der Partikel ohne das Basisverb, PTL) in Abbildungen 21 (Lem.) und 23 (Tok.) gegenübergestellt. Jans CDS-Verläufe sind in Abbildungen 22 (Lem.) und 24 (Tok.) zu finden. Bei den INPUT-Kindern sind die Entwicklungsverläufe nach Erhebungszeitpunkt (DP1–4) den Abbildungen 25 (Lem.) und 27 (Tok.) zu entnehmen, die jeweiligen CDS-Verläufe den Abbildungen 26 (Lem.) und 28 (Tok.).

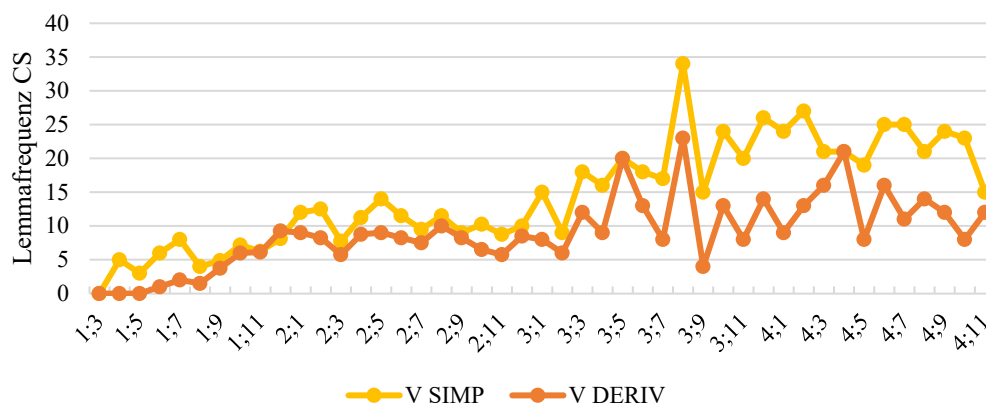


Abbildung 19. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Verben (V DERIV) vs. Simplex-Verben (V SIMP) in Lemmas (1;3–4;11)

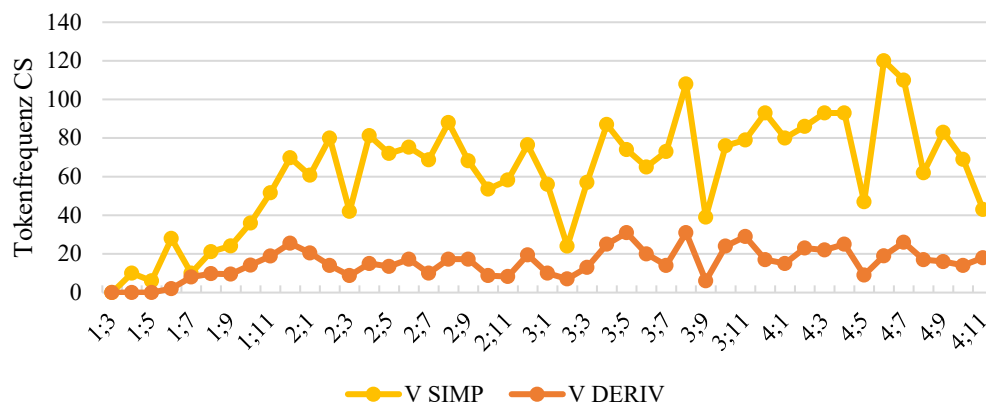


Abbildung 20. JAN-Korpus CS: Monatlicher Erwerbsverlauf derivierter Verben (V DERIV) vs. Simplex-Verben (V SIMP) in Tokens (1;3–4;11)

Sowohl bei einfachen als auch derivierten Verben ist bei allen Kindern eine stete Zunahme im Entwicklungsverlauf in Lemmas und Tokens erkennbar (vgl. Abbildungen 19, 20, 25 und 27). Überdies zeigt sich eine Präferenz für Simplex-Verben in Tokens, die sich bei Jan bis 4;11 auch in Lemmas findet. Bei allen INPUT-Kindern hingegen dominiert die Anzahl derivierter Verben in Lemmas. Das bedeutet, dass die HSES- und LSES-Kinder sehr viele unterschiedliche abgeleitete Verblemmas produzieren, diese jedoch nicht so häufig wie Simplex-Verben wiederholen. Im Fall von Jans CS verläuft die Kurve derivierter Verben zu Simplex-Verben in Lemmas über den gesamten Zeitraum ziemlich parallel. Gesamt betrachtet fallen die höheren Lemma- und Tokenfrequenzen der HSES-Gruppe im Vergleich zu den niedrigeren Frequenzen der LSES-Gruppe sowohl in CS als auch in CDS auf (vgl. Abbildungen 25, 26, 27 und 28).

Derivierte Verben treten im Vergleich zu derivierten Substantiven (vgl. Kap. 5.3.2) etwas früher in der Sprachproduktion der Kinder auf, wenngleich im selben Altersmonat bei Jan mit 1;6, siehe dazu Abbildung 19 und in Anhang A5.1 Tabelle 53, und bei Lena mit 1;8, siehe Anhang A5.3 Tabelle 55. Überdies entwickeln sich derivierte Verben schneller als derivierte Substantive, was sich an dem stärkeren Anstieg in Lemmas und Tokens zeigt (vgl. Abbildungen 19 und 20 zwischen 1;6 und 2;0). Lemma- und Tokenfrequenzen der V DERIV stehen in direkter Relation zu jenen der Partikelverben, da diese die überwiegende Mehrheit an derivierten Verben ausmachen (vgl. die parallelen Verläufe in CS und CDS in Abbildungen 22, 24–28).

Bevor Kinder komplexe deverbale Partikelverben bilden, kommt es zuerst zur isolierten Verwendung der Partikeln, wie bereits auf S. 54 in Kap. 2.6.2, eingeführt. Bei Jan zeigt sich die bloße Verwendung der Partikeln von Beginn an im Alter von 1;3. Dabei tritt die isolierte Partikel vor dem Partikelverb auf.²⁴ Gemäß Sommer-Lolei et al. (2021: 126) verwendet Jan im Alter von 1;6 zuerst die Partikel an# für an#hab-en, später im gleichen Monat das Partikelverb an#halt-en und erst mit 2;10 das Präfixverb be-halt-en. Es kommen verschiedene Partikeln in großer Häufigkeit

²⁴ Der Vergleich der Verwendung von Partikeln als Adverbien oder Präpositionen ergibt bei allen Partikeln, die ebenfalls isoliert oder in Partikelverben auftreten, eine spätere Emergenz, z. B. zu#: isolierte Partikel mit 1;4, im Partikelverb mit 1;7, als Adverb mit 2;0 und als Präposition mit 2;3.

(Tokens) vor, die sich jedoch mit Einsetzen der ersten Partikelverben rasch zurückbilden (vgl. Abbildungen 21 und 23).

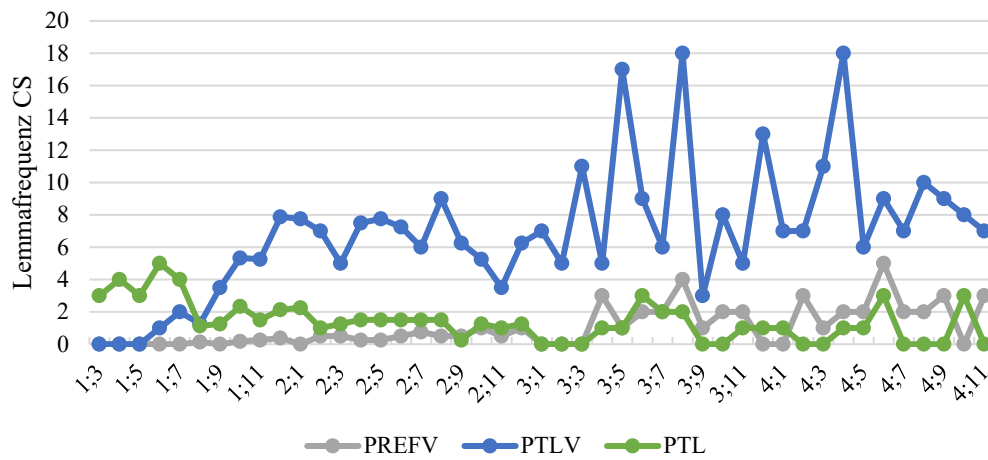


Abbildung 21. JAN-Korpus CS: Monatlicher Erwerbsverlauf von Präfixverben (PREFV), Partikelverben (PTLV) und Partikel (PTL) in Lemmas (1;3–4;11)

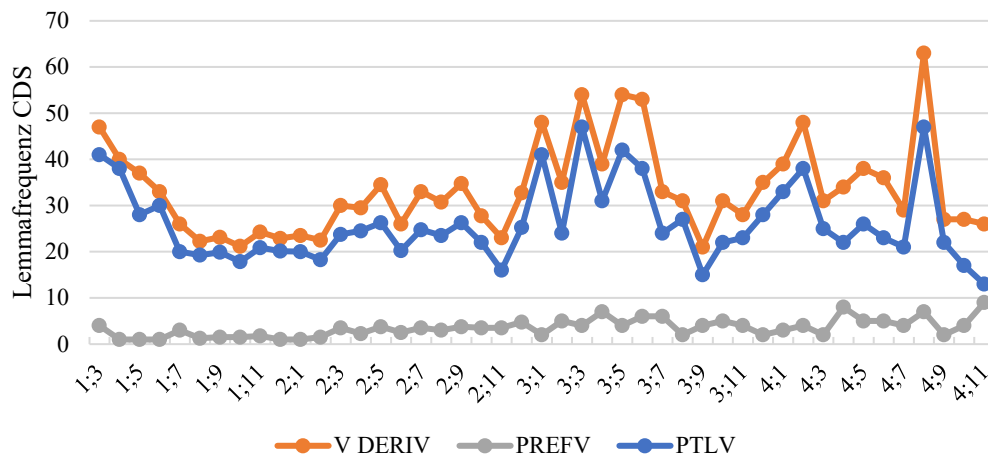


Abbildung 22. JAN-Korpus CDS: Verlauf derivierter Verben (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) in Lemmas (1;3–4;11)

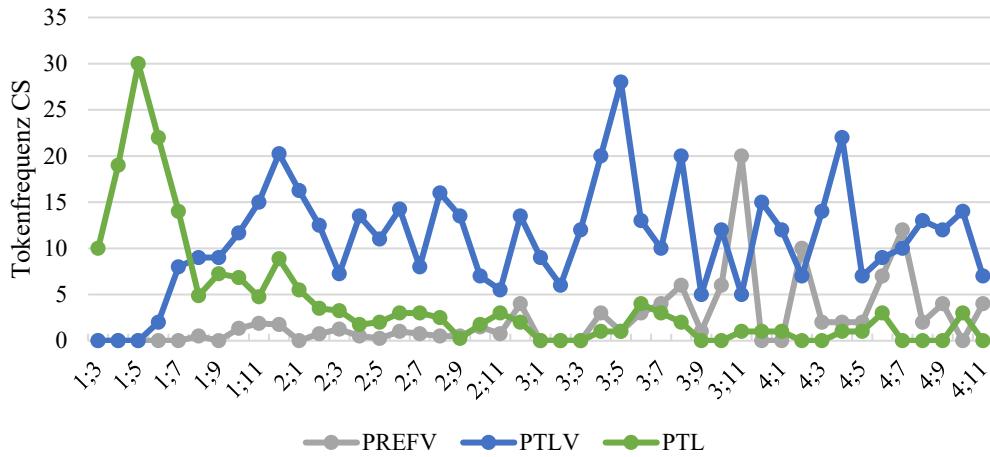


Abbildung 23. JAN-Korpus CS: Monatlicher Erwerbsverlauf von Präfixverben (PREFV), Partikelverben (PTLV) und Partikel (PTL) in Tokens (1;3–4;11)

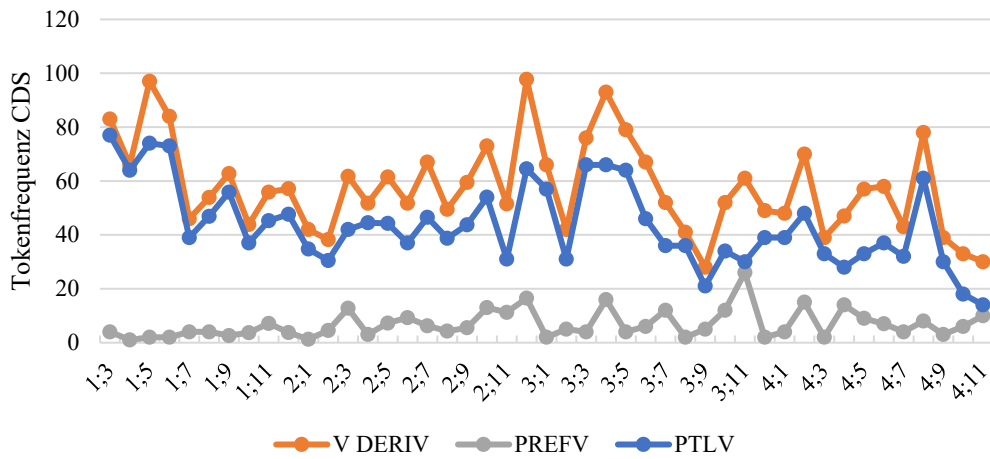


Abbildung 24. JAN-Korpus CDS: Verlauf derivierter Verben (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) in Tokens (1;3–4;11)

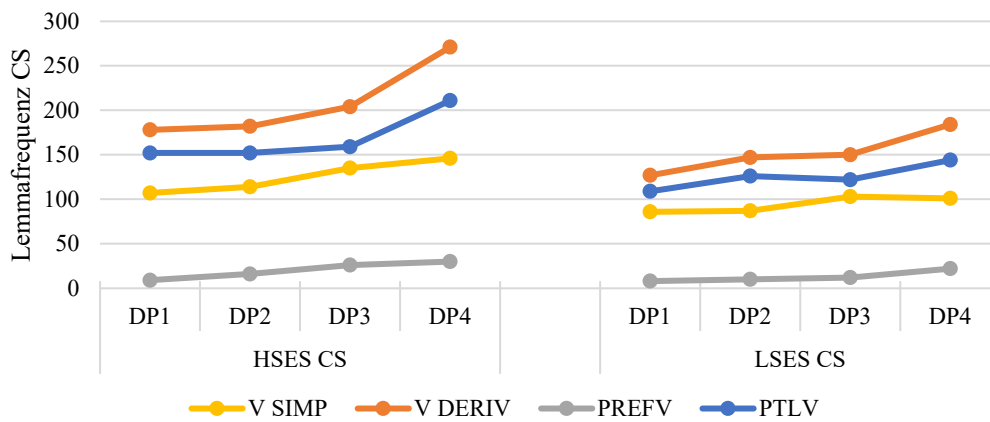


Abbildung 25. INPUT-Korpus CS: Erwerbsverlauf derivierter Verben gesamt (V DERIV) vs. Simplex-Verben (V SIMP), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Lemmas

Der frühe Erwerb der deutschen Derivationsmorphologie

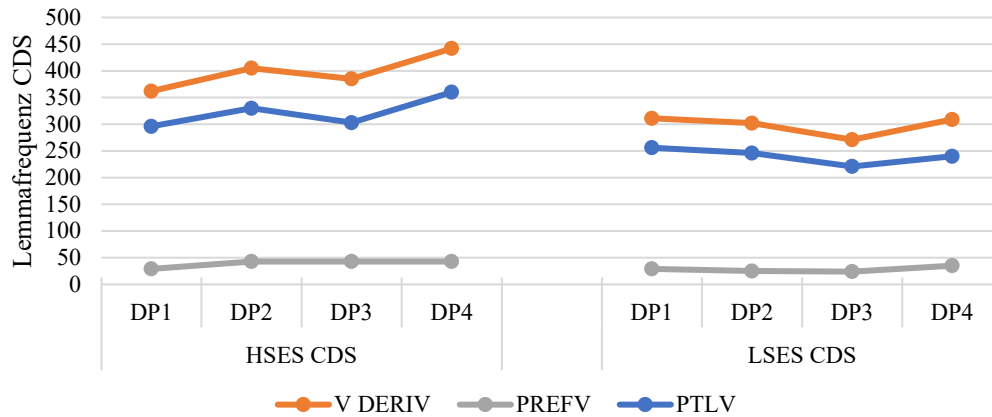


Abbildung 26. INPUT-Korpus CDS: Verlauf derivierter Verben gesamt (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Lemmas

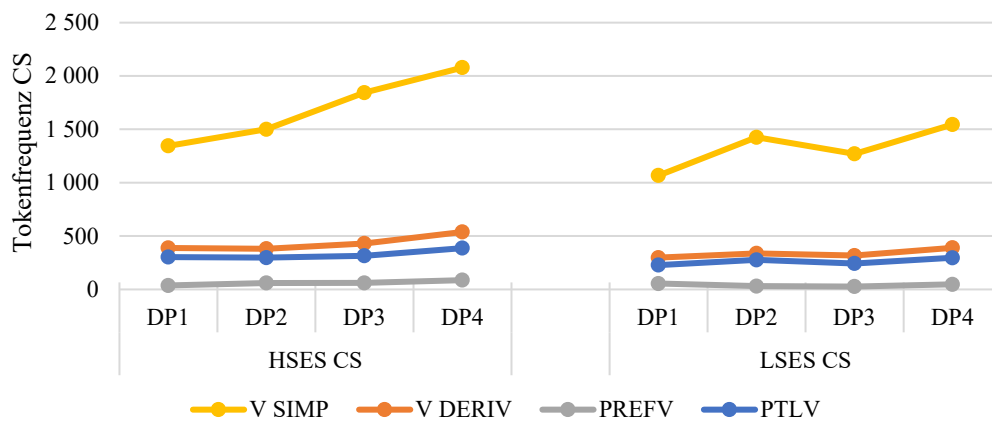


Abbildung 27. INPUT-Korpus CS: Erwerbsverlauf derivierter Verben gesamt (V DERIV) vs. Simplex-Verben (V SIMP), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Tokens

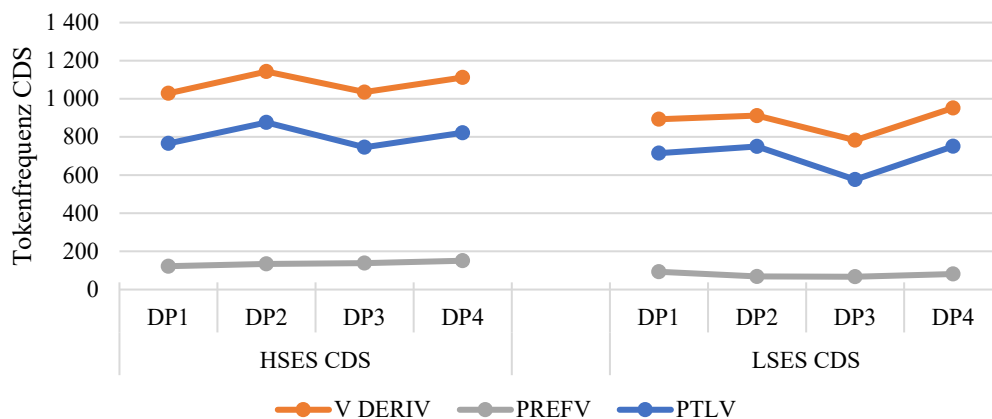


Abbildung 28. INPUT-Korpus CDS: Verlauf derivierter Verben gesamt (V DERIV), Präfixverben (PREFV) und Partikelverben (PTLV) nach Erhebungszeitpunkt (DP1–4) in Tokens

In Bezug auf den Erwerbsverlauf von Präfixverben (PREFV) und Partikelverben (= mit trennbarem Präfix, PTLV) ist feststellbar, dass PREFV bei Jan ab einem Alter von 3;4 an Variation (mehr verschiedene Lemmas) und ab 3;8 an Tokenfrequenz zunehmen, die sich auch bei den INPUT-Kindern wiederfindet, deren Aufnahmen ab einem Alter von durchschnittlich 3;1 beginnen. Die Tokenfrequenz der PREFV hingegen bleibt bei den LSES-Kindern in etwa gleich, während sie bei den HSES-Kindern schwach zunimmt (vgl. Abbildungen 21, 23, 25 und 27). Dies ist auch in der INPUT-CDS erkennbar (vgl. Abbildung 28).

Die Analyse der Differenz der Lemmafrequenzen derivierter Verben bei HSES- vs. LSES-Kindern zeigt, dass sich der Abstand im Lauf der Zeit zwischen den beiden Gruppen vergrößert, wie dies ebenso bei den derivierten Substantiven der Fall ist (vgl. Kap. 5.3.2). Die LSES-Kinder verfügen jeweils über weniger Verbderivationen als die HSES-Kinder (vgl. Abbildung 25). Die geringsten Differenzen zu den HSES-Kindern weisen die LSES-Kinder zu Datenzeitpunkt 2 auf, also in einem durchschnittlichen Alter von 3;4, wobei hier hauptsächlich weniger Partikelverben gebildet werden. Die größten Differenzen zu den HSES-Kindern finden sich zu Datenzeitpunkt 4 im Alter von 4;8. Diese Differenzen sind zu einem großen Teil den geringeren Frequenzen der Partikelverben geschuldet und zu einem wesentlich geringeren Teil jenen der Präfixverben und suffigierten Verben. Die Zunahme der Differenz zwischen den beiden untersuchten Gruppen lässt darauf schließen, dass die LSES-Kinder im direkten Vergleich zu den HSES-Kindern im Zeitraum zwischen 3;4–4;8 zunächst abfallen und nicht, wie dies zu erwarten gewesen wäre, rasch aufholen.

Bei genauerer Betrachtung der höchsten Werte in der größten Gruppe derivierter Verben, der Partikelverben in Jans Input, ist, wie bereits bei den derivierten Substantiven, ein Zusammenhang zu CS feststellbar. Höhere Lemma- und Tokenfrequenzen in CDS führen zwei bis drei Monate danach ebenfalls zu höheren Frequenzen in CS (vgl. Verlauf in Lem. in Abbildungen 21 und 22: PTLV CDS 3;1/3;3/4;2 vs. CS 3;3/3;5/4;4, in Tok. in Abbildungen 23 und 24: PTLV CDS 1;9/3;3 vs. CS 2;0/3;5).

Die Unterschiede derivierter Verben zu Simplex-Verben zwischen CS und CDS sind für das JAN-Korpus in Kap. 5.7.1.2 berechnet. Die mittleren Unterschiede

derivierter Verben zwischen den Kindern (CS vs. CS, S. 134) und den Hauptbezugspersonen (CDS vs. CDS, S. 136), werden nach SES-Zugehörigkeit in Kap. 5.8.1 auf statistisch signifikante Zusammenhänge untersucht. Für Gruppenunterschiede des INPUT-Korpus (CS vs. CDS) siehe Kap. 5.8.2.1, für die Mittelwertdifferenzen zweier Datenzeitpunkte (DP 1+2 vs. DP 3+4) siehe Kap. 5.8.2.2.

5.5 Wortbildung der Adjektive in CS und CDS

Die Auswertung der Durchschnittswerte in CDS aller Korpora, der Daten ab einem Alter von 3;0 bis 4;11,²⁵ zeigt, dass sich die Anzahl der Adjektive, die auf Wortbildung basieren, auf 49,46 % beläuft. Die Kategorie der suffigierten Adjektive ist dabei mit Abstand die größte Gruppe (27,85 %). Die restlichen Muster belaufen sich zusammen auf 21,61 % und bestehen aus: 9,05 % als Adjektive verwendete Perfektpartizipien (PP), 4,77 % Adjektivkomposita (ADJ KOMP), 3,91 % präfigierte Adjektive, 2,73 % Präsenspartizipien und 1,15 % Konversionen. Das bedeutet, dass 44,69 % aller der in CDS an das Kind gerichteten Adjektivlemmas abgeleitet sind. Im Vergleich zu den Daten der Kinder Jan, Kathi und Lena bis zu einem Alter von 3;0 (vgl. 26,10 % in Sommer-Lolei et al. 2021: 130) ergibt sich dadurch eine Erhöhung der abgeleiteten Adjektive in Lemmas zwischen 3;0 und 4;11 von 71,21 %.

Verteilung der Wortbildungsmuster bei Adjektiven

Die Verteilung derivierter, zusammengesetzter und Simplex-Adjektive in den Korpora, also die Frequenzen in Lemmas und Tokens, sowie der jeweilige prozentuale Anteil an der Gesamtheit der Adjektive in CS und CDS, finden sich im Detail in Anhang A4.3 Tabelle 52, jeweils über den gesamten Zeitraum, für alle Korpora dargestellt.

Die prozentuale Verteilung der Adjektive bei den drei Kindern Jan, Kathi und Lena und ihren Müttern sind in Abbildung 29, jene der 29 Kinder aus dem INPUT-Korpus und ihrer Hauptbezugspersonen (HSES und LSES) in Abbildung 30 übersichtlich dargestellt.

²⁵ Die Durchschnittswerte basieren auf den CDS-Daten des INPUT-Korpus (gesamt) und des JAN-Korpus (3;0–4;11).

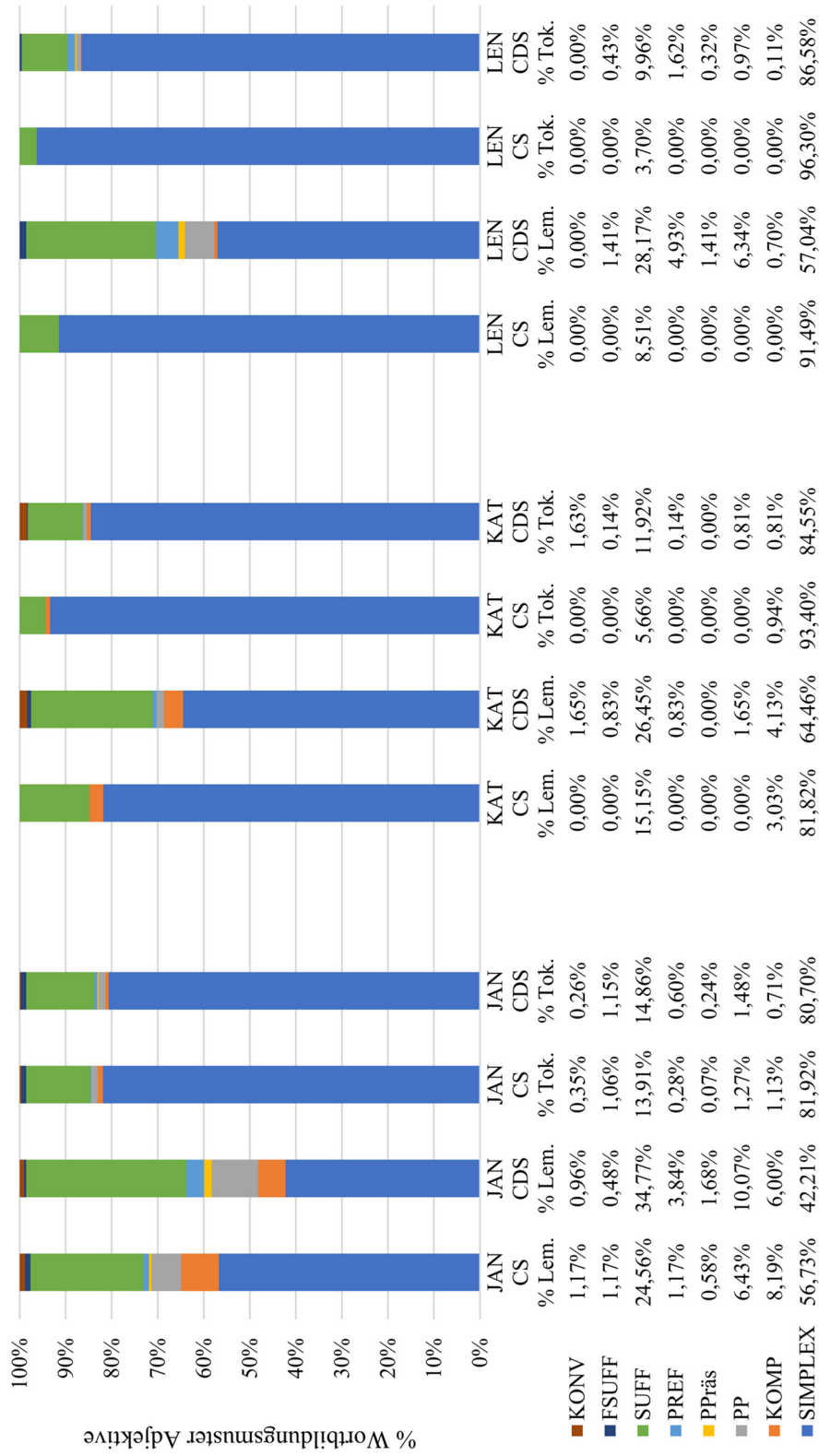


Abbildung 29. Prozentuale Verteilung der Wortbildungsmuster (Lemmas und Tokens) bei Adjektiven in CS und CDS in den Spontansprachkorpora von JAN, KAT und LEN (gesamter Zeitraum)

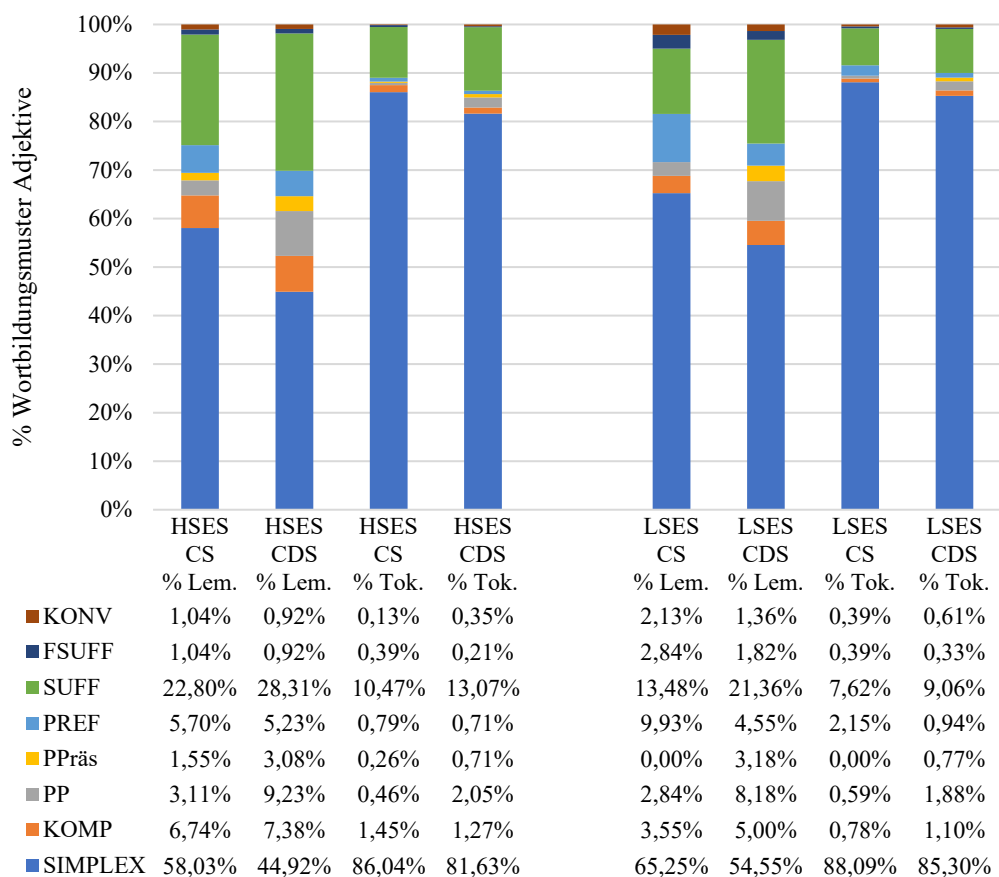


Abbildung 30. Prozentuale Verteilung der Wortbildungsmuster (Lemmas und Tokens) bei Adjektiven in CS und CDS im gesamten INPUT-Korpus (HSES und LSES)

Wie den Abbildungen 29 und 30 zu entnehmen ist, produzieren die Kinder und ihre Hauptbezugspersonen vorwiegend Simplex-Adjektive, wobei die Anteile in den frühen Daten von Kathi und Lena mit über 80 % Lem./90 % Tok. in CS besonders hoch sind. Suffixe zur Bildung derivierter Adjektive sind das häufigste Wortbildungsmuster mit durchschnittlich 20 % Lem./11 % Tok. in CS und 28 % Lem./12 % Tok. in CDS.²⁶

Generell ist festzuhalten, dass derivierte und zusammengesetzte Adjektive nur einen geringen Anteil aller verwendeten Adjektive bis 4;11 ausmachen. Selbst vor Schuleintritt im Alter von 6 Jahren nehmen abgeleitete Adjektive eine Randposition ein und zeigen sich auf wenige Muster beschränkt (vgl. Rainer 2010: 161; Mattes 2018: 313). Es zeigt sich bei der prozentualen Verteilung, dass Adjektivkomposita und Perfektpartizipien (PP) häufiger in CS und CDS auftreten als Konversionen, Fremdsuffigierungen, Präsenspartizipien und Präfigierungen. Wobei letztere zur

²⁶ Durchschnittswerte für CS und CDS errechnet aus den Werten von JAN, HSES und LSES.

Bildung von Adjektiven in HSES- und LSES-CS vergleichsweise häufig vertreten sind (siehe Abbildung 30).

Sofern Partizipien als Adjektive in den Korpora vorkommen, ist die Verwendung von Perfektpartizipien in CS und CDS wesentlich häufiger als jene von Präsenspartizipien. Bei den LSES-Kindern treten keine Präsenspartizipien auf, bei Kathi und Lena (CS) treten bis 3;0 keine Partizipien auf, ebenso wenig wie Konversionen, Fremdsuffigierungen und Präfigierungen.

Die einzelnen Wortbildungsmuster sind in absteigender Reihenfolge in CDS in Abbildung 31 für das JAN-Korpus, in Abbildung 32 für das INPUT-Korpus HSES und in Abbildung 33 für das INPUT-Korpus LSES übersichtlich dargestellt. Dabei wurden die einzelnen Wortbildungsmuster in Relation zur Grundgesamtheit an Adjektiven berechnet.

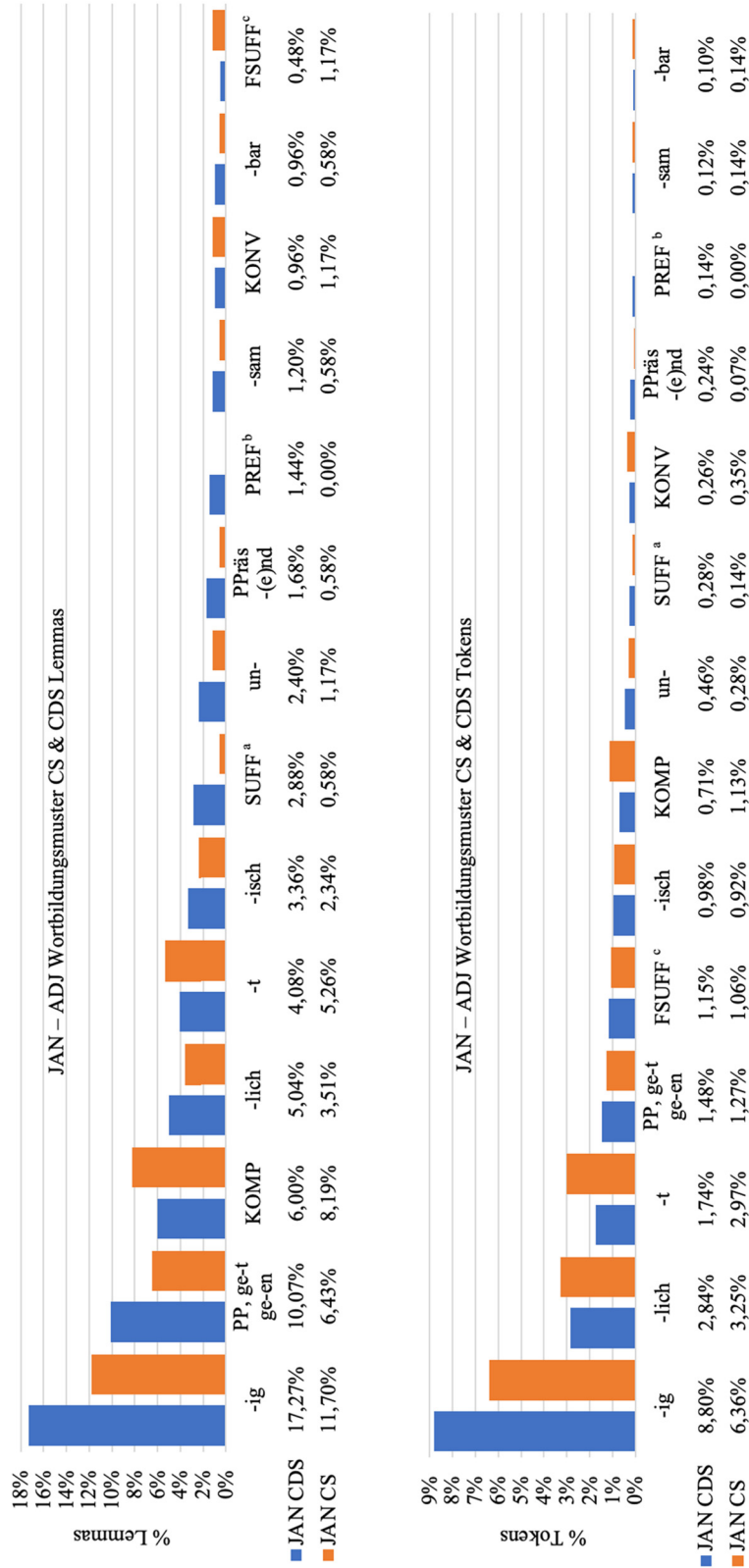
Daraus wird ersichtlich, dass -ig das am häufigsten verwendete Wortbildungssuffix in CS und CDS in Lemmas und Tokens zur Bildung von Adjektiven ist, mit Ausnahme der HSES-CS, wo -(s)t-Suffigierungen in Tokens prozentual häufiger sind. Danach folgen in CDS Perfektpartizipien und Adjektivkomposita, die über eine hohe Diversität verfügen. In CS hingegen werden mehr unterschiedliche Zusammensetzungen von Jan und den HSES-Kindern gebildet, während die LSES-Kinder präfigierte Bildungen mit ur- in größerer Vielfalt als ADJ KOMP verwenden.

Weitere produktive Wortbildungssuffixe in CS und CDS sind die bereits erwähnten Suffixe -t bzw. -st, zur Ableitung von Ordinalzahlen, die zwar in der Varietät der Lexeme weitgehend auf den Zahlenraum bis zwölf eingeschränkt sind, die einzelnen Bildungen jedoch sehr häufig wiederholt werden, vor allem von den HSES-Kindern (vgl. Tokenübersicht in Abbildung 32). Ordinalzahlen von Zehnern kommen in CDS in den Bildungen zehn-t, elf-t, zwölf-t, fünf/sech/siebzehn-t, zwei/drei/vier/fünf-undzwanzig-st und dreißig-st vor und nur in Jans CS in zehn-/elf-/zwölf-t und fünfzehn-t. Analog zur Bildung von Ordinalzahlen tritt wieviel-t isoliert in CDS auf (JAN 3;1, D31-AND 4;6). Überdies ist das Ableitungssuffix -lich erwähnenswert, da es in Jans CS und CDS nach Suffigierungen mit -ig und vor Ableitungen mit -(s)t am häufigsten in Tokens verwendet wird. Das Suffix

-isch tritt bei Jan und HSES in geringerer Anzahl und in allen Korpora in wesentlich niedrigerer Verwendungshäufigkeit im Vergleich zu -(s)t auf.

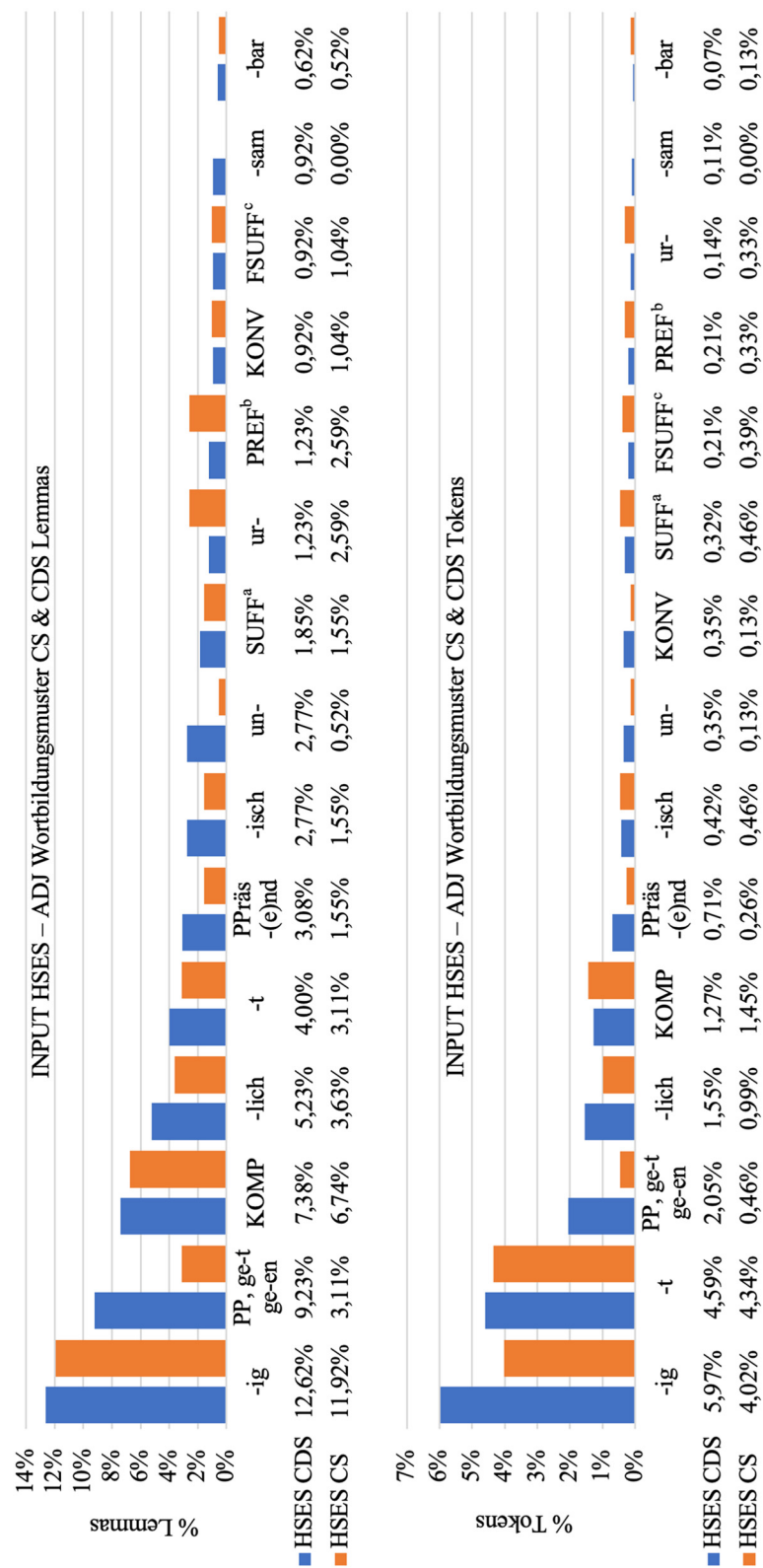
Andere Suffixe, darunter -bar, -en, -ern, -ert, -haft, -los, -sam, -voll, und Fremdsuffixe, wie -al, -ant, -iell und -ös treten nur in geringer Anzahl in CDS und CS auf, wobei der Anteil der fremdsuffigierten Adjektive in LSES-CS und -CDS jenen der anderen Suffixe in Lemmas und Tokens übersteigt. Andere Fremdsuffixe als die genannten kommen in den Korpora nicht vor.

Präfigierungen werden im Vergleich zu ihren Hauptbezugspersonen von den HSES- und LSES-Kindern in größerer Anzahl in Lemmas und Tokens verwendet, mit Ausnahme des Präfixes un-, das in CDS wesentlich frequenter ist.



a. andere Suffixe in CS: -e; in CDS: -en, -ern, -ert, -los, -voll
 b. andere Präfixe in CDS: aller-, an-, ur-
 c. Fremdsuffixe in CS & CDS: -al-, -ant

Abbildung 31. JAN-Korpus: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum (1;3–4;11) nach absteigender Häufigkeit in CDS

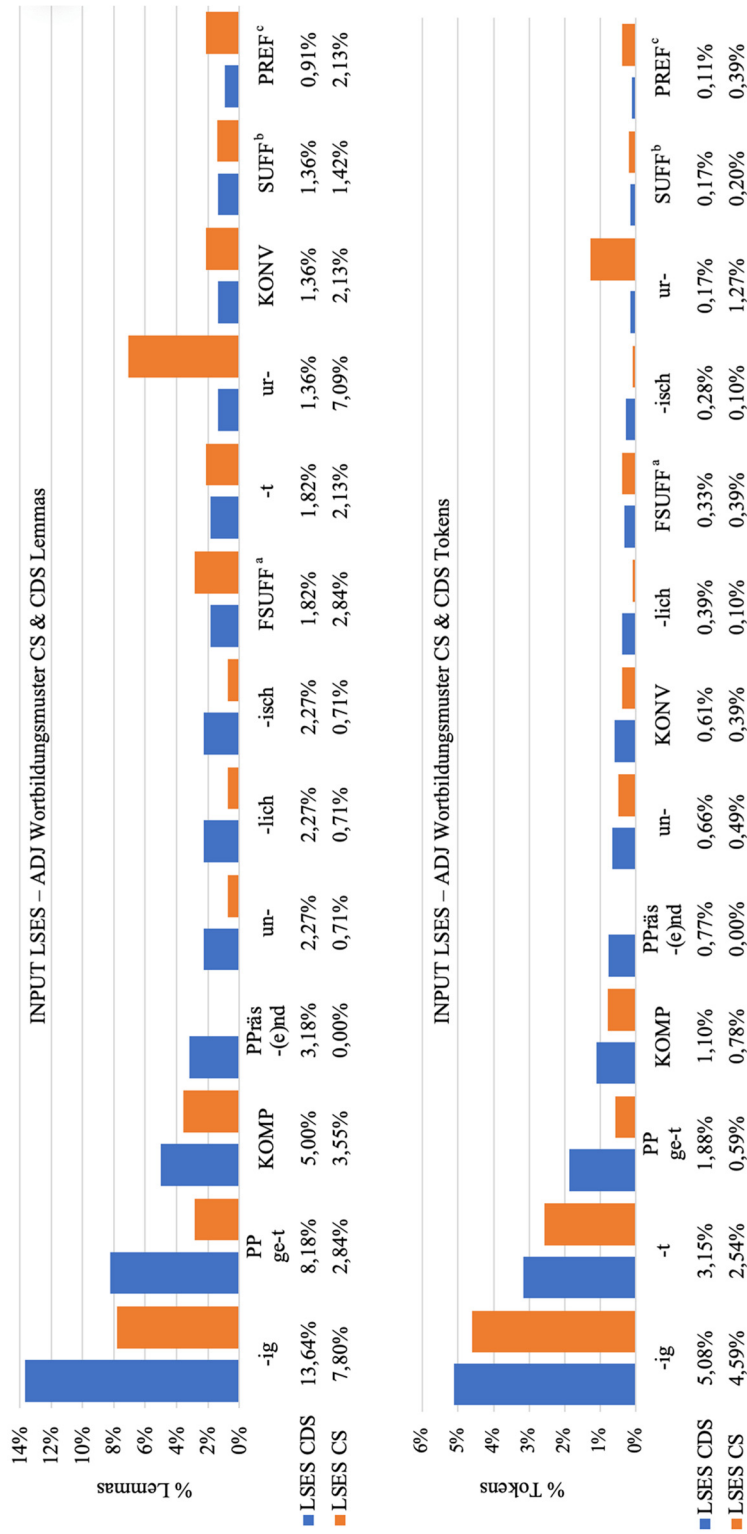


a. andere Suffixe in CS: -en, -ert; in CDS: zusätzlich -ern, -haft, -voll

b. andere Präfixe in CS: aller; in CDS: zusätzlich an-

c. Fremdsuffixe in CS: -al, -ico; in CDS: zusätzlich -ant, -iell

Abbildung 32. INPUT-Korpus HSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS



a. Fremdsuffixe in CS & CDS: -al, -ant, -iell, -ös
 b. andere Suffixe in CS & CDS: -en, -ert, -sam
 c. andere Präfixe in CS: aller-, mini-; in CDS: zusätzlich an-

Abbildung 33. INPUT-Korpus LSES: Prozentuale Verteilung (Lemmas und Tokens) der Wortbildungsmuster der Adjektive in CS und CDS im gesamten Zeitraum nach absteigender Häufigkeit in CDS

5.6 Emergenz und potenzielle Produktivität

Die Häufigkeit eines Wortbildungsmusters im Output eines Kindes (CS) alleine gibt nur ungenügend Aufschluss über die Produktivität desselben im kindlichen Sprachsystem. Für die Produktivität ist die Frequenz eines Musters zwar bedeutend (vgl. Bybee 2006), jedoch wie Cordes (2014: 117) feststellt, nicht der einzig wesentliche Aspekt im Sprachlernprozess. In Bezug auf die Häufigkeiten eines Musters im Input (CDS) wurde bereits festgestellt, dass Lemmafrequenz im Bereich der Entwicklung der Grammatik im Zuge des Spracherwerbs eine größere Rolle spielt als Tokenfrequenz (Bybee 1995: 430). Die Frequenz, bzw. die Häufigkeit eines Musters als feststellbare Größe, spielt in jedem Fall insofern eine Rolle, als eine kritische Masse in Lemmas in CDS vorhanden sein muss (vgl. Marchman & Bates 1994; Van der Schuit et al. 2011; Serra i Raventós 2014). Darüber hinaus spielt beim Spracherwerb die morphologische Produktivität eines Musters (siehe Kap. 2.5.2) eine Rolle (vgl. Korecky-Kröll et al. 2017; Dressler, Sommer-Lolei, et al. 2019; Sommer-Lolei et al. 2021). Als Indikator zur Vorhersage, ob ein Muster tendenziell eher früher oder später in CS auftritt und ab welchem Zeitpunkt es als gemeistert gilt, hat sich die Inputfrequenz im Rahmen der Studie von Sommer-Lolei et al. (2021: 130 ff.) für Derivationen im Spracherwerb des Deutschen tatsächlich als wenig verlässlich erwiesen.

Zusätzlich zur Analyse der Lemma- und Tokenfrequenzen in CS und CDS finden sich im Folgenden Altersangaben zum ersten Auftreten bestimmter Wortbildungsmuster, sowie überdies die jeweiligen Zeitpunkte der (partiellen) Erfüllung des Mini-Paradigma-Kriteriums (siehe zur Methodik Kap. 4.5), bei Jan bis zu einem Alter von 4;11 (siehe Tabelle 9) in Erweiterung zu den Daten in Sommer-Lolei et al. (2021: 127–130). Für die INPUT-Kinder, deren Aufnahmen erst in einem Alter von 3 Jahren beginnen, können die ersten Emergenzen ebenso wenig bestimmt werden wie die Erfüllung des Mini-Paradigma-Kriteriums zur Feststellung der potenziellen Produktivität eines Musters, daher werden sie in diesem Kapitel nicht behandelt.

Tabelle 9. Emergenzen (Em.) und Zeitpunkte der Erfüllung des Mini-Paradigma-Kriteriums (MPK) der relevantesten Derivationsmuster bei Substantiven, Verben und Adjektiven in CS im gesamten Zeitraum²⁷

	Muster	JAN		KAT		LEN		Beispiel
		Em.	MPK	Em.	MPK	Em.	MPK	
Substantive	-er (NAG)	1;8	1;10	1;11	2;10	1;9	2;3	Reit-er
	-er (NIN)	1;8	1;10	2;3	2;10	1;9	2;3	Flitz-er, Last-er
	-e	1;9	2;0	2;6	2;6	2;5	(2;10) ^a	Rutsch-e
	Konversion	1;8	1;10	2;8	3;0	2;4	–	Sitz, Fall
	-ung	2;0	2;3	2;9	–	2;7	(2;11)	Rett-ung Wohn-ung
Verben	PTLV	1;6 (<i>an</i>)	1;9	2;3 (<i>ein</i>)	2;6	1;8 (<i>auf</i>)	2;1	auf#heb-en
	Konversion	1;6	2;1	2;3	3;0	2;2	–	stempel-n
	Implizit	1;8	2;0	2;3	2;8	2;10	(2;11)	leg-en, fürcht-en
	ent-...-ig	2;6	(3;10)	–	–	–	–	ent-schuld-ig-en
	ver-	1;8	2;6	2;10	–	2;7	(2;7)	ver-steck-en
	(ge-)	1;11	2;9	2;6	(2;8)	2;6	–	ge-fall-en
	be-	2;10	3;0	2;9	–	2;8	–	be-komm-en
	-el	1;9	3;7	2;10	–	–	–	kling-el-n
	-ier	1;11	3;6	2;8	–	–	–	prob-ier-en
-er	2;11	(2;11)	–	–	–	–	steig-er-n	
Adjektive	-ig	1;8	1;11	2;8	–	2;9	–	durst-ig rutsch-ig
	-lich	1;9	4;6	2;9	–	–	–	glück-lich
	-(s)t	1;8	2;11	2;8	3;0	2;7	(2;5)	zwei-t, vierzig-st
	PP	2;7	2;8	–	–	–	–	ab-ge-riss-en
	-isch	2;9	(4;0)	–	–	3;0	–	quadrat-isch
	un-	2;6	(3;11)	–	–	–	–	un-fair

a. Alter zum Zeitpunkt der partiellen Erfüllung des Kriteriums wird in Klammern angeführt, vgl. Kap. 4.5, Beispiel (4).

Die ersten Emergenzen der relevantesten Derivationsmuster bei Jan, Kathi und Lena, sowie die Zeitpunkte der Erfüllung des MPK, sind mit Beispielen in Tabelle 9 übersichtlich dargestellt. Zusammengefasst sind die ersten Wortbildungsmuster, die in CS tendenziell zuerst auftreten, Partikelverben und -er-Suffigierung zur Bildung von Nomina Agentis und Nomina Instrumenti. Mit Ausnahme von -ig, und -t, die bei Jan ebenfalls sehr früh auftreten, kommen die meisten Wortbildungsmuster für Adjektive erst später im Spracherwerb vor. Alle Kinder erfüllen zuerst das MPK für Partikelverben, was auf die Wichtigkeit dieses Wortbildungsmusters im Spracherwerb hinweist: Jan mit 1;9, Kathi mit 2;6 und Lena mit 2;1. Danach für -er Nomina Instrumenti und Nomina Agentis: Jan mit 1;10, Kathi mit 2;10 und Lena mit 2;3. Die folgenden Kapitel behandeln die

²⁷ Modifiziert übernommen aus: Sommer-Lolei et al. (2021: 127 f.).

Emergenzen der Wortbildungsmuster und die Erfüllung des Mini-Paradigma-Kriteriums nach Wortklasse im Detail.

5.6.1 Derivierte und zusammengesetzte Substantive

5.6.1.1 Emergenz

Das erste Suffix in CS ist in allen drei Kindern -er zur Bildung deverbalen oder denominaler Nomina Agentis (z. B. Reit-er, Musik-er) und Nomina Instrumenti (z. B. Weck-er, Last-er) vgl. Tabelle 9. Deverbale -er-Suffigierung tritt bei Kathi im Alter von 1;11 (NAG) und 2;3 (NIN), bei Lena zeitgleich in den beiden Wortbildungskategorien mit 1;9, ein Monat nach ihrem ersten Substantivkompositum (N KOMP), und bei Jan im Alter von 1;8 zusammen mit seiner ersten Konversion und seinem ersten nominalen Kompositum auf. Kathi bildet ihr erstes N KOMP im Alter von 2;1. Das zweite Muster, das tendenziell früh auftritt, ist deverbale -e-Suffigierung, als Nomina Instrumenti (z. B. Rutsch-e, Wieg-e): bei Jan mit 1;9 und bei Kathi mit 2;6, zwei Monate vor ihrer ersten Konversion (2;8). Bei Lena tritt das Suffix -e im Alter von 2;5, ein Monat nach ihrer ersten Konversion (2;4) auf. Das Suffix -ung zur deverbalen Ableitung von Substantiven (z. B. Rett-ung, Wohn-ung) tritt bei allen Kindern später auf: bei Jan mit 2;0, bei Kathi mit 2;9 und bei Lena mit 2;7.

5.6.1.2 Erfüllung des Mini-Paradigma-Kriteriums

Jan und Lena erfüllen das MPK für -er von allen untersuchten Suffixen in abgeleiteten Substantiven als erstes im Alter von 1;10 (JAN), gleichzeitig mit Konversionen (z. B. Sitz, Fall) und ein Monat später für N KOMP (1;11), und im Fall von Lena als einziges Ableitungssuffix vollständig (2;3), neben N KOMP mit 2;4. Kathi erfüllt das Kriterium zuerst für nominale Komposita (2;4), für -e-Derivationen (2;6) und vier Monate danach für -er (2;10), für Konversionen erfüllt sie es mit 3;0. Das Suffix -ung erreicht nur bei Jan im Alter von 2;3 potenzielle Produktivität, und teilweise bei Lena mit 2;11, durch die Verwendung des Suffixes mit drei verschiedenen Basen (vgl. Tabelle 9). Wie in Sommer-Lolei et al. (2021: 127) bereits festgestellt, handelt es sich bei den frühen Substantiven auf -ung um sehr opake Bildungen (z. B. Ahn-ung) oder um konkrete Objekte (z. B. Heiz-ung). Es zeigt sich bei Konversionen eine frühe Erfüllung des Kriteriums bei Jan (1;10) und später im Alter von 3;0 bei Kathi. Das Kriterium wird nicht erfüllt für

Konversionen bei Lena und bei Kathi für -ung-Derivationen. Mit Ausnahme der potenziellen Produktivität der -er-Ableitungen ist ein heterogenes Bild feststellbar.

5.6.2 Derivierte Verben

5.6.2.1 Emergenz

Wie bereits festgestellt, sind Partikelverben das frequenteste Muster in CS und CDS zur Ableitung von Verben im ganzen untersuchten Zeitraum (vgl. Abbildungen 14 und 15) bei allen Kindern und ihren Hauptbezugspersonen. Es handelt sich ebenso um jenes Muster, das als erstes in CS auftritt (vgl. Tabelle 9): sehr früh bei Jan mit 1;6 und bei Lena mit 1;8, etwas später bei Kathi mit 2;3. Gleichzeitig mit dem ersten Partikelverb (= trennbares Präfix, z. B. an#schau-en, zu#mach-en) tritt auch bei Jan und Kathi die erste Verbkonversion (z. B. stempel-n) auf, bei Lena vier Monate danach mit 2;2. Als die wesentlichsten untrennbaren Präfixe in CS haben sich ver-, und be-, z. B. ver-steck-en, be-halt-en herausgestellt sowie das Pseudopräfix ge-, wie in ge-fall-en.

Das Zirkumfix ent-...-ig, wie in ent-schuld-ig-en, tritt allein bei Jan im Alter von 2;6 auf. Die beiden Mädchen bilden bis zum Ende der Aufnahmen mit 3;0 keine zirkumfigierten Verben.

5.6.2.2 Erfüllung des Mini-Paradigma-Kriteriums

Das MPK wird von allen drei Kindern für Partikelverben drei Monate nach der ersten Emergenz erfüllt (vgl. Tabelle 9). Jan erfüllt es mit 1;9 und damit ein Monat vor -er-Suffigierung bei Substantiven, Kathi mit 2;6, zeitgleich mit deverbalen Ableitungen auf -e, und Lena mit 2;1, zwei Monate vor abgeleiteten Substantiven auf -er. Bei Lena ist dies das einzige Ableitungsmuster für Verben, das potenziell produktiv wird. Partuell erfüllt sie das Kriterium nur für das Präfix ver- mit 2;7, durch die Verwendung der Basis in drei verschiedenen Typen.

Für Verbkonversionen, z. B. schaukel-n, erfüllen Jan mit 2;1 und Kathi mit 3;0 das MPK. Bei präfigierten Verben (untrennbare Präfixe) erfüllt nur Jan alle erforderlichen Anforderungen mit 2;6 für ver-, für das Pseudopräfix ge- mit 2;9 und für be- mit 3;0, dabei verwendet er zunächst die Basis in drei Typen, bevor er das Präfix mit drei verschiedenen Basen kombiniert (vgl. Sommer-Lolei et al. 2021: 129). Das Pseudopräfix ge- erfüllt Kathi nur teilweise mit 2;8 (Basis in drei Typen).

In Bezug auf implizite Derivationen, die eine Ab- oder Umlautung des Stammvokals oder eine Stammmodifikation beinhalten, z. B. Umlautung in Futter → fütter-n, zeigt sich das Kriterium bei Jan mit 2;0 und bei Kathi mit 2;8 erfüllt, obwohl es sich in CS und CDS um ein vergleichsweise selten vorkommendes Wortbildungsmuster handelt (vgl. Abbildung 14). Dies begründen Sommer-Lolei et al. (2021) mit hochfrequenten Verben des unproduktiven Ablautungsmusters i(e) → e, wie lieg-en → leg-en und sitz-en → setz-en, die dieser Kategorie angehören.

Suffigierte und zirkumfigierte Verben zeigen sich bei Jan erst später als andere verbale Wortbildungsmuster potenziell produktiv. Dabei erfüllt Jan das MPK einerseits für das Suffix -ier-en wie z. B. in telefon-ier-en mit 3;6, andererseits für -el-n wie z. B. in streich-el-n mit 3;7. Teilweise zeigt sich das Kriterium erfüllt für das Suffix -er-n wie z. B. in steig-er-n mit 2;11, ebenso wie für das Zirkumfix ent-...-ig, durch die Verwendung der Basis in drei Typen. Beide Mädchen haben keine zirkumfigierten Verben bis 3;0, und entweder keine (LEN) oder zu wenige (KAT) suffigierte Verben, um das Kriterium zu erfüllen.

5.6.3 Derivierte Adjektive

5.6.3.1 Emergenz

Zur Ableitung von Adjektiven treten in CS bis zu einem Alter von 4;11 bei Jan, und bis 3;0 bei den beiden Mädchen zunächst zwei Suffixe auf: -ig und -t (zeitgleich bei Jan 1;8 und Kathi 2;8, mit zwei Monaten Unterschied bei Lena -t mit 2;7 und -ig mit 2;9, vgl. Tabelle 9). Bei Kathis einzigem Auftreten des Suffixes -ig handelt es sich um die fehlerhafte Bildung *teu-ig-er statt teuer, vermutlich in Anlehnung an das Antonym billig-er. Im Alter von 3;0 kommt es bei Jan zu dem Neologismus *bahn-ig-es²⁸, was ebenso die Dominanz dieses Wortbildungsmusters bei Adjektiven widerspiegelt. Das Suffix -lich tritt ein Monat später bei Jan (1;9) und Kathi (2;9) auf, nicht jedoch bei Lena. Abgeleitete Adjektive auf -isch treten vereinzelt bei Jan (2;9) und Lena (3;0) auf. Bis zu einem Alter von 4;11 verwendet Jan das Präfix un- zur Bildung von Adjektiven zweimal, zuerst in un-sicht-bar (2;6) und wesentlich später in un-fair (4;1). Die beiden Mädchen bilden bis 3;0 keine

²⁸ In der Äußerung (3;0): *JAN: ja, das is(t) was bahniges@n. Die Markierung @n kennzeichnet auf der morphologischen Codierungszeile den Neologismus.

präfigierten Adjektive. Weitere Suffixe, die bei Jan bis 4;11 isoliert auftreten sind: -bar (2;0, sicht-bar) und -sam (2;2, lang-sam).

5.6.3.2 Erfüllung des Mini-Paradigma-Kriteriums

Das Mini-Paradigma-Kriterium wird von Jan zuerst für -ig (1;11) erfüllt, was zugleich das in CS und CDS häufigste Muster zur Ableitung von Adjektiven ist (vgl. Abbildungen 31 und 32, sowie Tabelle 9)). Das MPK für Perfektpartizipien (PP, z. B. ab-ge-riss-en) erfüllt er mit 2;8, danach folgen die Suffixe -(s)t (2;11) und -lich (4;6). Partiiell erfüllt Jan das Kriterium für un- mit 3;11 (Basis in drei Typen) und -isch mit 4;0 (Suffix mit drei verschiedenen Basen). Das MPK wird von den beiden Mädchen nur für -t (partiiell) erfüllt: bei Kathi mit 3;0 und bei Lena teilweise mit 2;5, aufgrund der Verwendung der Basis in drei verschiedenen Typen.

5.6.4 Frühe Emergenz vs. späte Meisterung

Die Analyse der Kindersprache in den vorangegangenen Kapiteln 5.6.1, 5.6.2 und 5.6.3 hat für Substantive, Verben und Adjektive gezeigt, dass die frühe Emergenz eines Wortbildungsmusters in CS nur in einigen Fällen zu potenzieller Produktivität führt. Dabei ist die potenzielle Produktivität im Sinne der regelhaften Anwendung eines Wortbildungsmusters nicht mit einer etwaigen Meisterung desselben gleichzusetzen. Erste Emergenzen derivierter Lexeme sind als Ausgangspunkt zu verstehen, von wo aus sich die sprachliche Entwicklung eines Kindes fortsetzt und schließlich ab einem bestimmten Punkt von diesem vollständig gemeistert wird (*full mastery*, Berman 2004). Daher ist der Zeitpunkt der Meisterung eines Wortbildungsmusters als wesentlich wichtiger anzusehen als dessen erstes Auftreten.

Als gemeistert gilt ein Wortbildungsmuster, wenn es ab einem bestimmten Zeitpunkt in der Sprachentwicklung konsistent korrekt gebildet wird. Basierend auf diesen Annahmen wurden die Daten des Kindes Jan und der INPUT-Kinder auf Meisterung am Beispiel der Präfixverbbildung (trennbar und untrennbare Präfixe) untersucht. Anhand der Fehlercodierungen der Daten kann für die Partikelverbbildung und die Präfixverbbildung festgestellt werden, dass dieses Wortbildungsmuster (deverbale Ableitung mit einem trennbaren oder untrennbaren Präfix) bis 4;11 von den Kindern nicht vollständig gemeistert wird, da es bis zum Ende des Untersuchungszeitraums immer wieder zu morphologischen Fehlern in

den Derivaten kommt. Ein Vergleich der HSES- und LSES-Kinder zeigt, dass die LSES-Kinder mehr Fehler in beiden Wortbildungsmustern produzieren. Die Präfixverbbildung, sowohl trennbare als auch untrennbare Präfixe, wird daher bis 4;11 auf keinen Fall im Sinne von Berman (2004) und nicht einmal im begrenzten Sinne von Kelić und Dressler (2019) von den Kindern gemeistert.

5.7 Statistische Auswertung der longitudinalen Spontansprachdaten

5.7.1 Chi-Quadrat-Anpassungstest

Für die statistische Auswertung der Unterschiede zwischen CS und CDS in den einzelnen Subkategorien der Wortklassen im Vergleich zu Simplex-Lexemen wurde für die longitudinalen Daten der Chi-Quadrat-Anpassungstest (*Goodness-of-Fit-Test*) herangezogen, der testet, wie gut eine beobachtete Häufigkeitsverteilung einer erwarteten entspricht.

Für die Berechnungen in Excel wurde jeweils das gesamte Korpus im gesamten Untersuchungszeitraum herangezogen, also alle Monate in den longitudinalen Spontansprachdaten. Dabei wurden die Kategorien der derivierten Substantive, Verben und Adjektive, sowie nominale Komposita und Adjektivkomposita auf Unterschiede untersucht. Für die Auswertung der zur Anwendung kommenden Signifikanzniveaus siehe Kap. 5.7.2, für einen zusammenfassenden Überblick der Signifikanzen in CS vs. CDS siehe Tabelle 10.

5.7.1.1 Simplex- vs. derivierte und zusammengesetzte Substantive

N SIMP vs. N DERIV

Die Auswertung der Unterschiede im Bereich der Simplex- vs. derivierten Substantive (N DERIV), zeigt, dass sich CS in Lemmas (JAN, LEN) und Tokens (JAN, KAT, LEN) signifikant von CDS unterscheidet: Lemmas: JAN: $\chi^2 = 7,310$, $df = 1$, $p = ,007$; LEN: $\chi^2 = 7,962$, $df = 1$, $p = ,005$; Tokens: JAN: $\chi^2 = 10,042$, $df = 1$, $p = ,002$; LEN: $\chi^2 = 32,718$, $df = 1$, $p < ,001$; KAT: $\chi^2 = 11,270$, $df = 1$, $p < ,001$.

N SIMP vs. N KOMP

Ebenso verhält es sich bei einem Vergleich der Anteile von Simplex- vs. zusammengesetzten Substantiven (N KOMP). Hier unterscheidet sich CS

signifikant von CDS in Lemmas und Tokens in allen untersuchten Korpora, mit der einzigen Ausnahme von Kathi in Tokens, die eine ähnliche Verwendung wie ihre Mutter zeigt. Lemmas: JAN: $\chi^2 = 27,852$, $df = 1$, $p < ,001$; LEN: $\chi^2 = 20,681$, $df = 1$, $p < ,001$; KAT: $\chi^2 = 5,913$, $df = 1$, $p = ,015$; Tokens: JAN: $\chi^2 = 10,588$, $df = 1$, $p = ,001$; LEN: $\chi^2 = 99,892$, $df = 1$, $p < ,001$.

N SIMP vs. synth. KOMP

Die Auswertung der Anteile von Simplex-Substantiven vs. synthetischer Komposita (synth. KOMP) ergibt im Vergleich ein heterogenes Bild. Während sich bei allen Korpora, außer bei Kathi, signifikante Unterschiede zwischen CS und CDS in Lemmas ergeben, sind die Unterschiede in Tokens nur bei Lena höchst signifikant. Lemmas: JAN: $\chi^2 = 14,497$, $df = 1$, $p < ,001$; LEN: $\chi^2 = 17,102$, $df = 1$, $p < ,001$; Tokens: LEN: $\chi^2 = 27,364$, $df = 1$, $p < ,001$.

5.7.1.2 Simplex- vs. derivierte Verben

Der Vergleich der Anteile von Simplex- vs. derivierten Verben zeigt hoch signifikante Unterschiede zwischen CS und CDS in allen Korpora in Lemmas: JAN: $\chi^2 = 24,238$, $df = 1$, $p < ,001$; LEN: $\chi^2 = 10,695$, $df = 1$, $p = ,001$; KAT: $\chi^2 = 12,097$, $df = 1$, $p < ,001$; und höchst signifikante Unterschiede in Tokens bei Lena $\chi^2 = 28,213$, $df = 1$, $p < ,001$. Jans und Kathis CS unterscheidet sich von CDS in derivierten Verbtokens nicht signifikant.

5.7.1.3 Simplex- vs. derivierte und zusammengesetzte Adjektive

ADJ SIMP vs. ADJ DERIV

In Lemmas ergibt sich der Vergleich der Anteile von Simplex- vs. derivierten Adjektiven in CS vs. CDS als hoch oder höchst signifikant in Lemmas: JAN: $\chi^2 = 12,785$, $df = 1$, $p < ,001$; LEN: $\chi^2 = 18,193$, $df = 1$, $p < ,001$. Bei Kathi zeigen sich die Unterschiede in Lemmas nicht signifikant. In Tokens hingegen ergeben sich höchst signifikante Unterschiede zu CDS bei LEN: $\chi^2 = 14,069$, $df = 1$, $p < ,001$ und signifikante Unterschiede bei KAT: $\chi^2 = 6,378$, $df = 1$, $p = ,012$. Bei Jan zeigen sich in Tokens keine signifikanten Unterschiede zu CDS.

ADJ SIMP vs. *ADJ KOMP*

In allen drei Korpora ergeben sich keine signifikanten Unterschiede bei einem Vergleich der Anteile von Simplex- vs. zusammengesetzten Adjektiven (*ADJ KOMP*).

5.7.2 Zusammenfassender Überblick der Signifikanzen

Die in den vorhergehenden Kapiteln (vgl. Kap. 5.7.1.1, 5.7.1.2, 5.7.1.3) beschriebenen Ergebnisse der Chi-Quadrat-Anpassungstests werden in Tabelle 10 je Korpus (*JAN*, *KAT*, *LEN*) dargestellt. Darin sind die einzelnen Signifikanzen (für Lemmas und Tokens) zwischen CS und CDS in den Wortklassen (*N*, *V*, *ADJ*) und nach Wortbildungsart (*DERIV*, *KOMP*) aufgelistet und übersichtlich angeordnet.

Signifikanzniveau

Weist der *p*-Wert ein Signifikanzniveau von 0,05 oder darunter auf ($\leq 0,05$), so ist der Wert statistisch signifikant, liegt er darüber, gilt er als nicht signifikant. Werte gleich oder unter 0,01 sind hoch signifikant, Werte gleich oder unter 0,001 sind höchst signifikant (vgl. Tabelle 10).

Tabelle 10. Überblick Signifikanzen der Chi-Quadrat-Anpassungstests in den Korpora *JAN*, *KAT* und *LEN*: Simplex- vs. derivierte/zusammengesetzte Lexeme (Lemmas, Tokens) in CS vs. CDS im gesamten Zeitraum je Wortklasse

CS vs. CDS	N DERIV	N KOMP	synth. KOMP	V DERIV	ADJ DERIV	ADJ KOMP
	vs. N SIMP			vs. V SIMP	vs. ADJ SIMP	
	<i>p</i> -Wert ^a	<i>p</i> -Wert	<i>p</i> -Wert	<i>p</i> -Wert	<i>p</i> -Wert	<i>p</i> -Wert
JAN Lem.	0,006859	0,000000	0,000140	0,000001	0,000349	<i>0,964350</i>
JAN Tok.	0,001530	0,001138	<i>0,312056</i>	<i>0,195330</i>	<i>0,170102</i>	<i>0,121525</i>
KAT Lem.	<i>0,113893</i>	0,015033	<i>0,371743</i>	0,000505	<i>0,058762</i>	<i>0,619680</i>
KAT Tok.	0,000788	<i>0,174724</i>	<i>0,720863</i>	<i>0,383470</i>	0,011553	<i>0,963794</i>
LEN Lem.	0,004775	0,000005	0,000035	0,001074	0,000020	<i>0,467188</i>
LEN Tok.	0,000000	0,000000	0,000000	0,000000	0,000176	<i>0,633423</i>

a. *p*-Werte: **fett** = (hoch, höchst) signifikant, *kursiv* = nicht signifikant

Von den Adjektivkomposita abgesehen, die in allen Korpora, sowohl in Lemmas als auch in Tokens, keine signifikanten Unterschiede zwischen CS und CDS zeigen, sind die Unterschiede in Lemmas in der Mehrheit der untersuchten Fälle signifikant. Auffallend ist, dass sich bei Lena in Lemmas und Tokens hoch oder höchst signifikante Unterschiede zu CDS ergeben, was bedeutet, dass Lena

signifikant weniger derivierte und zusammengesetzte Lexeme in Lemmas und Tokens bildet als ihre Mutter. Bei Kathi hingegen ist feststellbar, dass sich die Unterschiede zumeist nicht signifikant zeigen, was darauf zurückzuführen ist, dass sie über den gesamten Zeitraum berechnet zwar weniger derivierte und zusammengesetzte Lexeme bildet, der Unterschied zu CDS jedoch nicht signifikant ist.

Die Berechnungen, die in Tabelle 10 zusammengefasst sind, ergeben, dass Jan, Kathi und Lena signifikant weniger derivierte und zusammengesetzte Substantivlemmas und -tokens (N DERIV, N KOMP) aufweisen im Vergleich zu ihren Müttern, mit Ausnahme von Kathi, deren Lemmafrequenz in N DERIV und Tokenfrequenz in N KOMP sich nicht signifikant von ihrer Mutter unterscheiden. Bei den synthetischen Komposita (synth. KOMP) zeigt sich, dass die beiden Kinder Jan und Lena signifikant weniger synthetische Komposita als ihre Mütter produzieren, nur Lena verwendet diese auch signifikant weniger in Tokens. Kathi bildet ähnlich viele synthetische Komposita wie ihre Mutter und unterscheidet sich daher nicht signifikant.

Alle drei Kinder weisen signifikant weniger derivierte Verblemmas (V DERIV) auf verglichen mit ihren Müttern, allerdings produziert einzig Lena ebenfalls signifikant weniger derivierte Verbtokens als ihre Mutter (vgl. Tabelle 10).

Die beiden Kinder Jan und Lena verfügen über signifikant weniger derivierte Adjektivlemmas im Vergleich zu ihren Müttern. Kathi und Lena produzieren signifikant weniger derivierte Adjektive (ADJ DERIV) in Tokens als ihre Mütter, während sich die Tokenfrequenz der ADJ DERIV bei Jan nicht von dessen CDS unterscheidet.

5.8 Statistische Auswertung der Paneldaten

Die statistischen Berechnungen erfolgten für die Paneldaten des INPUT-Korpus mit dem Softwarepaket IBM® SPSS® Statistics, Version 27. Die Rohdaten der in Excel codierten Daten wurden nach SPSS exportiert und jede Variable individuell beschriftet, bestimmten Werten und dem korrekten Messniveau zugeordnet. Um die Gruppenunterschiede nach sozioökonomischem Status statistisch festzustellen, wurden unabhängige T-Tests durchgeführt (siehe Kap. 5.8.1), für die Testungen der

Gruppenunterschiede innerhalb einer SES-Gruppe zwischen CS und CDS wurden der Datenstruktur zufolge gepaarte T-Tests errechnet (siehe Kap. 5.8.2). Aufgrund der leicht unterschiedlichen Gruppengrößen wurde mit Mittelwerten gerechnet. Außerdem wurden die Gruppenvergleiche infolge multiplen Testens Benjamini-Hochberg-korrigiert.

5.8.1 Unabhängige T-Tests

Wie bereits bei den longitudinalen Spontansprachdaten werden die Ergebnisse der Spontansprache auch für die Paneldaten in Bezug auf Derivationen und Komposita in den Wortklassen Substantive, Verben und Adjektive statistisch ausgewertet. Für die Berechnungen wurde das gesamte Korpus, also alle vier Erhebungszeitpunkte herangezogen. Mittels unabhängiger T-Tests sollen die Unterschiede zwischen den Kindern (CS vs. CS) und den Hauptbezugspersonen (CDS vs. CDS) in den beiden SES-Gruppen (*high, low*) festgestellt und berechnet werden.

Gruppenvariable Sozioökonomischer Status

Da die Variable des sozioökonomischen Status für diese Arbeit wesentlich ist, um Unterschiede und Ähnlichkeiten im spontansprachlichen Verhalten der Kinder und ihren Hauptbezugspersonen zu erkennen, werden im Folgenden die Variablen der derivierten Substantive, Verben und Adjektive, und überdies der Substantiv- und Adjektivkomposita bei HSES- und LSES-Kindern, sowie bei HSES- und LSES-Hauptbezugspersonen auf Basis der Gruppenvariable des sozioökonomischen Status' gegenüberstellend untersucht.

CS vs. CS

Bei den Mittelwertvergleichen der Häufigkeiten in Lemmas in den beiden Kindergruppen nach sozioökonomischem Status ergeben sich statistisch signifikante Unterschiede zwischen HSES ($N = 15$) und LSES ($N = 14$) bei derivierten Substantiven und Adjektiven, sowie bei nominalen Komposita und synthetischen Komposita (siehe dazu Tabelle 11). Die stärksten Effekte zeigen sich bei derivierten und zusammengesetzten Substantiven (N KOMP $r = 0,61$; N DERIV $r = 0,56$), mittlere Effekte bei synthetischen Komposita ($r = 0,40$) und derivierten Adjektiven ($r = 0,43$). In Tokens sind statistisch signifikante Unterschiede in CS begrenzt auf die Kategorie der nominalen Komposita ($r = 0,57$), siehe Tabelle 12.

Tabelle 11. Mittelwertdifferenzen in CS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche SES | Lemmas

Variablen	LSES CS		HSES CS		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	16,29	4,983	23,07	5,378	-3,515	27	0,00157	sig. ^a	0,56
V DERIV	55,00	22,604	69,60	24,295	-1,672	27	0,10605	n. sig. ^a	–
ADJ DERIV	4,64	3,815	7,67	2,769	-2,455	27	0,02080	sig. ^a	0,43
N KOMP	15,93	9,119	33,93	14,114	-4,047	27	0,00039	sig. ^a	0,61
synth. KOMP	2,71	2,894	5,73	4,183	-2,244	27	0,03321	sig. ^a	0,40
ADJ KOMP	0,43	0,938	0,93	1,033	-1,375	27	0,18054	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)Tabelle 12. Mittelwertdifferenzen in CS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Tokens)

Gruppenvergleiche SES | Tokens

Variablen	LSES CS		HSES CS		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	42,86	27,431	55,20	14,833	-1,522	27	0,13969	n. sig. ^a	–
V DERIV	95,86	48,118	115,87	44,139	-1,168	27	0,25299	n. sig. ^a	–
ADJ DERIV	8,14	8,160	12,53	7,029	-1,556	27	0,13145	n. sig. ^a	–
N KOMP	27,07	18,395	54,07	21,339	-3,637	27	0,00115	sig. ^a	0,57
synth. KOMP	3,79	3,827	9,27	7,695	-2,400	27	0,02353	n. sig. ^a	–
ADJ KOMP	0,57	1,284	1,47	2,386	-1,245	27	0,22398	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Die LSES-Kinder bilden in allen Wortklassen durchschnittlich weniger verschiedene derivierte und zusammengesetzte Lexeme (Lemmas) im Vergleich zu den HSES-Kindern, mit den höchsten, statistisch signifikanten mittleren Differenzen mit starken Effekten bei nominalen Komposita, wo die LSES-Kinder um 18 Lemmas weniger nominale Komposita verwenden im Vergleich zu den HSES-Kindern: $-18,0$ (95 %-KI[-27,13; -8,88], $t(27) = -4,05$, $p < 0,001$, $r = 0,61$) und derivierten Substantiven, mit einer durchschnittlich um 7 Lemmas geringeren Frequenz: $-6,78$ (95 %-KI[-10,74; -2,82], $t(27) = -3,52$, $p < 0,01$, $r = 0,56$). Ebenfalls signifikante Unterschiede, allerdings mit mittleren Effekten zeigen sich in der Bildung derivierter Adjektive $-3,02$ (95 %-KI[-5,55; -0,50], $t(27) = -2,46$, $p = 0,02$, $r = 0,43$) und synthetischer Komposita $-3,02$ (95 %-KI[-5,78; -0,26], $t(27) = -2,24$, $p < 0,03$, $r = 0,40$), wo die HSES-Kinder durchschnittlich 3 Lemmas

mehr produzieren als die LSES-Kinder. Für die mittlere Anzahl derivierter Verben konnten keine statistisch signifikanten Unterschiede zwischen HSES und LSES festgestellt werden: $t(27) = -1,67$; $p = 0,11$, ebenso wenig wie für Adjektivkomposita: $t(27) = -1,38$; $p = 0,18$.

In Bezug auf die Verwendungshäufigkeit (Tokens) zeigt sich einzig die Kategorie der nominalen Komposita als statistisch signifikant unterschiedlich, wobei die LSES-Kinder durchschnittlich 12 N KOMP-Lemmas weniger bilden, als die gleichaltrigen HSES-Kinder: $-12,34$ (95 %-KI[-42,23; -11,76], $t(27) = -3,64$ $p = 0,001$, $r = 0,57$). Wie in Tabelle 12 ersichtlich, gibt es in CS in Abhängigkeit des sozioökonomischen Status' in allen anderen Kategorien keine signifikanten Unterschiede in Tokens, was bedeutet, dass die Kinder sich zwar in der Diversität der Lexeme in derivierten Substantiven und Adjektiven, sowie synthetischen Komposita signifikant unterscheiden (vgl. Tabelle 11), nicht jedoch in deren Verwendungshäufigkeiten.

CDS vs. CDS

Die Vergleiche der Mittelwerte ergeben für die beiden SES-Gruppen der Hauptbezugspersonen statistisch signifikante Unterschiede zwischen HSES ($N = 15$) und LSES ($N = 14$). Bei Lemmas unterscheiden sich die beiden Gruppen in allen Kategorien signifikant, mit Ausnahme von Adjektivkomposita (vgl. dazu Tabelle 13). In Tokens hingegen zeigen sich statistisch signifikante Differenzen bei der Verwendungshäufigkeit nominaler Komposita, synthetischer Komposita und derivierter Adjektive (siehe Tabelle 14). Starke Effekte zeigen sich, wie bereits für CS festgestellt, auch in CDS bei derivierten und zusammengesetzten Substantiven, sowie überdies bei synthetischen Komposita (N KOMP Lem. $r = 0,72$, Tok. $r = 0,70$; N DERIV Lem. $r = 0,55$, synth. KOMP Lem. $r = 0,53$, Tok. $r = 0,60$), mittlere Effekte bei derivierten Adjektiven (Lem. $r = 0,49$, Tok. $r = 0,47$) und derivierten Verben (Lem. $r = 0,42$).

Tabelle 13. Mittelwertdifferenzen in CDS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche SES Lemmas									
Variablen	LSES CDS		HSES CDS		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	26,43	7,743	37,47	9,357	-3,447	27	0,0019	sig. ^a	0,55
V DERIV	117,07	34,755	142,73	22,407	-2,380	27	0,0246	sig. ^a	0,42
ADJ DERIV	10,71	5,553	17,47	6,675	-2,950	27	0,0065	sig. ^a	0,49
N KOMP	28,86	11,272	59,33	21,875	-4,761	21,25	0,0001	sig. ^a	0,72
synth. KOMP	6,71	2,972	11,73	6,123	-2,837	20,55	0,0100	sig. ^a	0,53
ADJ KOMP	1,14	1,292	1,93	1,387	-1,585	27	0,1247	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Tabelle 14. Mittelwertdifferenzen in CDS nach SES-Zugehörigkeit, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Tokens)

Gruppenvergleiche SES Tokens									
Variablen	LSES CDS		HSES CDS		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	64,36	30,017	90,87	35,532	-2,162	27	0,03963	n. sig. ^a	–
V DERIV	252,86	65,957	287,93	65,981	-1,431	27	0,16396	n. sig. ^a	–
ADJ DERIV	17,50	9,517	32,13	17,324	-2,790	27	0,00955	sig. ^a	0,47
N KOMP	48,93	23,103	104,07	33,824	-5,089	27	0,00002	sig. ^a	0,70
synth. KOMP	9,00	4,151	18,00	9,761	-3,268	19,18	0,00401	sig. ^a	0,60
ADJ KOMP	1,43	1,604	2,40	2,197	-1,351	27	0,18779	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Im Vergleich zu den LSES-Hauptbezugspersonen bilden die HSES-Hauptbezugspersonen in allen Kategorien und Wortklassen durchschnittlich mehr unterschiedliche Derivationen und Komposita in Lemmas und weisen ebenfalls eine durchschnittlich höhere Verwendungshäufigkeit in Tokens auf. Dabei zeigen sich statistisch signifikante Unterschiede und starke Effekte in Lemmas und Tokens, wie den Tabellen 13 und 14 zu entnehmen ist, bei nominalen Komposita, wo die LSES-Hauptbezugspersonen durchschnittlich um 30 Lemmas und 55 Tokens weniger N KOMP produzieren im Vergleich zu den HSES-Hauptbezugspersonen: Lem. -30,48 (95 %-KI[-43,78; -17,17], $p < 0,001$, $r = 0,72$), Tok. -55,14 (95 %-KI[-77,37; -32,91], $p < 0,001$, $r = 0,70$), und bei synthetischen Komposita, wo die HSES-Hauptbezugspersonen rund 5 Lem./9 Tok. mehr synth. KOMP aufweisen: Lem. -5,02 (95 %-KI[-8,70; -1,33], $p = 0,01$, $r = 0,53$), Tok.

-9,0 (95 %-KI[-14,76; -3,24], $p < 0,01$, $r = 0,60$). Im Bereich der derivierten Adjektive verwenden die LSES-Hauptbezugspersonen durchschnittlich 7 Lem./15 Tok. weniger als die HSES-Hauptbezugspersonen bei mittleren Effekten: Lem. -6,75 (95 %-KI[-11,45; -2,06], $p < 0,01$, $r = 0,49$), Tok. -14,63 (95 %-KI[-25,40; -3,87], $p < 0,01$, $r = 0,47$). Außerdem weist die LSES-CDS durchschnittlich 11 Lem./27 Tok. weniger derivierte Substantive und 26 Lem./35 Tok. derivierte Verben als die HSES-CDS auf, was sich in beiden Wortklassen jedoch ausschließlich in Lemmas als statistisch signifikant unterschiedlich erweist: N DERIV -11,04 (95 %-KI[-17,61; -4,47]), $p = 0,001$, $r = 0,55$, V DERIV -25,66 (95 %-KI[-47,79; -3,54]), $p = 0,02$, $r = 0,42$.

Es gibt keine statistisch signifikanten Unterschiede zwischen HSES-CDS und LSES-CDS nach der Benjamini-Hochberg-Korrektur bei der mittleren Anzahl einerseits von Adjektivkomposita in Lem. $t(27) = -1,59$; $p = 0,12$ und Tok. $t(27) = -1,35$; $p = 0,19$), andererseits bei derivierten Substantiv- und Verbtokens: N DERIV Tok. $t(27) = -2,16$; $p = 0,04$, V DERIV Tok. $t(27) = -1,43$; $p = 0,16$.

Die zuvor bei der Auswertung der CS-Daten bei den HSES-Kindern festgestellten durchschnittlich höheren Lemmafrequenzen zeigen sich ebenfalls in HSES-CDS. Hauptbezugspersonen mit höherem sozioökonomischen Status weisen im Vergleich zu jenen mit niedrigerem SES signifikant mehr unterschiedliche Lemmas in allen untersuchten Wortklassen und Wortbildungsarten, außer bei Adjektivkomposita auf und verwenden im Speziellen nominale und synthetische Komposita, sowie derivierte Adjektive signifikant häufiger.

5.8.2 Gepaarte T-Tests

Für die Berechnungen der Gruppenunterschiede der Kindersprache und der an sie gerichteten Sprache innerhalb einer SES-Gruppe war es aufgrund der Datenstruktur nötig, gepaarte T-Tests durchzuführen.

Jeder Teilnehmer des INPUT-Projekts stellt einen einzelnen Datensatz in SPSS dar, der sämtliche das Kind betreffende Daten enthält. Dies inkludiert nicht nur dessen individuelle Metadaten, wie Alter, Geschlecht oder SES-Zugehörigkeit, sondern ebenfalls alle sprachlichen Daten, die das Kind selbst oder aber seine Hauptbezugsperson betreffen. In anderen Worten sind alle relevanten Daten in zahlreichen Variablen innerhalb eines Datensatzes erfasst, weshalb eine

Differenzierung zwischen CS und CDS nur mittels eines gepaarten T-Tests möglich ist.

Überdies wurden gepaarte T-Tests dazu verwendet Mittelwerte der Gruppen, die zu zwei verschiedenen Zeitpunkten erhoben wurden, zu differenzieren. Für die Paneldaten bedeutet dies eine statistische Untersuchung der durchschnittlichen Frequenzen in den einzelnen Datenzeitpunkten, um festzustellen, ob es eine Veränderung der mittleren Anzahl derivierter oder zusammengesetzter Lexeme vom ersten zum zweiten Messzeitpunkt gibt.

5.8.2.1 CS vs. CDS

Die Gruppenunterschiede der Kindersprache und ihrer CDS wurden nach SES-Zugehörigkeit ermittelt und in allen Variablen für Lemmas und Tokens berechnet. Die Gruppenvergleiche und Werte der HSES-Dyaden sind in Tabelle 15, und die Ergebnisse der LSES-Gruppe in Tabelle 16 übersichtlich zusammengefasst.

Tabelle 15. Gruppenvergleiche HSES, CS vs. CDS, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas und Tokens), gesamter Zeitraum
Gruppenvergleiche HSES

Variablen	CS		CDS		t	df	p^a	Sig. ^b	r^c
	M	SD	M	SD					
N DER. Lem.	23,07	5,378	37,47	9,357	-5,985	14	0,000	sig. ^a	0,85
N DER. Tok.	55,20	14,833	90,87	35,532	-3,985	14	0,001	sig. ^a	0,73
V DER. Lem.	69,60	24,295	142,73	22,407	-9,524	14	0,000	sig. ^a	0,93
V DER. Tok.	115,87	44,139	287,93	65,981	-9,227	14	0,000	sig. ^a	0,93
ADJ DER. Lem.	7,67	2,769	17,47	6,675	-5,518	14	0,000	sig. ^a	0,83
ADJ DER. Tok.	12,53	7,029	32,13	17,324	-3,977	14	0,001	sig. ^a	0,73
N KOMP Lem.	33,93	14,114	59,33	21,875	-8,336	14	0,000	sig. ^a	0,91
N KOMP Tok.	54,07	21,339	104,07	33,824	-8,776	14	0,000	sig. ^a	0,92
syn. KOMP Lem.	5,73	4,183	11,73	6,123	-4,608	14	0,000	sig. ^a	0,78
syn. KOMP Tok.	9,27	7,695	18,00	9,761	-4,727	14	0,000	sig. ^a	0,78
ADJ KOMP Lem.	0,93	1,033	1,93	1,387	-2,415	14	0,030	sig. ^a	0,54
ADJ KOMP Tok.	1,47	2,386	2,40	2,197	-1,210	14	0,246	n. sig. ^a	–

a. p -Werte 0,000 entsprechen $p < 0,001$

b. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

c. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Tabelle 16. Gruppenvergleiche LSES, CS vs. CDS, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas und Tokens), gesamter Zeitraum

Gruppenvergleiche LSES

Variablen	CS		CDS		t	df	p^a	Sig. ^b	r^c
	M	SD	M	SD					
N DER. Lem.	16,29	4,983	26,43	7,743	-4,745	13	0,000	sig. ^a	0,80
N DER. Tok.	42,86	27,431	64,36	30,017	-2,435	13	0,030	sig. ^a	0,56
V DER. Lem.	55,00	22,604	117,07	34,755	-6,122	13	0,000	sig. ^a	0,86
V DER. Tok.	95,86	48,118	252,86	65,957	-7,624	13	0,000	sig. ^a	0,90
ADJ DER. Lem.	4,64	3,815	10,71	5,553	-4,852	13	0,000	sig. ^a	0,80
ADJ DER. Tok.	8,14	8,160	17,50	9,517	-3,997	13	0,002	sig. ^a	0,74
N KOMP Lem.	15,93	9,119	28,86	11,272	-3,699	13	0,003	sig. ^a	0,72
N KOMP Tok.	27,07	18,395	48,93	23,103	-2,737	13	0,017	sig. ^a	0,60
syn. KOMP Lem.	2,71	2,894	6,71	2,972	-3,778	13	0,002	sig. ^a	0,72
syn. KOMP Tok.	3,79	3,827	9,00	4,151	-3,216	13	0,007	sig. ^a	0,67
ADJ KOMP Lem.	0,43	0,938	1,14	1,292	-2,219	13	0,045	sig. ^a	0,52
ADJ KOMP Tok.	0,57	1,284	1,43	1,604	-2,375	13	0,034	sig. ^a	0,55

a. p -Werte 0,000 entsprechen $p < 0,001$

b. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

c. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Die gepaarten T-Tests ergeben für HSES und LSES statistisch signifikante Unterschiede mit starken Effekten zwischen CS und CDS in allen getesteten Variablen, mit Ausnahme der Adjektivkomposita in Tokens (ADJ KOMP Tok.) in der HSES-Gruppe, die sich als einzige Variable als nicht signifikant unterschiedlich erweist $t(14) = -1,21$, $p = 0,25$ (vgl. Tabelle 15). Unabhängig von ihrer SES-Zugehörigkeit bilden die Kinder insgesamt signifikant weniger derivierte und zusammengesetzte Lexeme im Vergleich zu ihren Hauptbezugspersonen und verwenden diese auch in signifikant geringerem Ausmaß.

Die größten durchschnittlichen Differenzen lassen sich für CS in erster Linie bei derivierten Verben feststellen, wo die HSES-Kinder 73 Lem./172 Tok. und die LSES-Kinder 62 Lem./157 Tok. weniger V DER. im Vergleich zu CDS produzieren: HSES Lem. -73,13 (95 %-KI[-89,60; -56,66]), Tok. -172,07 (95 %-KI[-212,06; -132,07]), LSES Lem. -62,07 (95 %-KI[-83,98; -40,17]), Tok. -157,0 (95 %-KI[-201,49; -112,52]), sowie des Weiteren in nominalen Komposita und derivierten Substantiven: HSES N KOMP Lem. -25,40 (95 %-KI[-31,94; -18,87]), Tok. -50,0 (95 %-KI[-62,22; -37,78]), N DER. Lem. -14,4 (95 %-KI[-19,56; -9,24]), Tok. -35,67 (95 %-KI[-54,86; -16,47]); LSES N KOMP Lem. -12,93

(95 %-KI[-20,48; -5,38]), Tok. -21,86 (95 %-KI[-39,11; -4,61]), N DER. Lem. -10,14 (95 %-KI[-14,76; -5,53]), Tok. -21,5 (95 %-KI[-40,57; -2,43]). Geringere Unterschiede ergeben sich bei derivierten Adjektiven und synthetischen Komposita, die geringsten bei Adjektivkomposita, da diese in CS und CDS nur vereinzelt vorkommen: HSES ADJ DER. Lem. -9,8 (95 %-KI[-13,61; -5,99]), Tok. -19,6 (95 %-KI[-30,17; -9,03]), syn. KOMP Lem. -6,0 (95 %-KI[-8,79; -3,21]), Tok. -8,73 (95 %-KI[-12,70; -4,77]), ADJ KOMP Lem. -1,0 (95 %-KI[-1,89; -0,11]), Tok. -0,93 (95 %-KI[-2,59; 0,72]); LSES ADJ DER. Lem. -6,07 (95 %-KI[-8,78; -3,37]), Tok. -9,36 (95 %-KI[-14,41; -4,30]), syn. KOMP Lem. -4,0 (95 %-KI[-6,29; -1,71]), Tok. -5,21 (95 %-KI[-8,72; -1,71]), ADJ KOMP Lem. -0,71 (95 %-KI[-1,41; -0,02]), Tok. -0,86 (95 %-KI[-1,64; -0,08]).

5.8.2.2 DP 1+2 vs. DP 3+4

Um festzustellen, ob es eine Veränderung in der Verwendung derivierter Lexeme und Komposita bei einem Vergleich eines früheren und eines späteren Datenzeitpunktes gibt, wurden gepaarte T-Tests mit den Daten der HSES- und LSES-Kinder und ihrer Hauptbezugspersonen durchgeführt. Dabei wurden die Daten aus den Erhebungszeitpunkten 1 und 2 (Durchschnittsalter 3;1 und 3;4), sowie aus den Datenzeitpunkten 3 und 4 (Durchschnittsalter 4;4 und 4;8) zusammengefügt, um zu sehen, welche Veränderungen nach einem Jahr Aufnahmepause bei derivierten Substantiven, Verben und Adjektiven, und bei Substantiv- und Adjektivkomposita bei den Kindern und ihrer CDS statistisch signifikant unterschiedlich sind. Die Mittelwertdifferenzen der untersuchten Kategorien zu den beiden Messzeitpunkten, zusammen mit den jeweiligen Signifikanzen und gegebenenfalls der Effektgrößen sind in Tabelle 17 für die HSES-Kinder (HSES-CS), in Tabelle 19 für die HSES-Hauptbezugspersonen (HSES-CDS), in Tabelle 18 für die LSES-CS und in Tabelle 20 für die LSES-CDS dargestellt.

Gepaarter T-Test DP 1+2 vs. DP 3+4 in CS

Tabelle 17. Mittelwertdifferenzen zweier Datenzeitpunkte in HSES-CS, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche HSES CS | Lemmas

Variablen	DP 3+4		DP 1+2		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	14,80	3,707	12,60	4,983	1,256	14	0,2297	n. sig. ^a	–
V DERIV	43,40	17,422	34,93	13,344	2,273	14	0,0393	n. sig. ^a	–
ADJ DERIV	5,93	2,712	2,60	1,724	4,537	14	0,0005	sig. ^a	0,43
N KOMP	18,47	9,628	17,07	8,259	0,514	14	0,6152	n. sig. ^a	–
synth. KOMP	3,27	3,035	2,87	2,588	0,474	14	0,6429	n. sig. ^a	–
ADJ KOMP	0,73	0,884	0,20	0,561	1,948	14	0,0717	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Tabelle 18. Mittelwertdifferenzen zweier Datenzeitpunkte in LSES-CS, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche LSES CS | Lemmas

Variablen	DP 3+4		DP 1+2		t	df	p	Sig. ^a	r^b
	M	SD	M	SD					
N DERIV	10,79	3,286	8,29	4,027	2,654	13	0,0199	n. sig. ^a	–
V DERIV	33,79	17,669	27,93	12,443	1,463	13	0,1671	n. sig. ^a	–
ADJ DERIV	3,43	2,848	1,71	2,400	3,618	13	0,0031	sig. ^a	0,71
N KOMP	8,29	4,565	8,36	6,356	-0,049	13	0,9619	n. sig. ^a	–
synth. KOMP	1,36	2,274	1,36	1,598	0,000	13	1	n. sig. ^a	–
ADJ KOMP	0,36	0,745	0,07	0,267	1,749	13	0,1039	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Die Gruppenvergleiche der beiden Zeiträume DP 1+2 und DP 3+4 in CS ergeben nur für derivierte Adjektive statistisch signifikante Unterschiede vom früheren zum späteren Zeitpunkt in beiden SES-Gruppen. In allen anderen Kategorien zeigen sich keine statistisch signifikanten Unterschiede. Das bedeutet, dass sich ausschließlich bei der mittleren Anzahl derivierter Adjektive signifikante Veränderungen zwischen 3;1 und 4;8 beobachten lassen (was aufgrund des späteren Erwerbs von Adjektiven zu erwarten war, vgl. Tribushinina et al. 2015). Bei den HSES-Kindern entspricht dies einer Zunahme von rund 3 derivierten Adjektiven und einem mittleren Effekt: 3,33 (95 %-KI[1,76; 4,91]), $p < 0,001$, $r = 0,43$. Bei den LSES-

Kindern fällt die Steigerung mit durchschnittlich 2 derivierten ADJ etwas geringer aus mit einem starken Effekt: 1,71 (95 %-KI[0,69; 2,74]), $p < 0,01$, $r = 0,71$.

Die durchschnittlichen Anzahlen derivierter und zusammengesetzter Lexeme nehmen in Lemmas in HSES-CS vom ersten (DP 1+2) zum zweiten Zeitpunkt (DP 3+4) in allen Kategorien zu (vgl. Tabelle 17). Dies trifft mit Ausnahme der nominalen Komposita (N KOMP, synth. KOMP) ebenfalls auf die LSES-CS zu (siehe Tabelle 18). Die mittlere Anzahl synthetischer Komposita bleibt in LSES-CS über den Vergleichszeitraum gleich ($M = 1,36$) und annähernd gleich in nominalen Komposita $t(13) = -0,05$, $p = 0,96$.

Gepaarter T-Test DP 1+2 vs. DP 3+4 in CDS

Tabelle 19. Mittelwertdifferenzen zweier Datenzeitpunkte in HSES-CDS, Signifikanzen (p , 2-seitig) und Effektstärken (r) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche HSES CDS | Lemmas

Variablen	DP 3+4		DP 1+2		t	df	p	Sig. ^a	r ^b
	M	SD	M	SD					
N DERIV	22,00	6,665	24,00	7,644	-0,889	14	0,3893	n. sig. ^a	–
V DERIV	90,40	17,484	90,33	16,312	0,015	14	0,9885	n. sig. ^a	–
ADJ DERIV	12,73	3,973	9,20	5,074	4,026	14	0,0013	sig. ^a	0,73
N KOMP	32,53	12,247	31,13	14,252	0,379	14	0,7107	n. sig. ^a	–
synth. KOMP	6,33	3,478	5,93	4,234	0,404	14	0,6921	n. sig. ^a	–
ADJ KOMP	1,00	1,195	0,93	0,961	0,155	14	0,8792	n. sig. ^a	–

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

b. Effektstärke ($r = 0,1$ schwach, $r = 0,3$ mittel, $r = 0,5$ stark)

Tabelle 20. Mittelwertdifferenzen zweier Datenzeitpunkte in LSES-CDS und Signifikanzen (p , 2-seitig) der Derivationen und Komposita je Wortklasse (Lemmas)

Gruppenvergleiche LSES CDS | Lemmas

Variablen	DP 3+4		DP 1+2		t	df	p	Sig. ^a
	M	SD	M	SD				
N DERIV	16,36	5,611	15,29	5,497	0,778	13	0,4503	n. sig. ^a
V DERIV	72,29	22,818	76,00	23,442	-0,766	13	0,4573	n. sig. ^a
ADJ DERIV	6,14	3,035	6,07	4,028	0,084	13	0,9347	n. sig. ^a
N KOMP	14,86	7,059	16,36	7,153	-0,988	13	0,3412	n. sig. ^a
synth. KOMP	3,07	2,369	3,86	1,956	-0,928	13	0,3701	n. sig. ^a
ADJ KOMP	0,79	1,188	0,36	0,745	1,066	13	0,3059	n. sig. ^a

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

Die mittlere Anzahl zusammengesetzter und derivierter Lexeme nimmt in Lemmas in HSES-CDS vom ersten (DP 1+2) zum zweiten Zeitpunkt (DP 3+4) in allen Kategorien minimal zu, außer in derivierten Substantiven, deren durchschnittliche Anzahl leicht abnimmt (vgl. Tabelle 19). Eine statistisch signifikante Zunahme lässt sich ausschließlich bei derivierten Adjektiven feststellen, wo die HSES-Hauptbezugspersonen durchschnittlich 4 derivierte Adjektivlemmas mehr im Vergleich zu dem früheren Zeitraum bilden: 3,53 (95 %-KI[1,65; 5,42]), $p = 0,001$, $r = 0,73$. Dies geht mit den durchschnittlichen Zunahmen in HSES-CS konform.

Bei den Hauptbezugspersonen der LSES-Gruppe bleiben die mittleren Anzahlen aller Variablen in dem Untersuchungszeitraum mehr oder weniger konstant (+/- 1, vgl. t-Wert in Tabelle 20). Es sind demnach keine signifikanten Unterschiede zwischen den beiden Messzeitpunkten in LSES-CDS feststellbar, weshalb die Effektstärken für diese Gruppe nicht berechnet wurden. Auch die LSES-CS präsentiert ähnliche Varianzen wie ihre CDS, mit dem Unterschied, dass die mittlere Anzahl in derivierten Verben in CDS sinkt, während sie in CS zunimmt.

Für beide Gruppen (HSES- und LSES-CDS) kann aufgrund der Tests festgestellt werden, dass die an die Kinder gerichtete Sprache durch ihre Hauptbezugspersonen in dem untersuchten Zeitraum nicht signifikant an Diversität, im Sinne einer höheren Lemmafrequenz, zunimmt, sondern bereits im ersten Zeitraum (DP 1+2, Durchschnittsalter 3;1 und 3;4) eine Vielzahl an unterschiedlichen Lexemen in allen analysierten Variablen bereitstellt.

6 Ableitungsbasen bei Derivationen

Von welchen Basen sich Substantive, Verben und Adjektive in CS und CDS ableiten, wird im Folgenden nach Wortklasse in den beiden Korpora JAN und INPUT für die Zeiträume 3;0–4;11 untersucht. Es fließen alle derivierten Lexeme in die Analysen mit ein, unabhängig davon, mit welchem Wortbildungsmuster sie gebildet werden. Aufgrund dessen, dass Verben zu einem großen Teil als Ableitungsbasis dienen, wird zwischen einfachen Verben (V) und komplexen Verben, also Präfixverben (PREFV, mit untrennbarem Präfix), Partikelverben (PTLV, mit trennbarem Präfix), Zirkumfixverben (ZIRKV), und suffigierten

Verben (SUFFV) unterschieden, um die jeweiligen prozentualen Anteile sichtbar zu machen.

6.1 Basen abgeleiteter Substantive

Substantive leiten sich hauptsächlich von Verben und zu einem geringeren Anteil von Adjektiven und Substantiven ab, wie in Abbildung 34 (Lemmas) und in Anhang A7.1 Abbildung 64 (Tokens) dargestellt. Dabei werden einfache Verben gegenüber komplexen in CS und CDS in beiden Zeiträumen sowohl in Lemmas als auch in Tokens bevorzugt. Komplexe verbale Basen werden ebenso wie Zahlwörter und exogene Basen nur selten zur Ableitung von Substantiven herangezogen. Suffigierte Verben treten ausschließlich in CDS im letzten Zeitraum vereinzelt als Basen auf.

Die Anzahl nominaler Basen steigt bei den INPUT-Kindern vom ersten zum zweiten Zeitraum an, während sie bei Jan und in CDS abnimmt oder annähernd gleichbleibt. Insgesamt betragen die durchschnittlichen prozentualen Anteile substantivischer Basen im gesamten Zeitraum 10,2 % Lem./7,8 % Tok. in CS und 11,3 %/8,9 % in CDS. Die Kinder verwenden demnach in etwa gleich viele nominale Basen wie ihre Hauptbezugspersonen.

In Lemmas und Tokens verwenden Jan und die LSES-Kinder mehr unterschiedliche Partikel- als Präfixverbbasen. Dies liegt daran, dass Partikelverben vor allem in frühen Jahren des Spracherwerbs wesentlich frequenter als Präfixverben sind und letztere insgesamt im Untersuchungszeitraum bis fast 5 Jahre eine noch untergeordnete Rolle in CS einnehmen. Überdies sind Partikelverben positionell und prosodisch salienter und gewöhnlich auch morphosemantisch transparenter und können so von den Kindern schneller verarbeitet werden. Die HSES-Kinder verwenden im Vergleich zu den LSES-Kindern in etwa gleich viele Präfix- und Partikelverben als Basis, deren Lemma- und Tokenfrequenzen jedoch sowohl in CS als auch in CDS vom früheren (DP 1+2) zum späteren Zeitpunkt (DP 3+4) jeweils zunehmen. Dieser Trend zeigt sich ebenfalls bei Jans CS und CDS in Bezug auf Partikelverbbasen in Lemmas, ist hingegen bis 4;11 bei den LSES-Kindern und ihren Hauptbezugspersonen weder für Präfix- noch für Partikelverbbasen erkennbar.

Die in Anhang A7.1 in Abbildung 64 dargestellten hohen Tokenfrequenzen bei adjektivischen Basen sind den verschiedenen Farbbezeichnungen geschuldet, die die Kinder in diesem Alter häufig wiederholen, wie (die) Blau, (die) Gelb (JAN 3;0), neben vereinzelt Bildungen von Nomina Qualitatis: Näh-e ← nah, Wahrheit ← wahr (D31-AND 4;6), Güt-e ← gut (D25-PHG 4;4), Höh-e ← hoch (D08-FAJ 4;4) und Sicher-heit ← sicher (D27-JOZ 3;0).

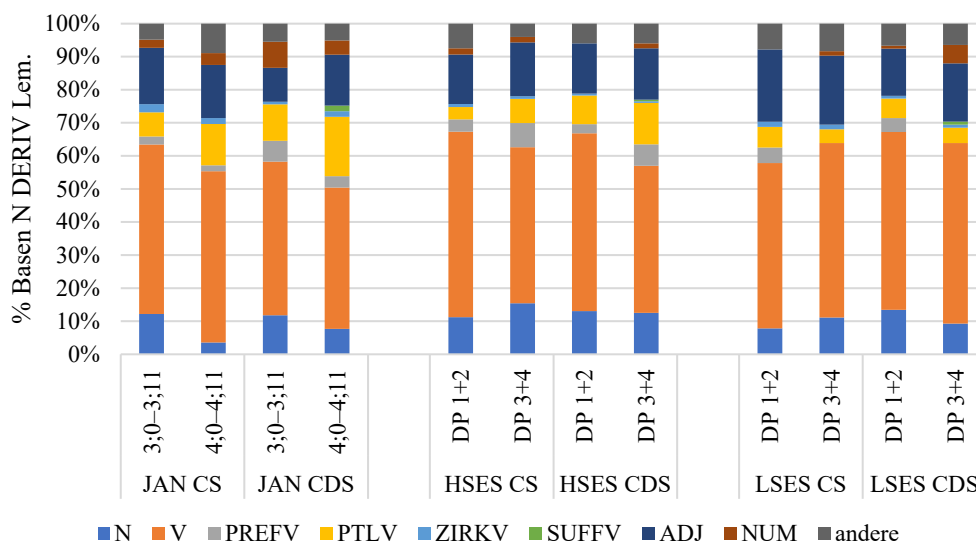


Abbildung 34. Ableitungsbasen derivierter Substantive (Lemmas) im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

Der Großteil derivierter Substantive zählt zu den wortartverändernden Derivationen, da sie sich hauptsächlich von Verben und zu einem geringeren Teil von Adjektiven ableiten. Insgesamt sind in etwa 10 % aller nominalen Derivationen in CS und CDS wortartbewahrend.

6.2 Basen derivierter Verben

Verben leiten sich für gewöhnlich deverbale, in der überwiegenden Mehrheit von einfachen Verben und äußerst selten von komplexen Verben ab. In Summe sind über 92,4 % in CS und über 91,3 % in CDS aller derivierten Verblemmas deverbale abgeleitet. Alle anderen Basen (Substantive, Adjektive, exogene Basen) verteilen sich auf die restlichen Anteile, wobei nominale Basen den größten Teil mit durchschnittlich 6,6 % in CS und CDS in Lemmas im gesamten Zeitraum einnehmen, wie in Abbildung 35 dargestellt.

Bei den deverbalen Ableitungen von einfachen Basen handelt es sich vorwiegend um Partikelverbbildungen, wie *drauf#bau-en* (D21-PAS 3;6) und *weg#geb-en* (D07-NIG 4;2), in wesentlich geringerem Ausmaß um untrennbare Präfigierungen: *be#folg-en* (D29-SAK 4;9), *er#kenn-en* (D28-MOH 3;3); Suffigierungen: *streich-el-n* (D23-SOK 3;3); und implizite Ableitungen: *leg-en* (D15-LUN 3;3) und *setz-en* (D27-JOZ 4;3). Verbderivationen von komplexen verbalen Basen (Präfix- oder Suffixverben) sind ausschließlich Partikelverbbildungen, wie *hin#ge#hör-en* (D13-JOP 4;8) und *aus#prob-ier-en* (D21-PAS 3;6).

Nominale Basen nehmen im Lauf der Zeit in ihrer Verwendungshäufigkeit (Tokenfrequenz) in allen Korpora vom ersten zum zweiten Zeitraum zu (vgl. dazu in Anhang A7.2 Abbildung 65), wobei die verschiedenen denominal abgeleiteten Verblemmas annähernd gleichbleiben oder ebenfalls leicht an Variation zunehmen (siehe Abbildung 35). Denominale verbale Ableitungen treten mit und ohne Umlaut auf, wie in *brösel-n* (CDS JAN 1;9), *fürcht-en* (JAN 2;11), *blätt-er-n* (CDS JAN 1;11) und *fotograf-ier-en* (D19-VAS 3;1).

Deadjektivische Verbderivationen treten vereinzelt in CS und CDS auf, wie in den Bildungen *leer-en* (D28-MOH 3;0), *wärm-en* (CDS D20-LUD 3;6), *blöd-el-n* (CDS D05-FLB 3;3) und *stolz-ier-en* (CDS D29-SAK 3;3).

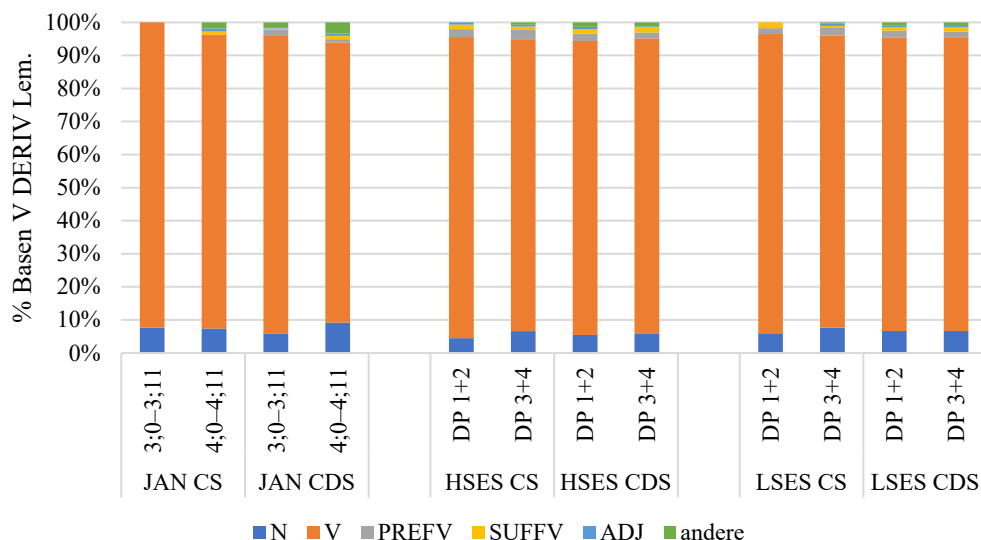


Abbildung 35. Ableitungsbasen derivierter Verben (Lemmas) im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

Derivierte Verben werden in der überwiegenden Mehrheit auf Basis anderer, einfacher oder komplexer, Verben gebildet und zählen daher zu den wortartbewahrenden Derivationen. Wortartverändernde verbale Ableitungen treten lediglich in rund 8 % Lem./9 % Tok. in CS und 9 % Lem./11 % Tok. in CDS aller verbalen Derivationen auf.

6.3 Basen derivierter Adjektive

Die prozentuale Verteilung der Ableitungsbasen zur Bildung von Adjektiven sind für Lemmas in Abbildung 36 und für Tokens in Anhang A7.3 Abbildung 66 in CS und CDS der Korpora JAN und INPUT übersichtlich dargestellt. Dabei zeigt sich, dass Adjektive in CS und CDS hauptsächlich denominal abgeleitet werden, wie z. B. glück-lich (D06-LAB 4;5), gold-en (D03-NIP 4;5), hungr-ig (CDS D29-SAK 4;9), künst-ler-isch (CDS D04-SOS 3;5). Nominale Basen zeigen sich für durchschnittlich 42 % Lem./41 % Tok. in CS und 44 % Lem./38 % Tok. in CDS verantwortlich.

Verbale Basen sind in CDS die zweithäufigste Wortklasse zur Ableitung von Adjektiven in Lemmas und Tokens. In CS hingegen werden Adjektive als Basen vor verbalen Basen in Lemmas bevorzugt. Vorwiegend handelt es sich bei deverbalen Bildungen um Partizipien in ihrer Verwendung als Adjektive, wie in brenn-end (PPräs, CDS D26-JUB 4;5), glitzer-nd (PPräs, D07-NIG 4;5), ge-bog-en (PP, CDS D02-SIJ 4;11) und ge-spann-t (PP, CDS D22-LAS 3;3). Weitere deverbal abgeleitete Adjektive werden zumeist mit den Suffixen -lich und -ig gebildet, wie in empfind-lich (CDS D18-ALM 4;7), beweg-lich (CDS D15-LUN 4;7) und stink-ig (CDS D04-SOS 4;5).

Besonders in CS lassen sich hohe Lemmafrequenzen im Bereich adjektivischer Basen feststellen. Derivierte Adjektive, die sich von Adjektiven ableiten, werden entweder suffigiert oder präfigiert, wie in schwier-ig (mit Ablaut, D16-LEP 4;6), nack-ig (D15-LUN 4;7), lang-sam (D22-LAS 2;11), aller-schlimm-ste (D06-LAB 4;5), ur-schnell (D18-ALM 3;0) und un-fair (D05-FLB 4;4). Die Kinder verwenden viele unterschiedliche deadjektivische Lexeme, die sie jedoch weniger oft als denumerale Bildungen wiederholen. Die Tokenfrequenz numeraler Basen ist mit durchschnittlich 28 % nach nominalen Basen die zweithäufigste in CS. Diese hohen

Frequenzen sind auf suffigierte Bildungen von Numeralia mit und ohne (irreguläre) Stammmodifikation zurückführbar: vier-t (D03-NIP 4;8), fünf-t (D31-AND 3;3), zehn-t (JAN 2;0), dri-tt (D23-SOK 4;3), er-st (D21-PAS 3;6).

Bildungen auf Basis von Adverbien, wie in derart-ig (ADV, CDS JAN 2;10) kommen nur vereinzelt in CS und CDS vor.

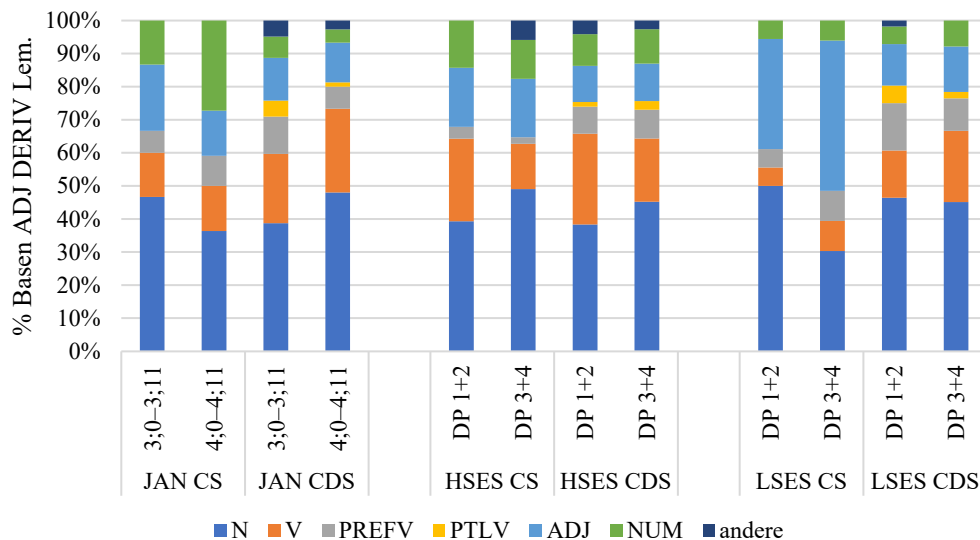


Abbildung 36. Ableitungsbasen derivierter Adjektive (Lemmas) im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

Derivierte Adjektive können sowohl wortartbewahrend als auch wortartverändernd sein, wobei der Großteil abgeleiteter Adjektive zu zweiter Gruppe zählt, da sie sich in erster Linie von Substantiven und Verben ableiten, und zu einem geringeren Anteil von Adjektiven. Durchschnittlich 12 % Lem./Tok. aller adjektivischen Derivationen in CDS und 25 % Lem./17 % Tok. in CS in Lemmas sind wortartbewahrend.

Für eine cross-linguistische Übersicht zum Erwerb von Adjektiven siehe Tribushinina et al. (2015), zum Erwerb im Deutschen siehe Korecky-Kröll und Dressler (2015).

7 Komplexität in der Wortbildung

Die morphotaktische Komplexität betreffend, folgt die Verfasserin der Annahme des Modells der Komplexität (siehe Zurek 1990; Miestamo et al. 2008), das besagt, dass die Anzahl der Wortbildungsprozesse, die zur Bildung eines Wortes nötig sind, den Grad der Komplexität bestimmen. Wird zur Bildung eines Lexems ein einziges Wortbildungsverfahren, wie etwa Affigierung, Ab-/Umlaut, Stammmodifikation oder Komposition appliziert, wie z. B. in *fahr-* → *Fahr-er*, *Kamm* → *kämm-*, *Haus* und *Tier* → *Haus+tier*, dann ist dies einfacher zu erfassen und zu verarbeiten, als wenn mehrere Wortbildungsprozesse eingesetzt werden, wie z. B. in *jag-* → *Jäger-er*, *Busch* → *Ge-büsch*. Die morphotaktische Komplexität spielt für den Spracherwerb eine wesentliche Rolle, da die zunehmende Komplexität grammatischer Strukturen zu längeren Verarbeitungszeiten führen. Eine besondere Art der Komplexität weisen daher synthetische Komposita auf, da diese derivationale mit kompositionellen Wortbildungsprozessen vereinen, z. B. *Auto+fahr-er*, *Be-dien-ungs+an#leit-ung*.

7.1 Komplexitätsgrade der Derivate und Komposita

Die morphologische Komplexität eines Derivats oder eines Kompositums wird von der Anzahl der benötigten Wortbildungsprozesse bestimmt, die zu dessen Bildung nötig sind, und steht ebenso in direkter Relation zur Komplexität seiner Basis, wie bereits unter Kap. 4.4.1 (S. 78) aufgegriffen. Steigt die Komplexität der Basis, etwa wenn ein Präfix- oder Partikelverb als Ableitungsbasis dient oder ein mehrgliedriges oder synthetisches Kompositum gebildet wird, so erhöht sich der Komplexitätsgrad des Lexems entsprechend. Der morphologische Komplexitätsgrad ist nicht an der Anzahl der Morpheme ablesbar, sondern allein an der Anzahl der Wortbildungselemente, die sich in einem Lexem feststellen lassen. Für eine gute Vergleichbarkeit der Ergebnisse orientieren sich die Grade der morphologischen Komplexität an der Untersuchung von Mattes (2018: 123). Alle Lexeme in allen Korpora wurden einem Komplexitätsgrad zugeordnet (0–4, einfach–komplex, vgl. Übersicht der Komplexitätsgrade in Tabelle 5, S. 78). Einfache Lexeme, die Komplexitätsgrad 0 (Null) aufweisen, sind Simplizia oder Lexeme, die rein formal ein Affix aufweisen, aber über keine autonome Basis

verfügen, z. B. gewinn-, vergess-. Lexeme, die einen Wortbildungsprozess (Komposition, Affigierung, Konversion, Stammmodifikation, Stammvokalveränderung) beinhalten, werden Komplexitätsgrad 1 zugeordnet, unabhängig davon, ob sie monomorphemisch sind (z. B. leg-, Fall, tank-) und somit als komplex gewertet. In gleicher Weise werden alle lexikalisierten und morphosemantisch opaken Lexeme (z. B. ver-steh-en, be-such-en, Zeit-ung) kategorisiert. Dass die Opazität in deutschen Lexemen keinen nachteiligen Einfluss auf die Verarbeitung derselben hat, wurde im Rahmen experimenteller Studien anhand komplexer Verben gezeigt (vgl. Smolka et al. 2014; Smolka et al. 2019). Alle komplexen Bildungen ab Komplexitätsgrad 4 wurden der letzten Gruppe zugeordnet, um wie bereits erwähnt eine gute Vergleichbarkeit der Analysen zu gewährleisten.

7.2 Morphologische Komplexität im Erwerbsverlauf

In Bezug auf die morphologische Komplexität von Nomina Agentis und Instrumenti in der Kindersprache konnten Sommer-Lolei und Dressler (akzeptiert: Kap. 5.5) anhand von Bildungen mittels -er, -e und Konversion zeigen, dass komplexe Ableitungen bereits sehr früh in CS vorkommen und mehr als die Hälfte aller Lemmas in diesen Kategorien ausmachen. Einfache Bildungen mit einem einzelnen Wortbildungsmuster (ein Suffix, Konversion, Stammmodifikation oder Änderung des Stammvokals), treten der Untersuchung zufolge chronologisch etwas früher auf als komplexe Formen, die mehr als einen Wortbildungsprozess beinhalten und sind ein wichtiger Ausgangspunkt für die Entwicklung komplexerer Derivationen. Ein konstanter Anstieg konnte in nur einem der drei untersuchten Kinder (KAT) festgestellt werden.

7.2.1 Anteile der Wortbildungsarten je Komplexitätsgrad

Die Komplexitätsgrade zusammengesetzter und abgeleiteter Substantive, Verben und Adjektive in CS und CDS wurden in Relation zur jeweiligen Grundgesamtheit aller Substantiv-, Verb- und Adjektivlemmas und -tokens berechnet, um die prozentuale Verteilung in den Korpora darzustellen. Die morphologischen Komplexitätsgrade wurden anhand der in Tabelle 5 aufgelisteten Grade bestimmt.

Ein Gesamtüberblick über die durchschnittliche prozentuale Verteilung derivierter Substantive, Verben und Adjektive, sowie nominaler Komposita nach

Komplexitätsgraden in CS und CDS in allen Korpora wird in Tabelle 21 gegeben. Aufgrund der limitierten Anzahl zusammengesetzter Adjektive wurde auf die tabellarische Darstellung verzichtet. Die Zeitspanne wurde rund um 3 Jahre gewählt, um alle Kinder abbilden zu können.

Tabelle 21. Mittelwerte aller Korpora (in Prozent) der Komplexitätsgrade nominaler Komposita, derivierter Substantive, Verben und Adjektive in CS und CDS (Altersspanne 2;9–3;3)

		Grad 0	Grad 1	Grad 2	Grad 3	Grad 4
		%	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N KOMP	CS	–	14,50/9,70	5,64/3,28	1,91/1,13	0,67/0,39
	CDS	–	18,28/11,92	6,17/3,23	2,87/1,53	0,91/0,51
N DERIV	CS	21,16/18,34	14,81/12,00	3,37/4,40	0,64/0,33	–
	CDS	17,86/19,38	14,72/12,42	5,02/4,34	1,17/0,91	0,03/0,01
V DERIV	CS	2,38/3,69	48,02/19,49	9,30/2,87	0,64/0,22	–
	CDS	2,20/2,66	53,19/21,04	13,85/4,48	0,97/0,23	–
ADJ DERIV	CS	4,14/6,69	13,07/8,30	5,46/2,71	0,65/0,17	0,65/0,17
	CDS	5,46/5,77	18,06/10,26	9,22/4,22	2,23/0,64	1,28/0,61

Derivierte Verben, die mittels eines Wortbildungsprozesses (Komplexitätsgrad 1) gebildet werden, machen in CS und CDS durchschnittlich rund 50 % aller Verben aus (vgl. Tabelle 21) und treten damit mit Abstand am häufigsten in Tokens (CS 19,5 %, CDS 21 %) und Lemmas auf (CS 48 %, CDS 53,2 %). Die Mehrheit aller Bildungen sind dabei dem Wortbildungsmuster der Partikelverben zuzuschreiben (vgl. Kap. 7.2.4.2). Auch Bildungen mit zwei Wortbildungsprozessen (Komplexitätsgrad 2) sind in abgeleiteten Verben häufiger als in zusammengesetzten oder derivierten Substantiven und Adjektiven. Dies zeigt, dass die Verbderivation im Bereich der frühen Sprachentwicklung wesentlich ist und von größerer Bedeutung als Substantivderivation.

In CDS sind nominale Komposita (N KOMP) mit 18,3 % in Lemmas am zweithäufigsten vertreten, gefolgt von derivierten Adjektiven (ADJ DERIV) mit 18,1 %. In CS weist die Gruppe der derivierten Substantive (N DERIV) mit 14,8 % eine ähnliche Diversität wie N KOMP mit 14,5 % auf. Bildungen, die den Grad 3 oder 4 aufweisen, liegen sowohl in CS als auch in CDS in diesem Zeitraum weit hinter den einfacheren Bildungen (Grade 1 und 2) zurück. Pseudo-derivierte Lexeme (Grad 0) finden sich hauptsächlich in derivierten Substantiven und

Adjektiven, gleichermaßen in CS und CDS. Es gibt keine pseudo-zusammengesetzten Lexeme.

7.2.2 Anteile der Wortbildungsmuster je Komplexitätsgrad

In CS und CDS zeigt sich bei den beiden analysierten Korpora JAN und INPUT eine Vielfalt an Wortbildungsmustern zur Bildung von Substantiven, Verben und Adjektiven insbesondere in Grad 1 und 2, wohingegen mit steigender Komplexität eine Abnahme der verschiedenen Wortbildungsmuster in den Wortklassen zu beobachten ist. Im Folgenden wird die Verteilung der verschiedenen Wortbildungsmuster bei Substantiven, Verben und Adjektiven bei den beiden Korpora in zwei Zeitintervallen untersucht.

7.2.2.1 Anteile nominaler Wortbildungsmuster je Komplexitätsgrad

Abbildungen 37 und 38 stellen die jeweiligen Anteile sämtlicher nominaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS und CDS der Korpora JAN und INPUT nach zwei Zeitabschnitten getrennt dar, und ist berechnet an der Gesamtheit derivierter und zusammengesetzter Substantive. Für eine Übersicht in Tokens siehe Anhang A8.1 Abbildungen 67 und 68.

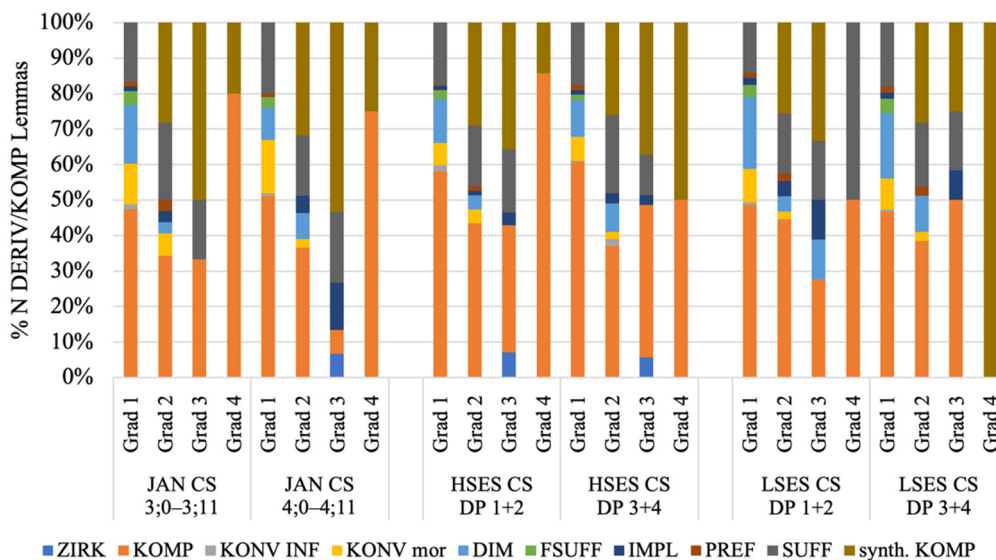


Abbildung 37. Prozentuale Anteile nominaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

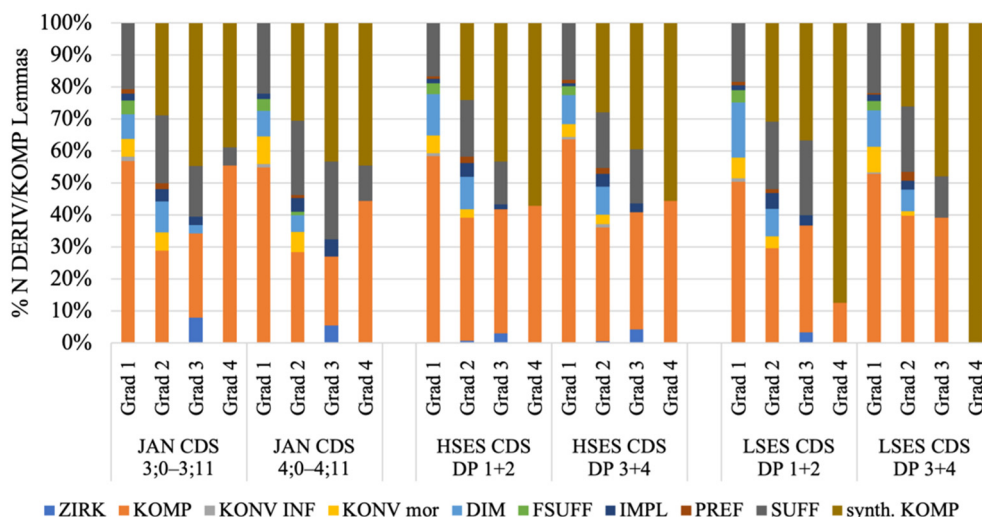


Abbildung 38. Prozentuale Anteile nominaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

Wortbildungen mit geringerer Komplexität (Grad 1 und 2) weisen eine höhere Anzahl an verschiedenen Wortbildungsmustern in CS und CDS in allen Korpora auf. Bildungen mit dem Komplexitätsgrad 4 sind beschränkt auf komplexe Substantivkomposita mit einer Ableitung im Erstglied, Diminutivkomposita, synthetische Komposita und zu einem wesentlich geringeren Anteil auf Suffigierungen. Konversionen (KONV mor, KONV INF), Diminutivsuffixe (DIM), Fremdsuffixe (FSUFF) und Präfigierungen (PREF) treten in den Graden 1 und 2 auf und nicht darüber, da komplexere Bildungen zwar meist eines der genannten Wortbildungsmuster beinhalten, es sich dabei jedoch nicht um den relevanten Wortbildungsprozess handelt (siehe Erläuterung in Beispiel (1) und Kap. 4.4.4). Zirkumfigierte Substantive (Ge-e, Ge-t) mit Komplexitätsgrad 2, wie in Ge-red-e, Ge-heul-e oder Ge-flech-t kommen einzig in CDS der HSES-Kinder vor. Werden sie im Zuge der Zirkumfigierung zusätzlich umgelautet (Grad 3), wie in Berg → Ge-birg-e, treten sie marginal in CDS und CS auf, mit Ausnahme der LSES-Kinder und ihrer CDS (DP 3+4), die über keine derartigen Bildungen verfügen.

Bei Jan, der aufgrund monatlicher Aufnahmen das umfangreichste Korpus darstellt, zeigt sich in Komplexitätsgrad 3 eine Erweiterung um zwei Wortbildungsmuster (ZIRK, IMPL) im Vergleich zu seinem früheren Zeitraum (vgl. Abbildung 37).

7.2.2.2 Anteile der Wortbildungsmuster bei Verben je Komplexitätsgrad

Abbildungen 39 und 40 zeigen die jeweiligen prozentualen Anteile verbaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS und CDS im JAN- und INPUT-Korpus, getrennt nach zwei Zeitintervallen, berechnet in Relation zur Gesamtheit derivierter Verben. Für die Verteilung in Tokens siehe Anhang A8.2 Abbildungen 69 und 70.

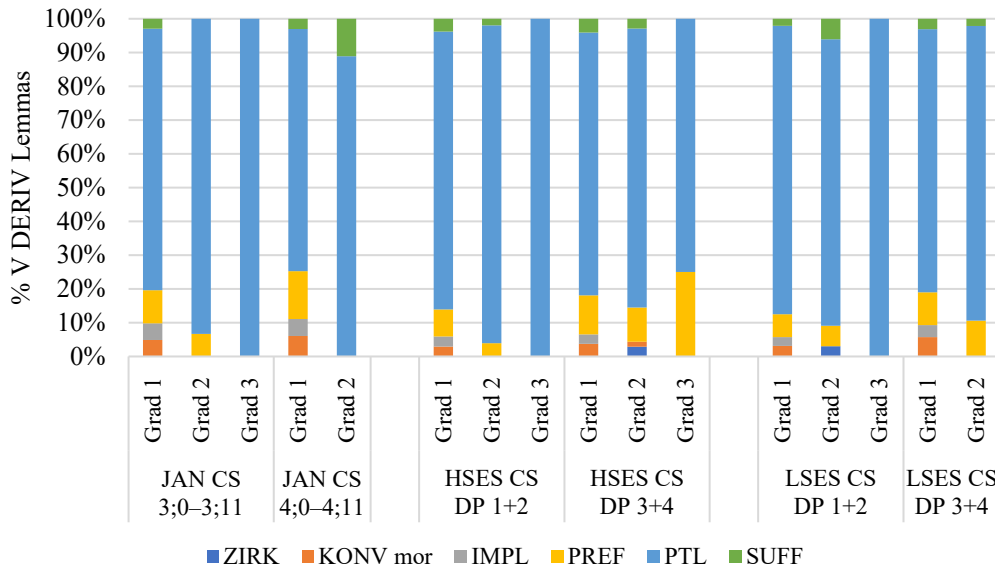


Abbildung 39. Prozentuale Verteilung verbaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

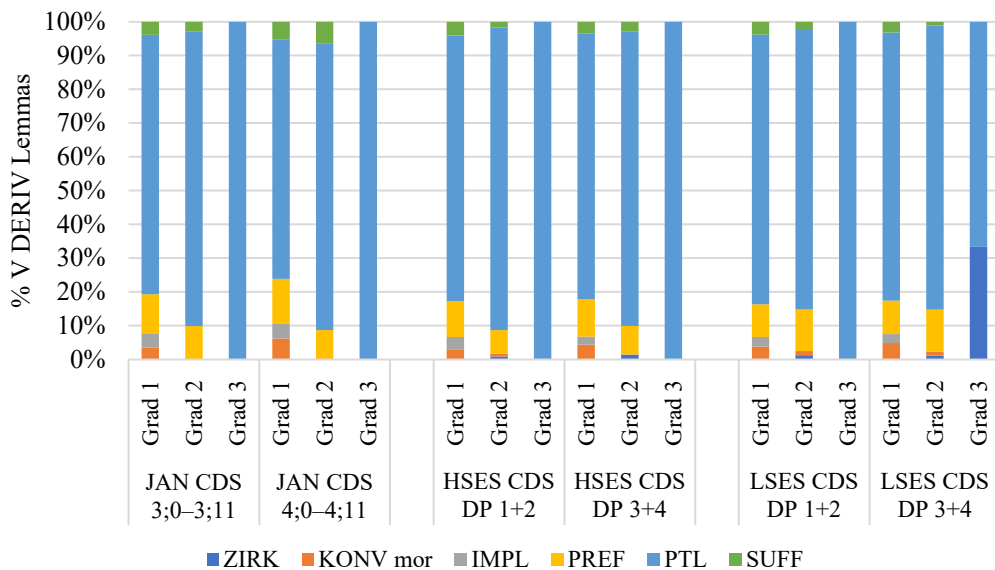


Abbildung 40. Prozentuale Verteilung verbaler Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

Zur Ableitung von Verben ist die Partikelverbbildung in allen Komplexitätsgraden mit Abstand das in CS und CDS dominanteste Wortbildungsmuster, da die Partikel positionell und prosodisch salient ist. Kein anderes verfügt über eine derartige Diversität und Frequenz. Dennoch zeigt sich im Bereich der Partikelverben, dass sich sowohl ihre Lemmafrequenz, als auch ihre Verwendungshäufigkeit (Tokens) im direkten Vergleich der Zeiträume in CS und CDS in allen Komplexitätsgraden zu Gunsten anderer, bzw. in CS neu auftretender Wortbildungsmuster verringert (vgl. Abbildungen 39 und 40, sowie in Anhang A8.2 Abbildungen 69 und 70).

Wie bereits im Bereich der Substantive festgestellt (vgl. Kap. 7.2.2.1), nimmt die Mustervielfalt mit zunehmender Komplexität der Derivate ab. Es gibt keine Bildungen mit dem Komplexitätsgrad 4. Bei Verben mit dem Komplexitätsgrad 3, handelt es sich größtenteils um Bildungen mit den Doppelpartikeln hinein#, herein#, hinüber#, herüber#, herum# in Kombination mit stell- und leg-, sowie um die nähere Bestimmung der Art des Blätterns wie in um#blätt-er-n, weiter#blätt-er-n und weg#blätt-er-n; bei dem zirkumfigierten Verb im Zeitraum DP 3+4 in LSES-CDS um be-schäft-ig-en.

Morphologische Konversionen und implizite Ableitungen sind in den frühen Daten nur in geringem Ausmaß vertreten und kommen hauptsächlich mit Komplexitätsgrad 1 vor. Komplexitätsgrad 2 findet sich vereinzelt bei den HSES-Kindern und den Hauptbezugspersonen aller INPUT-Kinder.

Präfigierte Verben treten, mit einer Ausnahme in HSES-CS in Komplexitätsgraden 1 und 2 auf und nehmen in den untersuchten Zeiträumen an Variation in CS und CDS in beiden Graden zu. Hiervon ausgenommen sind die prozentualen Anteile in Jans CS und CDS mit Grad 2, die an Variation abnehmen. In Bezug auf die Verwendungshäufigkeit einfacher präfigierter Verben (Grad 1) in CS und CDS lässt sich eine tendenzielle Verringerung vom ersten zum zweiten Zeitpunkt feststellen, außer bei den HSES-Kindern, deren Tokenfrequenzen in allen Komplexitätsgraden zunehmen. Die Verwendung von Präfixverben (Grad 2, Tokens) steigt bei den INPUT-Kindern und ihren Hauptbezugspersonen an, während sie bei Jan und seiner Mutter abnimmt. Diese Abnahme kann in Jans CS der zunehmenden Variation und Frequenz suffigierter Verben geschuldet sein (siehe unten). In CDS nimmt die Zahl der verschiedenen präfigierten Verben im Komplexitätsgrad 1 zu, während sich ihre

Frequenz in Tokens im gleichen Zeitraum verringert, was zeigt, dass die verschiedenen Bildungen weniger oft wiederholt werden.

Suffigierte Verben erreichen in den Daten einen maximalen Komplexitätsgrad von 2 und nehmen in Grad 1 bei allen Kindern, und in Grad 2 bei allen HSES-Kindern vom ersten zum zweiten Datenpunkt in CS an Vielfalt und Verwendungshäufigkeit zu. Die LSES-Kinder zeigen eine höhere Diversität und Frequenz an komplexen suffigierten Verben (Grad 2) zum ersten Datenzeitpunkt im Vergleich zum zweiten, was sich auch in ihrer CDS zeigt. In CDS zeigt sich insgesamt ein heterogenes Bild, denn in Jans Input steigen die prozentualen Anteile suffigierter Verben in Lemmas in beiden Komplexitätsgraden an, während in der LSES-CDS das Gegenteil der Fall ist. In HSES-CDS sinkt die Variation der einfach derivierten Verben und nimmt in Grad 2 zu. Die Verwendungshäufigkeit in Tokens nimmt ausschließlich in Jans CDS in Grad 2 und der CDS der LSES-Kinder in Grad 1 zu.

7.2.2.3 Anteile der Wortbildungsmuster bei Adjektiven je Komplexitätsgrad

Die prozentualen Anteile der Wortbildungsmuster bei Adjektiven je Komplexitätsgrad in CS und CDS in Lemmas für die beiden Korpora JAN und INPUT sind in den Abbildungen 41 und 42 in zwei Zeitabschnitten dargestellt. Die Anteile sind berechnet an der Gesamtheit derivierter und zusammengesetzter Adjektive. Für die Darstellungen in Tokens siehe Anhang A8.3 Abbildungen 71 und 72.

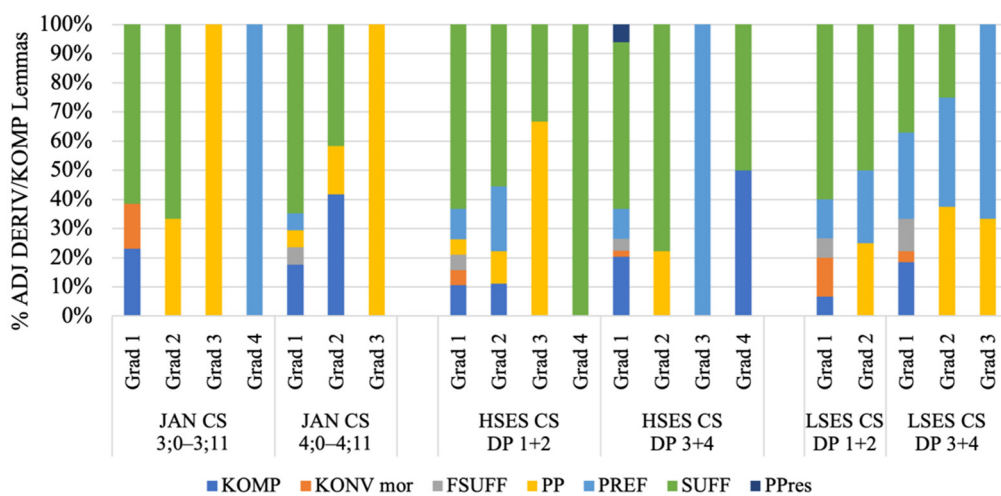


Abbildung 41. Prozentuale Verteilung adjektivischer Wortbildungsmuster in Lemmas je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

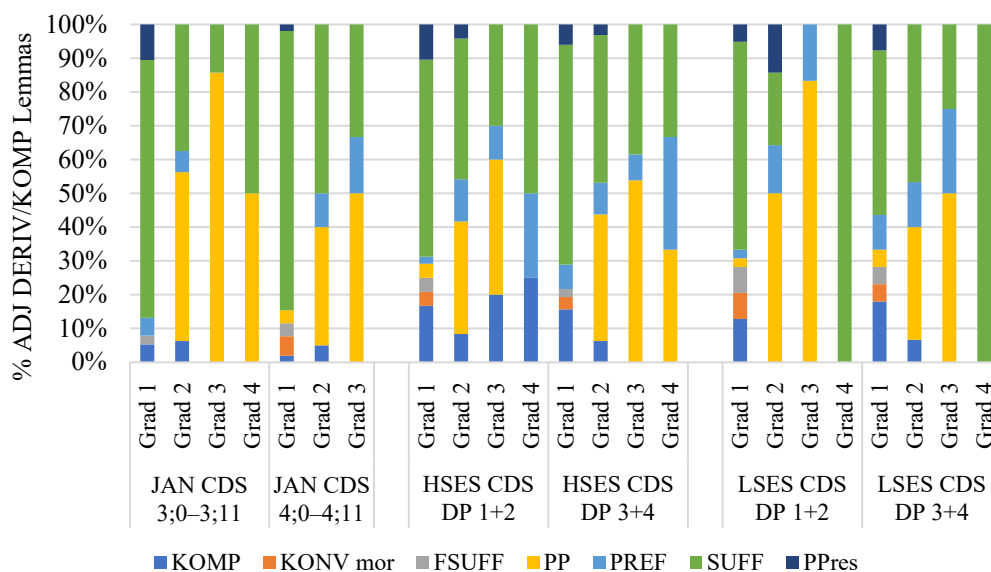


Abbildung 42. Prozentuale Verteilung adjektivischer Wortbildungsmuster in Lemmas je Komplexitätsgrad in CDS im JAN-Korpus (3;0-3;11) und INPUT-Korpus (DP 1+2, 3+4)

Die Variation der Wortbildungsmuster erhöht sich, wie zu erwarten, mit Abnahme des Komplexitätsgrades, wie bereits im Rahmen der Komplexität abgeleiteter und zusammengesetzter Substantive und derivierter Verben festgestellt (vgl. Kap. 7.2.2.1, 7.2.2.2). Es ist nicht überraschend, dass die meisten Adjektive durch Suffigierung gebildet werden.

Bei den Adjektivbildungen mit Komplexitätsgrad 2 handelt es sich deshalb auch hauptsächlich um suffigierte Adjektive, und überdies um Perfektpartizipien in ihrer Verwendung als Adjektive, präfigierte Adjektive und bei HSES auch um Adjektivkomposita. Letztere beschränken sich mit Ausnahme der HSES-CS (DP 3+4) und -CDS (DP 1+2) auf die Komplexitätsgrade 1 und 2. Morphologische Konversionen und Fremdsuffigierungen in CS und CDS sind ausnahmslos einfache Ableitungen mit Grad 1. Präsenspartizipien zeigen eine geringe Vielfalt und Tokenfrequenz in CDS in den Graden 1 und 2, in HSES-CS in Grad 1.

Die Vielfalt suffigierter Adjektive steigt in Jans CDS und HSES-CDS in den Komplexitätsgraden 1-3, in LSES-CDS in den Graden 2-3. Das bedeutet, dass in den zweiten Zeitabschnitten den Kindern mehr verschiedene Lexeme mit allen Komplexitätsgraden zur Verfügung gestellt werden. Die Verwendungshäufigkeit erweist sich im Vergleich dazu als wenig konstant, da die Tokenfrequenz in manchen Graden steigt, in anderen hingegen abfällt, wie in Jans CDS Grad 1, 4,

LSES-CDS Grad 1, 2 und HSES-CDS Grad 3, 4. Die verschiedenen Typen werden demnach nicht mehr so häufig wie im ersten Zeitraum wiederholt. In CS verringern sich sowohl die Lemma- als auch die Tokenfrequenzen in beinahe allen Komplexitätsgraden. Ausgenommen davon sind lediglich Jans CS Grad 1 (Lemmas) und HSES-CS Grad 2 (Lemmas und Tokens).

Die Komplexitätsgrade in Komposita und Derivationen in CS und CDS (Lemmas und Tokens) werden in den folgenden Kapiteln nach Wortklasse und Wortbildungsart getrennt behandelt und detailliert beschrieben.

7.2.3 Morphologische Komplexität der Komposita

Die morphologische Komplexität nominaler Komposita und Adjektivkomposita in CS und CDS wird im Folgenden nach Datenpunkten (DP 1–4, INPUT-Korpus) und Zeiträumen (JAN, KAT, LEN) je Wortklasse analysiert. Bei den prozentualen Berechnungen werden die zusammengesetzten Lexeme auf Basis der Gesamtanzahl an Substantiven oder Adjektiven berechnet. Enthalten sind alle zusammengesetzten Substantive, auch Diminutivkomposita und synthetische Komposita. Bei der Zuordnung zu einem Komplexitätsgrad wurden allein die wortbildenden Prozesse berücksichtigt, unabhängig von etwaigen Interfixen.

7.2.3.1 Morphologische Komplexität nominaler Komposita

Untersucht werden alle nominalen Komposita, die in CS und CDS in den verschiedenen Zeiträumen oder zu distinkten Datenpunkten auftreten. Dabei dienen die Zeiträume dazu zu erkennen, wann die ersten komplexen Formen in CS auftreten und ob es eine Zunahme komplexer Komposita im Erwerbsverlauf sowohl in Lemmas als auch in Tokens gibt. Auch soll gezeigt werden, ob sich dies im sprachlichen Verhalten der Hauptbezugspersonen widerspiegelt.

Die morphologische Komplexität nominaler Komposita in Lemmas sind für Jan in Abbildung 43, für Kathi und Lena in Abbildung 44 und für die INPUT-Kinder in Abbildung 45 übersichtlich dargestellt. Für die prozentuale Verteilung der Substantivkomposita siehe Anhang 9 Abbildungen 76, 81 und 85. Dabei wird die Präferenz für einfach zusammengesetzte Substantive in CS und CDS offensichtlich, z. B. Bunt+stift, Eis+bär, die in großer Variation und Häufigkeit in allen Korpora bis 3;0 bzw. 4;11 auftreten. Lemma- und Tokenfrequenzen komplexerer Komposita

(Grad 2) steigen in CS mit zunehmendem Alter der Kinder Jan, Kathi und Lena, wohingegen sie bei den HSES-Kindern von DP 1–DP 4 annähernd gleichbleiben und bei den LSES-Kindern an Variation und Frequenz abnehmen.

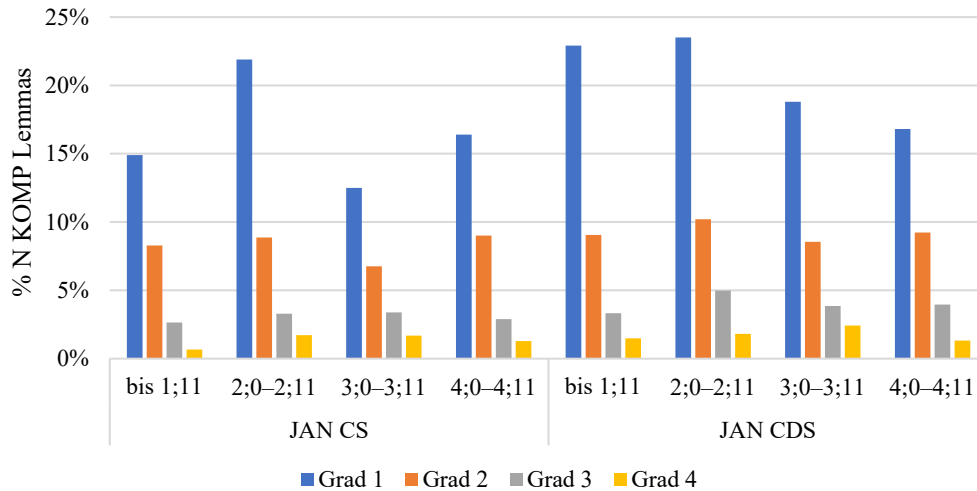


Abbildung 43. JAN-Korpus: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS

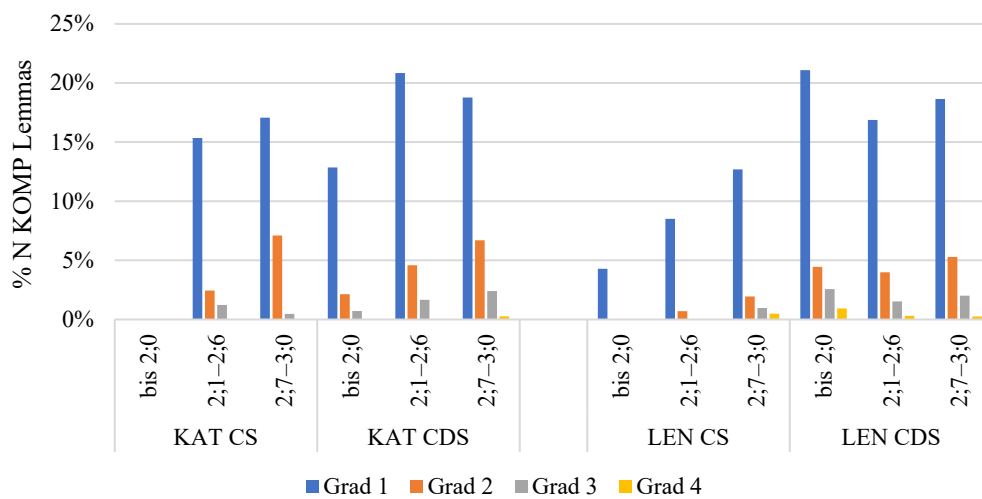


Abbildung 44. KAT- und LEN-Korpora: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS

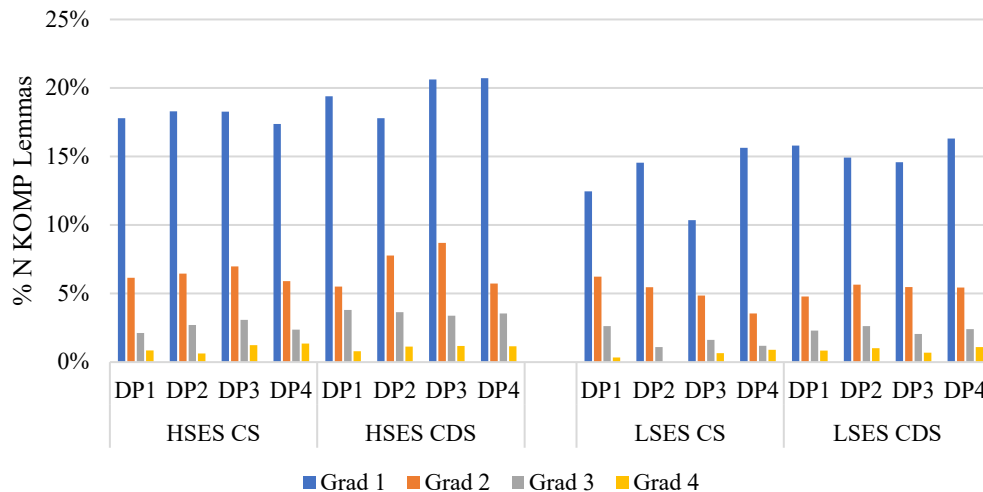


Abbildung 45. INPUT-Korpus: Komplexitätsgrade nominaler Komposita (Lemmas) in CS und CDS

Generell ist feststellbar, dass nur bei Jan ebenfalls komplexere Komposita in CS von Beginn an auftreten. Bei Kathi und Lena setzen die ersten komplexen N KOMP ab 2;1 ein und erhöhen sich stetig bis 3;0. Lenas Zeitraum 2;7–3;0 weist letztlich Bildungen bis zu einem Komplexitätsgrad von 4 auf, wenngleich auch in sehr geringem Ausmaß. Die INPUT-Kinder, deren Aufnahmen ab 3;0 beginnen, produzieren N KOMP in allen Komplexitätsgraden. In CS und CDS zeigt sich eine abnehmende Diversität und Verwendungshäufigkeit mit Zunahme der Komplexität.

Die Hauptbezugspersonen stellen ab der ersten Aufnahme zusammengesetzte Lexeme in allen Komplexitätsgraden zur Verfügung, mit Ausnahme von Kathis Mutter, deren komplexestes nominales Kompositum (Grad 4) im letzten Zeitraum auftritt.

Mehrgliedrige Komposita

Die ersten dreigliedrigen Komposita treten bei Jan bereits mit 1;8 zunächst noch in amalgamierter Form in Not+arzt+auto (notztauto) und gleich darauf in Doppel+decker+bus auf. Analog dazu bildet er mit 1;10 den Neologismus Doppel+decker+zug (vgl. Korecky-Kröll et al. 2017: 32). Dies setzt er mit weiteren Bildungen fort: Feuer+wehr+auto und Honomo+seil+bahn (1;11), Feuer+wehr+mann, Wind+schutz+scheibe, Kinder+zimmer+tisch, und dem Neologismus Bahn+hof+steg²⁹

²⁹ In folgender Äußerung (2;0): *JAN: naechste tation@ [: station] # bahnhofsteg.

(2;0). Jans letzte mehrgliedrige Lexeme sind Atom+kraft+werk (4;5), und die Neologismen Doppel+zug+karte (4;7) und Kunst+sport+schuh (4;8). Insgesamt bildet Jan bis 4;11 32 verschiedene dreigliedrige Komposita, die er 130-mal verwendet (vgl. Tabelle 22). Kathis einzige dreigliedrige Komposita im Alter von 2;8 sind Weih+nacht-s+markt und neologistisch *Weih+nacht-s+koppel, die immer in Verbindung mit dem Erstglied Weih+nacht auftreten, was bedeutet, dass sie dieses komplexe Wortbildungsmuster bis 3;0 nicht produktiv anwendet. Lena produziert bis zum Ende der Aufnahmen weder mehrgliedrige Komposita noch Phrasenkomposita.

Von den INPUT-Kindern, produzieren die HSES-Kinder 39 unterschiedliche KOMP 3-Lemmas, die sie 62-mal verwenden, die LSES-Kinder bilden mit 15 Lemmas deutlich weniger unterschiedliche Lexeme und verwenden diese mit 20 Tokens auch weniger oft. Neologistische Bildungen dreigliedriger Komposita finden sich sowohl bei HSES (7 Lem./7 Tok.) als auch bei LSES (3 Lem./3 Tok.): *Hunde+knochen+kette (D03-NIP 4;5), *Tür+steh-er+maus (D09-KAB 3;4), *Heiß+luft+ball (D07-NIG 2;11) in Analogie zu Heiß+luft+ballon, *Marien+käfer+spinne (D13-SOP 3;3) und *Tücher+hand+tasche (D18-ALM 4;7).

Viergliedrige Komposita kommen nur bei Jan und den HSES-Kindern des INPUT-Korpus vor. Jan bildet insgesamt vier Komposita mit vier Konstituenten, die er 5-mal verwendet: mit 2;0 die neologistische Form *Mist+haufen+durch+fahrt, gefolgt von dem Kopulativkompositum Motor+boot+dampf+schiff (2;5), Flug+hafen+feuer+wehr (2;8), und schließlich Spiel+zeug+u+sechs (3;8). Die HSES-Kinder produzieren drei viergliedrige Komposita: Haupt+feuer+wehr+mann (D11-DAG 3;4), und die beiden Neologismen *Rauch+fang+kehrer+schlange (D09-KAB 4;4) und *Easy+learn+spassi+gassi (D27-JOZ 4;7), die sie insgesamt 4-mal verwenden.

Phrasenkomposita

Einzig Jan produziert zwischen 2;0 und 2;6 erste Phrasenkomposita: Gute+nacht+geschichte (2;0), erweitert zu Dipsy+gute+nacht+geschichte und Po+gute+nacht+geschichte (2;3), sowie Goofy+und+Micky+buch (2;6), wie in Korecky-Kröll et al. (2017: 32) festgestellt. Bis 4;11 bildet Jan auch keine weiteren Phrasenkomposita. In CDS treten Phrasenkomposita ebenfalls nur vereinzelt auf, wie beispielsweise in

Ein+und+aus+schalt-er (LEN 1;7), Jan+und+Julia+geschichte (KAT 2;8) und Mensch+ärgerer+dich+nicht+maxl (JAN 2;10).

Die Lemma- und Tokenfrequenzen, sowie die prozentualen Anteile dreigliedriger (KOMP 3) und viergliedriger (KOMP 4) Substantivkomposita (inkl. Phrasenkomposita in Jans CS und allen CDS-Daten) gemessen an der Gesamtanzahl nominaler Komposita in CS und CDS wird in Tabelle 22 dargestellt. Dabei sind in der Gesamtanzahl der N KOMP alle Komposita des jeweiligen Korpus enthalten.

Tabelle 22. Prozentualer Anteil mehrgliedriger Substantivkomposita in Relation zur Gesamtanzahl nominaler Komposita in CS und CDS in allen Korpora

Korpus	JAN	KAT	LEN	INPUT HSES	INPUT LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N KOMP (gesamt)	375/1.959	77/138	49/106	479/950	219/432
KOMP 3	32/130	2/3	0/0	39/62	15/20
% KOMP 3/N KOMP	8,53/6,64	2,60/2,17	0,00/0,00	8,14/6,53	6,85/4,63
KOMP 4	7/9	0/0	0/0	3/4	0/0
% KOMP 4/N KOMP	1,87/0,46	0,00/0,00	0,00/0,00	0,63/0,42	0,00/0,00
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N KOMP (gesamt)	1.073/4.205	226/428	244/650	834/1.831	370/811
KOMP 3	123/263	11/16	15/22	80/134	23/35
% KOMP 3/N KOMP	11,46/6,25	4,87/3,74	6,15/3,38	9,59/7,32	6,22/4,32
KOMP 4	16/21	1/1	1/1	8/10	3/3
% KOMP 4/N KOMP	1,49/0,50	0,44/0,23	0,41/0,15	0,96/0,55	0,81/0,37

Wie aus Tabelle 22 ersichtlich, gibt es in CS und CDS einen höheren Anteil dreigliedriger Komposita im Vergleich zu viergliedrigen. Dies liegt zum einen an der morphologischen Komplexität der Lexeme, zum anderen an der damit einhergehenden schwereren Verständlichkeit. Viergliedrige Komposita kommen deshalb mit rund 0,8–1,5 % in CDS und mit 0,6–1,9 % in CS in Lemmas nur als Gelegenheitsbildungen vor, wie ein Vergleich der bis 4;11 reichenden Korpora JAN und INPUT zeigt. Dreigliedrige Substantivkomposita sind hingegen mit 6,9–8,5 % in CS und 6,2–11,5 % in CDS wesentlich diverser und werden auch häufiger verwendet (CS 4,6–6,6 %, CDS 4,3–7,3 %). Dabei bilden die LSES-Kinder und ihre Hauptbezugspersonen weniger mehrgliedrige Lexeme in Lemmas und Tokens als die vergleichbare HSES-Gruppe.

Synthetische Komposita

Eine besondere Stellung im Bereich der nominalen Wortbildung nehmen synthetische Komposita ein, da es sich wie bereits erwähnt, um zusammengesetzte Substantive mit einem deverbale abgeleiteten Letztglied handelt (vgl. Kap. 2.1.1). Für einen sprachübergreifenden Vergleich des Erwerbs synthetischer Komposita siehe Dressler, Sommer-Lolei, et al. (2019). In CDS sind durchschnittlich 7,64 % aller Substantivlemmas synthetische Komposita (siehe S. 90). Aufgrund ihrer internen Struktur weisen sämtliche Bildungen mindestens den Komplexitätsgrad 2 auf.

Tendenziell ist eine Zunahme der synthetischen Komposita in Lemmas und Tokens in CS und CDS in den Komplexitätsgraden 2 und 4 vom ersten zum zweiten Zeitraum (DP 1+2 zu DP 3+4; 3;0–3;11 zu 4;0–4;11) feststellbar (vgl. Abbildungen 37 und 38, und für Tokenfrequenzen Anhang A8.1 Abbildungen 67 und 68).

Sowohl die HSES- als auch die LSES-Kinder bilden neologistische synthetische Komposita. Beispiele für Neologismen in verschiedenen Komplexitätsgraden bei LSES (5 Lem./6 Tok.) sind *Bett+über-zieh-er statt Polster+über-zug (D08-FAJ 3;4), *Polizei+ge-räusch statt Sirene (D17-TOB 4;8), *Popo+klatsch ‚jemandem einen Klatsch auf den Popo geben‘ (D18-ALM 4;4), bei HSES (13 Lem./19 Tok.): *Apotheken+ver-käuf-er-in ‚wörtl. Verkäuferin in einer Apotheke‘ (D04-SOS 4;9), *Spring+falt-e statt Abschussrampe (für Autos) (D27-JOZ 4;3) und *Be-füll+stell-e statt Tankstelle (D28-MOH 4;8), siehe Beispiel (5), das gut zeigt, auf Basis welcher Überlegungen das Kind das Kompositum zusammensetzt:

(5) D28-MOH (4;8):

*PAR: nein is(t) nicht die tankstelle ?

*CHI: +< nein .

*CHI: da [//] das is(t) die befuellstelle .

*CHI: die befuellt dies(e)n wag(e)n .

7.2.3.2 Komplexitätsgrade der Adjektivkomposita

Adjektivkomposita dienen wie nominale Komposita dazu, den Wortschatz zu erweitern, indem Eigenschaften spezifiziert werden. ADJ KOMP sind in den frühen Daten in CS und CDS limitiert (durchschnittlich 4,77 % aller ADJ-Lemmas in CDS sind zusammengesetzt, siehe S. 116) und weisen mit wenigen Ausnahmen hauptsächlich Komplexitätsgrad 1 auf (vgl. Abbildung 46 für Lemmas und in

Anhang A9.1 Abbildung 77 für Tokens). Jans erstes Adjektivkompositum tritt im Alter von 2;0 in dem Lemma weiß+rot auf, Kathis einziges zusammengesetztes Adjektiv, braun+grün (2;3) ist ebenfalls ein Farbadjektiv. Auch die INPUT-Kinder bilden meist Farbadjektive, wie glitzer+gelb (D09-KAB 4;4) und dunkel+grün (D15-LUN 4;3). Lena bildet kein Adjektivkompositum. Auch in CDS ist die hauptsächliche Funktion der verwendeten Adjektivkomposita eine nähere Bestimmung verschiedener Farbnuancen: dunkel+gelb (JAN 3;0), rosa+rot (D23-SOK 4;5), hell+blond (D22-LAS 4;3).

Dabei sind Adjektivkomposita, die sich aus zwei Adjektiven zusammensetzen (AA) am häufigsten, neben Komposita, die mit einem Substantiv als Erstglied gebildet werden (NA). Dabei treten a) vergleichende Muster mit einfachen und komplexen Erstgliedern auf, wie in baby+blau (D31-AND 4;6), baby+leicht (D15-LUN 4;7), limetten+grün (CDS D23-SOK 4;3), erd+beer+rot (CDS KAT 2;5, Grad 2), *giftig+grün (CDS D04-SOS 3;5, Grad 2), b) intensivierende oder vermindernde Muster vom Typ riesen+gross (JAN 2;0), super+toll (CDS D05-FLB 3;0), super+cool (CDS D04-SOS 4;5), blitz+schnell (D03-NIP 4;8), wunder+schön (D29-SAK 3;5), halb+klein (D07-NIG 4;2), halb+fertig (D21-PAS 3;6), mittel+schwer (CDS JAN 2;9), wutze+klein (CDS D05-FLB 4;4), c) einfache Substantive als erste Konstituente, die das Letztglied semantisch näher bestimmen, wie in Jans CDS: farben+blind (1;11), wasser+scheu (1;10), wasser+dicht (2;5), wasser+lös-lich (4;8).

Die einzigen Zusammensetzungen von Zahlwörtern (NUM) und Adjektiven sind mit Komplexitätsgrad 2: drei+eck-ig (D09-KAB 3;1), sechs+eck-ig (CDS D29-SAK 4;9), und bei Jan (4;2) handelt es sich um graduelle Abstufungen mit dem Zweitglied lecker: zwei-t+lecker, fünf-t+lecker, sechs-t+lecker. NUMA-Komposita mit Komplexitätsgrad 3 sind in CDS: zwei+färb-ig (D29-SAK 3;5), und drei+färb-ig (D02-SIJ 3;2).

Bei den weiteren komplexen Bildungen mit Komplexitätsgrad 3 und 4 handelt es sich einerseits um das Lexem kurz+ärml-ig (D02-SIJ 4;11, Grad 4), sowie in Jans CDS um ADJ KOMP, die ein Perfektpartizip oder Präsenspartizip als Kopf aufweisen: weiß+rot+ge-streif-t (2;8, Grad 3), furcht+er-reg-end (2;2, Grad 3) und

weih+nachts+ge-schmück-t (2;9, Grad 4), die jedoch nur isoliert auftreten. Generell sind vielgliedrige Adjektive in CS und CDS als Ausnahme anzusehen.

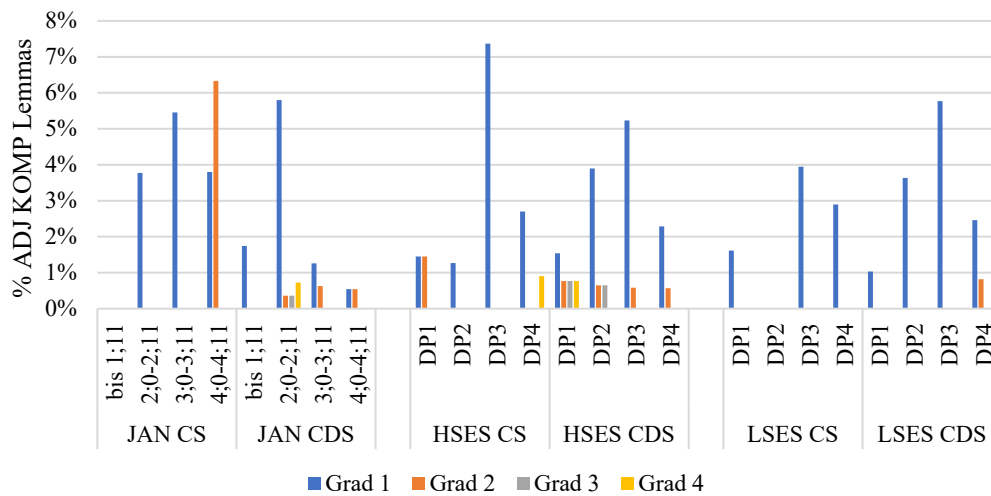


Abbildung 46. JAN- und INPUT-Korpus: Komplexitätsgrade von Adjektivkomposita (Lemmas) in CS und CDS

Zusammengefasst lässt sich eine eindeutige Präferenz für Adjektivkomposita feststellen mit Komplexitätsgrad 1, wobei auch hier ein klarer Anstieg in Lemmas und Tokens der ADJ KOMP im DP 3 (mit rund 4 Jahren) in CS erkennbar ist (vgl. Abbildung 46, und in Anhang A9.1 Abbildung 77). Komplexere ADJ KOMP kommen nur selten vor und liegen meist unter 1 %. Die hohe Diversität an Grad 2-Komposita bei Jan im letzten Zeitraum liegt an der vielfältigen Verwendung von Bildungen in Kombination mit lecker (siehe oben).

7.2.4 Morphologische Komplexität der Derivate

Die morphologische Komplexität der Derivationen in CS und CDS wird im Folgenden nach Zeitpunkten (DP 1–4, INPUT-Korpus) und Zeiträumen (JAN, KAT, LEN) nach Wortklasse getrennt untersucht. Wie bereits erwähnt errechnen sich die prozentualen Anteile der derivierten Lexeme auf Basis der jeweiligen Grundgesamtheit der Wortklasse.

7.2.4.1 Morphologische Komplexitätsgrade derivierter Substantive

In die Analyse fließen alle Substantive mit ein, die durch mindestens einen Wortbildungsprozess (Affigierung, Konversion, Stammmodifikation, Stammvokaländerung) entstehen. Es sind alle explizit und implizit abgeleiteten Substantive

enthalten, daher auch Diminutivbildungen. Überdies wurden auch pseudo-derivierte Substantive in die Analyse mitaufgenommen, da die Kinder zu Beginn des Spracherwerbs in der prämorphologischen Phase noch keine Regeln ableiten oder morphologische Muster erkennen, und eine Vielzahl an pseudo-derivierten Lexemen begünstigend zur Herausbildung eines Wortbildungsmusters beitragen kann.

Die prozentuale Verteilung der Komplexitätsgrade derivierter Substantive in Lemmas, berechnet an der Grundgesamtheit der Substantive, in CS und CDS sind für Jan in Abbildung 47, für Kathi und Lena in Abbildung 48 und für die INPUT-Kinder in Abbildung 49 veranschaulicht, für die Darstellung in Tokens siehe Anhang 9 Abbildungen 73 (JAN), 78 (KAT, LEN) und 82 (INPUT).

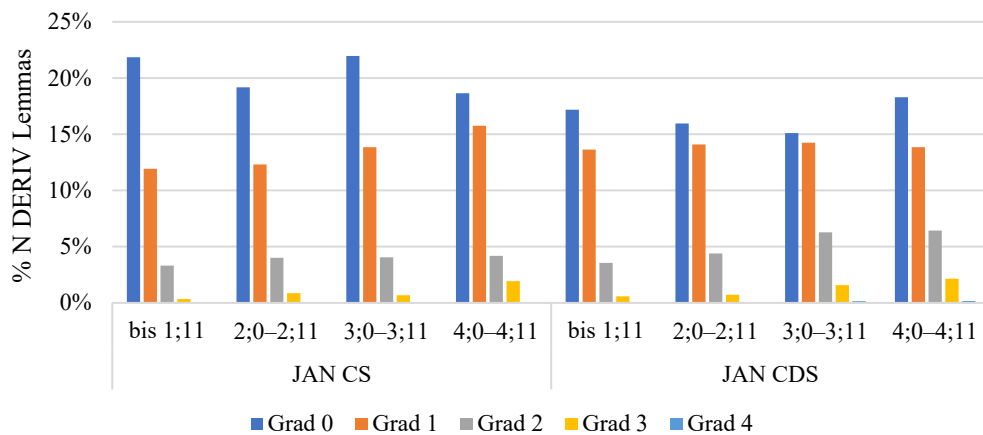


Abbildung 47. JAN-Korpus: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS

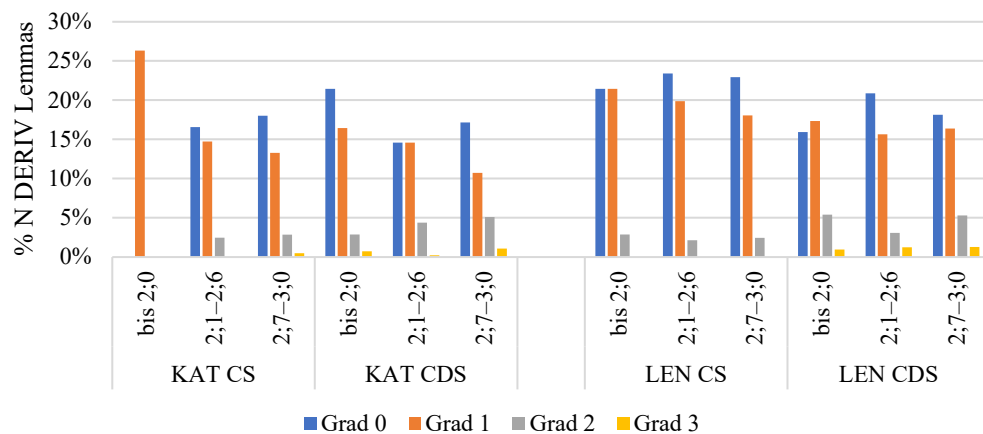


Abbildung 48. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS

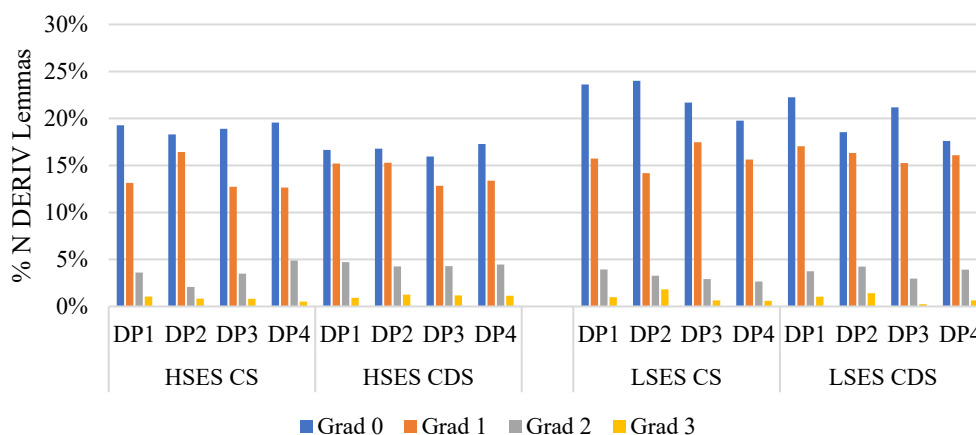


Abbildung 49. INPUT-Korpus: Komplexitätsgrade derivierter Substantive (Lemmas) in CS und CDS

Wie in Abbildungen 47, 48 und 49 dargestellt, sind nominale Ableitungen mit dem Komplexitätsgrad 1 von Beginn an in großer Variation vorhanden. In geringerem Ausmaß kommen jedoch auch bereits komplexere Formen in der Kindersprache vor. Rund um den dritten Geburtstag weisen die Kinder (JAN und INPUT ab 3;0, KAT und LEN 2;7–3;0) einen Anteil von mehr als 14 % an einfachen (Grad 1, mit einem Wortbildungsprozess) Derivationen, 3 % an komplexeren (Grad 2) und knapp 1 % an komplexen Bildungen (Grad 3) in Lemmas auf (vgl. Tabelle 21).

Generell treten einfache deverbale, nominale Derivationen (mit einem Affix, Komplexitätsgrad 1) bei Jan, Kathi und Lena früh und vor den ersten komplexen Formen (Grad 2 und 3) auf. Bei Jan: Schleck-er (1;8, Grad 1) vor Stieg-e (1;8, Grad 2) und An#fäng-er (2;8, Grad 3), bei Kathi: Reit-er (1;11, Grad 1) vor Zug (2;5, Grad 2) und Eingang (2;9, Grad 3), bei Lena: Hock-er, Deck-e (1;9, Grad 1) vor Wäsch-e (1;11, Grad 2). Lena weist bis zum Ende ihrer Aufnahmen keine derivierten Substantive mit Komplexitätsgrad 3 auf. Die INPUT-Kinder, deren Aufnahmen ab 3 Jahren beginnen, verwenden in allen Aufnahmen nominale Derivationen bis zu Komplexitätsgrad 3. In CDS zeigt sich die Verwendung derivierter Substantive in allen Komplexitätsgraden über den gesamten Zeitraum, wobei komplexere Ableitungen mit der Zeit an Diversität gewinnen.

Jan und Lena produzieren von Anfang an pseudo-derivierte Formen (Simplex-Substantive, die rein formal ein Affix aufweisen, Grad 0), z. B. Hase (JAN 1;3), Antenne (JAN 1;8), Wasser (LEN 1;10), Kathi hingegen erst ab dem Zeitraum 2;1–

2;6, da sie davor die Verwendung von Diminutiven (Grad 1) bevorzugt (vgl. Abbildung 48). Rund um 3 Jahre weisen jedoch alle Kinder und ihre Hauptbezugspersonen pseudo-derivierte Lexeme von über 16 % in Lemmas auf. Am häufigsten handelt es sich dabei um unbetontes -e im Auslaut, gefolgt von -er und -el, z. B. Kugel (D23-SOK 3;3).

Derivationen mit dem Komplexitätsgrad 1 sind in allen Daten sowohl in CS als auch in CDS in Lemmas und Tokens am frequentesten (vgl. Abbildungen 47, 48, 49, und in Anhang 9 Abbildungen 73, 78, 82). Dies liegt zum einen an der Einfachheit der Bildung, zum anderen an der morphosemantischen und -taktischen Transparenz sowie der positionellen Salienz von Suffigierungen, und auch, dass es sich in der Mehrheit um Konkreta, insbesondere um Instrumente (NIN) oder Personen (NAG) handelt. Suffigierungen verbaler oder nominaler Basen mit -e, und -er sind dabei häufiger als Konversionen, wie (die) Grün (JAN 2;7), oder implizite Ableitungen, wie Fluss (D12-JAS 4;6). Jan, der eine Affinität für Fahrzeuge aufweist, bildet eine Vielzahl an nominalen Ableitungen, die mit dieser Kategorie in Verbindung stehen: Flitz-er (1;8), Trakt-or (FSUFF 1;8), Transport-er (1;8), Last-er (1;8), Roll-er (1;10) und Fahr-er (1;10).

Nominale Ableitungen des Komplexitätsgrades 2 sind Bildungen, die mittels zweier Wortbildungsprozesse gebildet werden. Dies können Ableitungen oder Konversionen von Präfix- oder Partikelverben sein, z. B. Be-mal-ung (D04-SOS 4;9), Be-such (D27-JOZ 3;0), Ab#wasch (D28-MOH 4;8), oder implizite Derivationen, die eine Veränderung des Wortstamms, von dem sie sich ableiten, mit sich bringen, oder/und im Zuge der Wortbildung zusätzlich zu einer Um- oder Ablautung des Stammvokals führen, z. B. Ge-büsch (D10-LER 4;7), Schnitt ← schneiden (D13-JOP 4;4), Schläg-er (LEN 2;9) und Hilf-e (D26-JUB 4;9), wobei es sich dabei hauptsächlich um deverbale Ableitungen handelt und Adjektive als Ableitungsbasen selten sind, da diese abstrakte Nomina qualitatis (NQU) bilden, z. B. Güt-e (D25-PHG 4;4), Größ-e (D18-ALM 3;3), Näh-e (D31-AND 4;6). Die Suffigierung eines bereits präfigierten Adjektivs kommt einmalig in Un-eben-heit (D11-DAG 4;9) vor.

Ebenso kommt es nur vereinzelt zu Bildungen, die zwei oder mehr derivationelle Suffixe enthalten. Dazu zählen Ableitungen auf -(l)er-in und -or-in, wie in Künst-

ler-in und Dokt-or-in, und bei Einschluss von Zirkumfigierung auch -ig-ung in Entschuld-ig-ung. Bis zu einem Alter von 5 Jahren kommt es weder in CS noch in CDS zu einer nennenswerten Anzahl nominaler Ableitungen mit zwei (oder mehr) Suffixen.

Wesentlich frequenter präsentiert sich die Verwendung mehrerer Affixe (Präfix und Suffix). Wenngleich es bei Kathi und Lena bis 3 Jahre zu keinen mehrfachen Affigierungen kommt, so zeigt sich bei Jan und den INPUT-Kinder mehr Variation. Dabei kommt doppelte Affigierung nach dem ersten Auftreten synthetischer Komposita vor, wenn auch im gleichen Altersmonat, wie im Beispiel von Jan: 1;8 Hub+schraub-er vs. An#häng-er. Bei Kathi und Lena lassen sich erste synthetische Komposita feststellen, jedoch keine Lexeme mit zwei Affixen (KAT 2;3, LEN 2;11). Bei den älteren INPUT-Kindern sind beide Muster vertreten: Spiel+an#leitung (D07-NIG 2;11) und Ab#sperr-ung (D07-NIG 4;2), Schraube-n+zieh-er (D27-JOZ 3;0) und Auf#fahr-t (D27-JOZ 3;0), Be-wach-er (D27-JOZ 4;7), wobei Ableitungen von Partikelverben zu Beginn wesentlich häufiger sind als von Präfixverben. Für die Analyse der Komplexität synthetischer Komposita in CS und CDS siehe Kap. 7.2.3.1, S. 164.

Neologistische Bildungen derivierter Substantive kommen bei den INPUT-Kindern in verschiedenen Komplexitätsgraden vor, wie in *Zu#sperr-ung³⁰ statt Schloss (D11-DAG 4;6, Grad 2), *Ver-schwemm-ung³¹ statt Überschwemmung (D04-SOS 4;5, Grad 3), *Bohr-e³² statt Bohrmaschine (D27-JOZ 3;0, Grad 1) und bei Jan als morphologische Konversionen in den kreierten Sprachbezeichnungen *Paul-inisch (2;2, Grad 2) und *Jan-isch (2;8, Grad 2), sowie in *Schwimm statt See (2;4, Grad 1) vor, siehe Beispiel (6):

- (6) JAN (2;4):
*MUT: # und # wo is(t) das da hm ?
*JAN: ein schwimm@c [: schwimm].
*MUT: ein schwimm@c ?
*JAN: ja .
*MUT: an einem see is(t) das .

³⁰ In der Äußerung (4;6): *CHI: ## und die &zuspre [//] zusperrung # bei [/] bei der # garage.

³¹ In der Äußerung (4;5): *CHI: wie der Pupsi nicht das wassa [: wasser] &a [//] zuggedreht hat # so is(t) das ganze haus &vo [//] von [*] &ver [//] verschwemmung@m gangen [*].

³² In der Äußerung (3;0): *CHI: vielleicht mit der bohre@n.

Movierung

Vorrangig wird im Deutschen zur Ableitung femininer Personenbezeichnungen das Suffix *-in* gebraucht. Auch in den untersuchten Korpora bis 5 Jahre ist *-in* das dominante Motionssuffix, wohingegen in den Daten bis 3 Jahre keine movierten Formen in CS auftreten.³³ Dies ist insofern zu erklären, als das Kind zuerst kognitiv dazu in der Lage sein muss, zwischen den Merkmalen weiblich und männlich zu differenzieren, was gemäß Kalédaité und Savickienė (2007) im Alter von rund 2;6 der Fall ist. Bei Jan kommt *-in* mit 3;7 erstmals in dem abgeleiteten Lexem *Lehr-er-in*, in Kombination mit dem Suffix *-er*, vor und bleibt in den Daten bis 4;11 auch seine einzige Movierung. Generell sind Bildungen von Basen, die auf *-er* enden in CS und CDS am gebräuchlichsten: *Käuf-er-in* (D09-KAB 3;4), *Lehr-er-in* (D31-AND 3;7) und *Ver-käuf-er-in* (D07-NIG 3;4). Movierte Neologismen finden sich in CDS in der deverbalen Ableitung *Stibitz-er-in* ← *stibitzen* (D10-LER 3;4) und denominal in *Brösl-er-in* ← *Brösel* (JAN 1;11). Feminine Berufsbezeichnungen, die allein in der maskulinen Form ohne das Motionssuffix *-in* benannt werden, kommen weder in CS noch in CDS vor.

Nominale Basen zur femininen Movierung mittels *-in* treten bei den INPUT-Kindern ab 3;3 auf: *Freund-in* (D22-LAS 3;3), *König-in* (D24-RAG 3;3), *Nachbar-in* (D08-FAJ 4;4), *Dokt-or-in* (D04-SOS 4;9) nach Fremdsuffix, und in *Kolleg-in* (D04-SOS 4;5) mit Entfall des Schwalauts. Erst nach dem Alter von 5 Jahren verwendet auch Jan nominale Ableitungsbasen, z. B. *Österreich-er-in* (5;4), *Künstler-in* (5;6), *Freund-in* (5;7) (vgl. Sommer-Lolei & Dressler akzeptiert).

Es gibt keine Movierungen von Basen, die in der maskulinen Form auf *er-er* enden, wie beispielsweise *Zauber-er* → *Zauber-in*, und daher die Tilgung von *-er* verlangen, weder in CS noch in CDS. Irreguläre Bildungen vom Typ *Prinz* → *Prinzess-in*, die eine Ableitung von einer nicht-autonomen lexikalischen, aus dem französischen entlehnten, Basis darstellen, kommen nur in dieser Form und ausschließlich in den INPUT-Daten, sowohl bei HSES als auch bei LSES vor. Die

³³ Dies entspricht Rainers (2010: 141) erster dokumentierter Movierung seiner Tochter Carmen im Alter von 3;1 (*Freund-in*).

auch in ADS überaus seltenen maskulinen Movierungen vom Typ -erich sind in CS nicht vorhanden und treten in CDS isoliert in Ente → Ent-erich (JAN 2;1) auf.

7.2.4.2 Morphologische Komplexitätsgrade derivierter Verben

In der Auswertung sind alle Verben enthalten, die mindestens einen Wortbildungsprozess beinhalten (Affigierung, morphologische Konversion, Stammvokaländerung, Stammmodifikation). Wie auch bei den Substantiven werden auch bei der Wortklasse der Verben pseudo-derivierte Lexeme in die Analyse integriert (vgl. Kap. 7.2.4.1).

Die prozentuale Verteilung der Komplexitätsgrade derivierter Verben (Lemmas), gemessen an der Grundgesamtheit an Verben, in CS und CDS ist für Jan in Abbildung 50, für Kathi und Lena in Abbildung 51 und für die INPUT-Kinder in Abbildung 52 dargestellt, für die prozentualen Anteile in Tokens siehe Anhang 9 Abbildungen 74 (JAN), 79 (KAT, LEN) und 83 (INPUT).

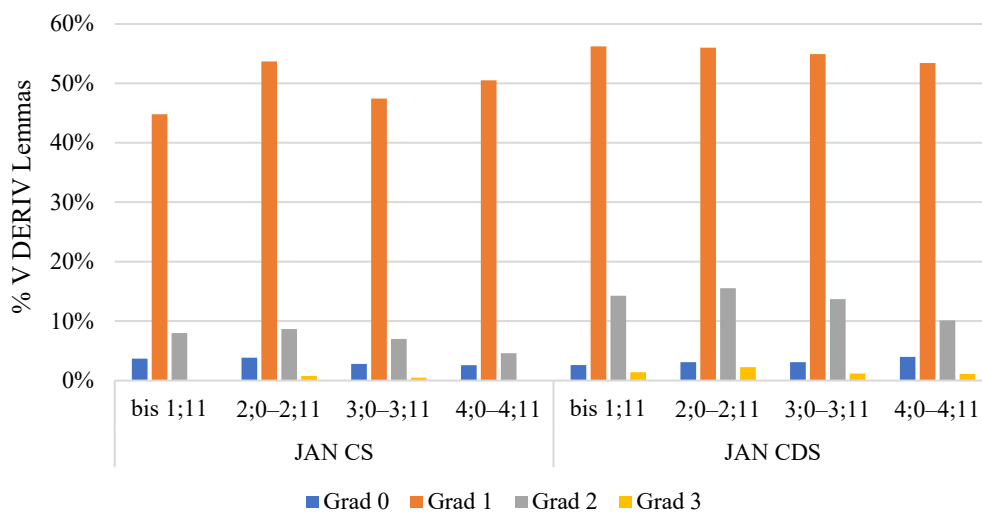


Abbildung 50. JAN-Korpus: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS

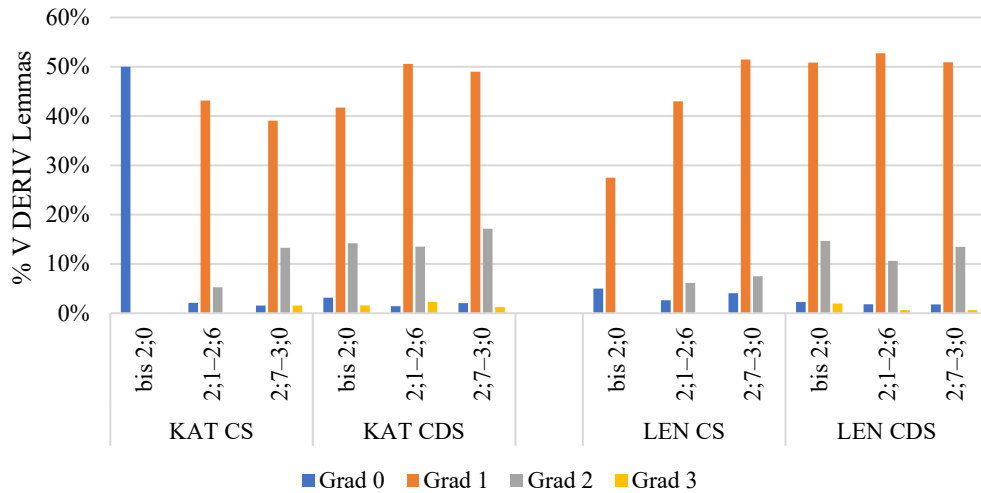


Abbildung 51. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS

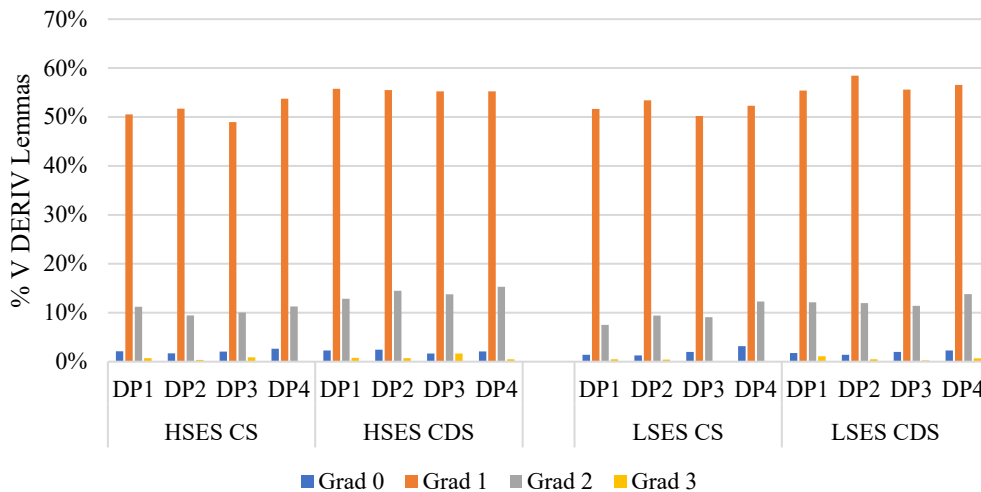


Abbildung 52. INPUT-Korpus: Komplexitätsgrade derivierter Verben (Lemmas) in CS und CDS

Aus den Abbildungen 50, 51 und 52 geht eindeutig hervor, dass abgeleitete Verben in der überwiegenden Mehrheit Komplexitätsgrad 1 aufweisen. Derivate, die einen einzigen Wortbildungsprozess beinhalten sind in allen Korpora sowohl in CS als auch in CDS vom ersten³⁴ bis zum letzten Datenpunkt in sehr hohem Ausmaß vorhanden und zeigen sich nicht nur an der enormen Vielzahl unterschiedlicher Lemmas, sondern auch an der höheren Verwendungshäufigkeit (Tokens). In wesentlich geringerem Ausmaß treten allerdings auch komplexere Lexeme mit Grad 2 in CS und CDS auf. Wie bereits in Kap. 7.2.1, in Tabelle 21 einleitend zusammengefasst, bilden die Kinder im Alter von rund 3 Jahren mehr als 48 % an

³⁴ Mit Ausnahme von Kathi (CS), deren einfache und komplexe Verbderivationen ab 2;1 auftreten.

einfachen Verbderivationen (Grad 1), über 9 % an Grad 2-Bildungen und unter 1 % an komplexen verbalen Ableitungen (Grad 3). Die Hauptbezugspersonen bieten den Kindern vom ersten Zeitpunkt abgeleitete Verben in allen Komplexitätsgraden.

Einfach abgeleitete Verben mit Komplexitätsgrad 1 treten bei Jan und Lena früher als die ersten derivierten Substantive auf (vgl. Kap. 7.2.4.1) und bei allen Kindern früher als die ersten komplexeren Verbformen (Grad 2 und 3). Bei Jan: an#halten (1;6, Grad 1) vor drauf#setzen (1;9, Grad 2) und hinein#ge-hören (2;5, Grad 3), bei Kathi: zu#decken (2;3, Grad 1) vor weg#räumen (2;4, Grad 2) und hinein#räumen (2;8, Grad 3), bei Lena: auf#schauen (1;8, Grad 1) vor hin#legen (2;3, Grad 2). Lena produziert keine abgeleiteten Verben mit Komplexitätsgrad 3. Die INPUT-Kinder verwenden ab DP 1 Verbderivationen bis Grad 3, wenngleich derart komplexe Bildungen nur vereinzelt vorkommen. Sämtliche derivierte Verben treten zuerst im Infinitiv auf, bevor sie in flektierter Form verwendet werden. Zeitgleich mit infiniten Formen derivierter Verben treten finite Simplex-Verbformen wie etwa lach-t ‚lach-3S‘ (JAN 1;5), geh-t ‚geh-3S‘ und fahr-en ‚fahr-3P‘ (JAN 1;8) früh in CS auf. Dies spricht gegen generative Annahmen einer einheitlichen *bare-infinitive*-Analyse (vgl. Laaha & Bassano 2013). Diese nicht flektierten Infinitive derivierter Verben treten hauptsächlich in satzfinaler Position, wie z. B. Mann reinsteig(e)n (JAN 1;7) auf und nur selten satzinitial, wie in aussteig(e)n Mann (JAN 2;0), was auf die größere positionelle Salienz am Satzende schließen lässt (vgl. Wijnen et al. 2001: 639; Gillis 2003).

Pseudo-derivierte Verblemmas, also Simplex-Verben, die nur formal ein Affix aufweisen, kommen in CS und CDS, mit Ausnahme von Kathis CS im ersten Zeitraum, nur sporadisch vor und bleiben über den gesamten Zeitraum hinweg in allen Korpora unter 5 % in Lemmas. Hauptsächlich handelt es sich um pseudo-suffigierte Verben wie **knabbern**, **flattern**, **stolpern**, sowie um Pseudo-Präfigierungen wie in **gewinnen**, **beginnen**, **verlieren**, **vergessen**.

Derivierte Verben mit Komplexitätsgrad 1 sind sowohl in CS als auch in CDS wie bereits erwähnt größtenteils Partikelverben. Es kommen jedoch auch Bildungen mit untrennbaren Präfixen wie be-halt-en (D16-LEP 4;11), er-rat-en (D06-LAB 4;8), ver-such-en (D24-RAG 4;7) und zer-stör-en (D18-ALM 3;0) vor, sowie morphologische Konversionen, die denominal abgeleitet werden, wie in angel-n

(D14-KEW 3;4), blut-en (D05-FLB 4;4), knie-n (D26-JUB 4;5), und deadjektivische Konversionen wie in leer-en (D28-MOH 3;0) und spitz-en (D12-JAS 4;9). Ebenso bilden die Kinder erste implizite Ableitungen mit Ablaut oder Stammmodifikation, wie leg-en (JAN 1;8), stell-en (JAN 1;9), setz-en (D19-VAS 4;7) sowie mit Umlaut: färb-en (D18-ALM 4;4), fütter-n (D12-JAS 4;6), kämpf-en (D27-JOZ 4;7) und glänz-en (D31-AND 4;6). Denominale und deverbale suffigierte Verben wie transport-ier-en (JAN 1;11), telefon-ier-en (D07-NIG 4;5), schütt-el-n (JAN 1;9) und streich-el-n (JAN 1;10) kommen in geringerer Variation und Frequenz vor (vgl. dazu Abbildungen 39 und 69).

Auch im Komplexitätsgrad 2 sind Partikelverbbildungen in CS und CDS am häufigsten. Andere Wortbildungsmuster, darunter Präfigierung, Suffigierung, Zirkumfigierung und morphologische Konversion kommen nur gelegentlich vor. Auffallend sind die zahlreichen Partikel, die die Verben setz-en und stell-en, näher bestimmen. Beispielsweise kombiniert Jan setz-en mit 7 Partikeln: auf#, rauf#, drauf#, rein#, nieder#, her#, zusammen#, und stell-en mit 6 Partikeln: um#, runter#, rauf#, auf#, hin#, her#. Beispiele für die wenigen Bildungen mit anderen Wortbildungsmustern sind bei Zirkumfigierung be-ruh-ig-en (D04-SOS 4;9) und ent-schuld-ig-en (D18-ALM 3;3), bei morphologischen Konversionen ärger-n (LEN 2;8) und frühstück-en (JAN 1;11), bei Suffigierung repar-ier-en (D29-SAK 3;5) und mit zusätzlicher Umlautung des Stammvokals: künst-el-n (D18-ALM 3;0) und blätt-er-n (D27-JOZ 4;7), bei präfigierten Verben: er-tränk-en (D02-SIJ 4;5), zer-leg-en (D11-DAG 4;6) und über-leg-en (JAN 3;8).

Abgeleitete Verben, die Komplexitätsgrad 3 aufweisen, sind größtenteils Partikelverben, die eine Doppelpartikel beinhalten wie in rund#herum#bau-en, hin#über#stell-en (D11-DAG 4;6), hin#ein#räum-en (KAT 2;8), hin#ein#ge-hör-en (JAN 2;5). Bildungen, die eine Partikel und ein Suffix enthalten, treten vereinzelt auf wie in um#blätt-er-n (JAN 2;6) und weiter#blätt-er-n (JAN 3;6), ebenso wie das isoliert auftretende präfigierte Verb ver-ab#schied-en (D04-SOS 4;5). Es gibt keine Verbderivationen mit Komplexitätsgrad 4 in CS und CDS.

Neologistische Verbderivationen treten nur isoliert und in verschiedenen Komplexitätsgraden auf. Bei Jan in *be-hol-en (3;0), *aus#tröpf-el-n (2;11), bei Kathi in *re-par-it-ier-en (2;10), bei Lena in *fern#an#schau-en (2;7) und *zu#leg-

en (3;0). Etwas mehr Variation ist bei den INPUT-Kindern zu finden: LSES-CS (3 Lem./4 Tok.) *über#legen³⁵ (D08-FAJ 4;4), *blut-isier-en³⁶ (D22-LAS 4;5) und in HSES-CS (7 Lem./9 Tok.) *her#be-reit-e³⁷ (D04-SOS 3;2), *stech-elt³⁸ (D13-JOP, 3;0), *ver-glitzern³⁹ (D29-SAK 4;9).

7.2.4.3 Morphologische Komplexitätsgrade derivierter Adjektive

Im Folgenden werden alle Adjektive analysiert, die durch Wortbildung entstehen. Es sind daher alle derivierten Adjektive enthalten, die zumindest ein Wortbildungsmuster enthalten (Affigierung, Konversion). Wie bereits bei den Substantiven und Verben werden ebenfalls bei den Adjektiven pseudo-derivierete Wortformen miteinbezogen (vgl. Kap. 7.2.4.1, 7.2.4.2).

Die prozentualen Anteile der Komplexitätsgrade abgeleiteter Adjektive in Lemmas in CS und CDS sind für Jan in Abbildung 53, für Kathi und Lena in Abbildung 54, und für das INPUT-Korpus in Abbildung 55 illustriert. Die prozentuale Verteilung der Adjektivtokens ist in Anhang 9 Abbildungen 75 (JAN), 80 (KAT, LEN) und 84 (INPUT) dargestellt.

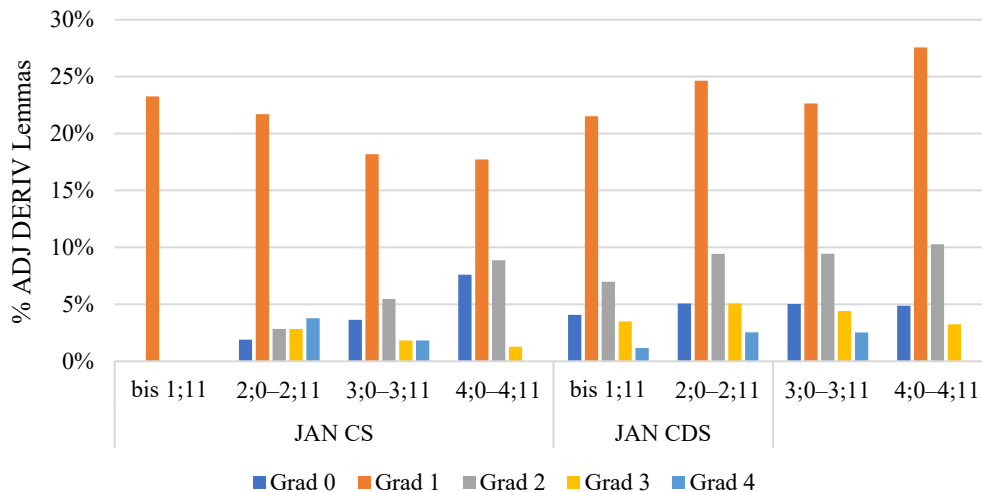


Abbildung 53. JAN-Korpus: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS

³⁵ In der Äußerung (4;4): *CHI: so dann xxx das legma [: legen wir] dann so ueber [*].

³⁶ In den Äußerungen (4;5): wo man ## blutisieren@n kann; so kann man blutisieren@n.

³⁷ In der Äußerung (3;2): *CHI: aber ich bereite schon die teller her [*] fuer die gaeste.

³⁸ In der Äußerung (3;0): *CHI: naemlich mami das da stechelt@n.

³⁹ In der Äußerung (4;9): *CHI: ## wir woll(e)n aba [: aber] alles verglitzern@n.

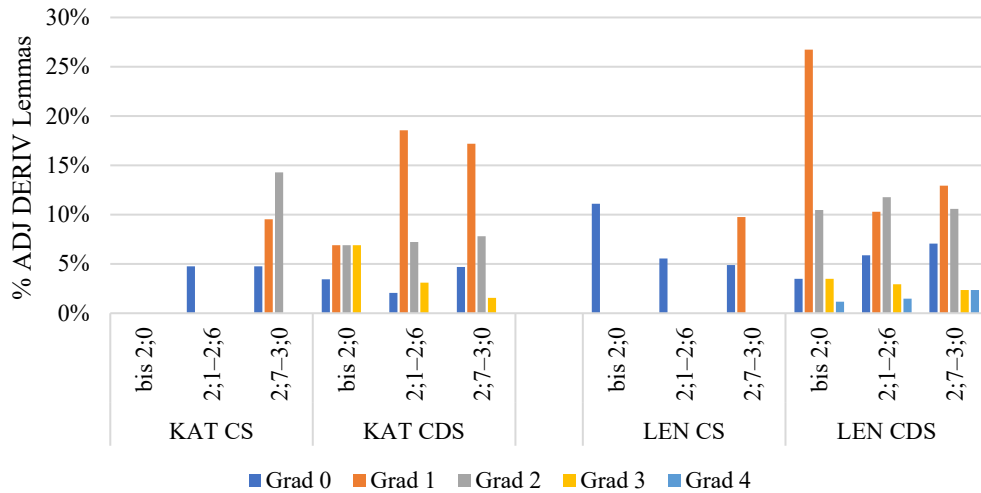


Abbildung 54. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS

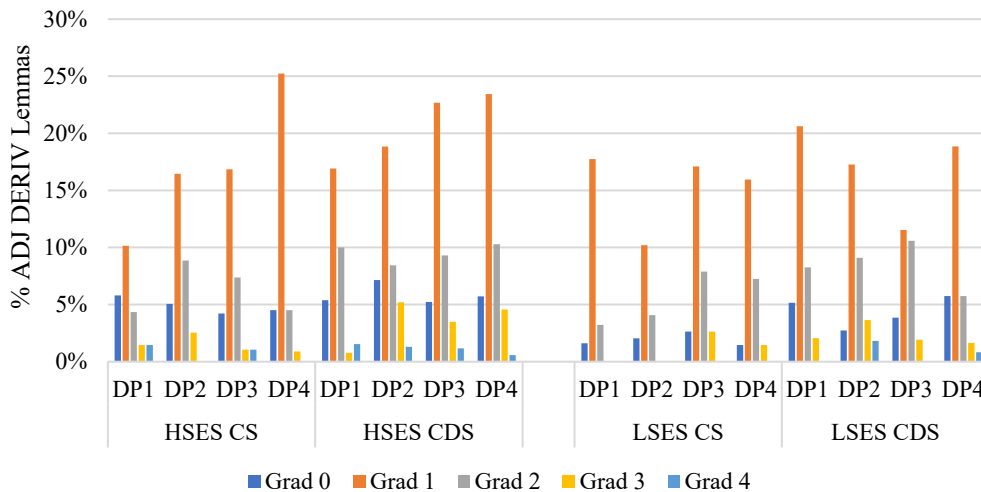


Abbildung 55. INPUT-Korpus: Komplexitätsgrade derivierter Adjektive (Lemmas) in CS und CDS

Derivierte Adjektive sind in allen Korpora, besonders in den frühen Daten von Kathi und Lena, sowohl in CS als auch in CDS, wie von der späten Emergenz von Adjektiven im Allgemeinen zu erwarten (vgl. Tribushinina et al. 2015), nur in geringer Anzahl vertreten, da Simplex-Adjektive in der Verwendung bevorzugt werden. Bei jenen derivierten Adjektiven, die in CS vorkommen, überwiegt bei Jan und den INPUT-Kindern die Zahl der einfach abgeleiteten Adjektive (Grad 1). In CDS werden den Kindern derivierte Adjektive bis zu einem Komplexitätsgrad von 3 bzw. 4 von Beginn an, wenn auch in geringer Variation und Frequenz, zur Verfügung gestellt.

Die Kinder produzieren im Alter von durchschnittlich 3 Jahren, zusammengekommen etwas mehr als 13 % an einfachen Adjektiven (Grad 1), mehr als 5 % an komplexeren abgeleiteten Adjektivlemmas (Grad 2) und rund 1 % an komplexen Formen mit Komplexitätsgrad 3 und 4 (vgl. Kap. 7.2.1, Tabelle 21).

Simplex-Adjektive, die pseudo-deriviert sind (Grad 0), kommen in suffigierter Form vor, wie beispielsweise in **fertig**, **wichtig**, **ewig** und **dämlich**. Derivierte Adjektive mit Komplexitätsgrad 1 treten bei Jan gleichzeitig mit den ersten nominalen Derivationen auf, bei Kathi und Lena später als die ersten derivierten Substantive und Verben (vgl. Kap. 7.2.4.1, 7.2.4.2). Kathi, die ihre ersten Verbderivationen ab 2;3 bildet, verwendet auch ihre ersten abgeleiteten Adjektive im Vergleich zu dem gleichaltrigen Kind Jan etwas später, nämlich im Zeitraum 2;7–3;0, ebenso wie Lena. Einfache Adjektivderivationen (Grad 1) treten bei Jan früher als die ersten komplexen Adjektivderivationen ab Grad 2 auf. Jan bildet **kitzlig** (1;8, Grad 1) vor **biss-ig** (2;2, Grad 2) und **sicht-bar** (2;0, Grad 3), Kathi: irreguläre Stammmodifikation der Numerale **ein-** in **er-st-er** (2;8, Grad 2) vor der ebenso adjektivischen Ordinalzahl **vier-t-en** (2;9, Grad 1) und **gefähr-lich** (2;9, Grad 2), Lena: **zwei-t-er** (2;7, Grad 1) vor **lust-ig** (2;9, Grad 1). Lena verwendet ausschließlich derivierte Adjektive mit Komplexitätsgrad 1. Bei den HSES-Kindern treten von Beginn an derivierte Adjektive bis Grad 4 auf, bei den LSES-Kindern nimmt die Komplexität der Bildungen ab DP 3 zu (vgl. Abbildung 55). Jans erstes präfigiertes Adjektiv ist **un-sicht-bar** (2;6), die beiden Mädchen bilden bis zum Ende ihrer Aufnahmen keines.

Im Bereich der Ableitung von Adjektiven sind in den Komplexitätsgraden 1–3 in CS und CDS hauptsächlich **-ig** und **-(s)t** maßgeblich an der Bildung beteiligt, während die Suffixe **-lich** und **-isch**, sowie sämtliche anderen Affixe und Wortbildungsmuster eine untergeordnete Rolle spielen. Die überwiegende Mehrheit derivierter Adjektive enthält ein Wortbildungselement (Grad 1). Dazu zählen Bildungen wie **sport-lich** (JAN 1;10), **lust-ig** (LEN 2;9), **mut-ig** (JAN 3;8), **narr-isch** (JAN 4;10), **zehn-t** (JAN 2;2), und vereinzelte Ableitungen wie **lang-sam** (JAN 2;2), **wunder-bar** (D09-KAB 4;7), **norm-al** (FSUFF, D19-VAS 3;1), **un-fair** (PREF, JAN 4;1), **sprech-end** (PPräs, JAN 2;10), **flieg-end** (PPräs, D27-JOZ 4;3) und **schuld** (KONV mor, D24-RAG 4;4).

Die Variation komplexer Adjektive mit Grad 2 nimmt in Jans CS von 2;0 bis 4;11 zu, während die Lemmafrequenzen komplexerer Bildungen im gleichen Zeitraum rückläufig sind. Dies zeigt sich ebenso in ihrer Verwendungshäufigkeit. Beispiele für derivierte Adjektive mit Komplexitätsgrad 2 sind *ur-lust-ig* (D17-TOB 4;8), *kräft-ig* (D04-SOS 4;5), *dri-tt-e* (D03-NIP 4;8) und das Perfektpartizip *verwuschel-t* (D22-LAS 4;3) und *ver-irr-t* (D13-JOP 4;8). Auch in Grad 2 ist das Suffix *-ig* das frequenteste Derivationsmuster. In CDS finden sich jedoch auch etliche Lexeme, die mit *-lich* abgeleitet werden, wie bei Jans Mutter in *künst-lich* (2;1), *häss-lich*, *gefähr-lich* (2;7), *natür-lich* (4;5) und *zusätz-lich* (4;9).

Komplexe Adjektivableitungen mit Komplexitätsgrad 3 sind entweder suffigiert wie in *sicht-bar* (JAN 2;0), *künst-ler-isch* (CDS D04-SOS 3;5), Perfektpartizipien (PP) wie in *ge-broch-en* (JAN 2;8), *er-kält-et* (D29-SAK 3;5) oder bei den INPUT-Kindern und in CDS präfigiert wie in *ur-schwier-ig* (D15-LUN 4;7), in dem Exzessiv *aller-er-st-er*⁴⁰ mit nicht-autonomer Basis *er-* (D07-NIG 4;2), *un-zu-stell-bar* (CDS JAN 2;11). Bildungen mit Komplexitätsgrad 4 sind bei Jan und den HSES-Kindern *vor-sicht-ig* (JAN 2;0, D23-SOK 2;11), *durch-sicht-ig* (D31-AND 4;6), und das präfigierte Adjektiv *un-sicht-bar* (JAN 2;6). Dies ist ebenso für CDS der Fall, wo *-ig*-Ableitungen in den gleichen Lexemen dominieren, neben vereinzelt Ableitungen auf *-lich* wie in *offen-sicht-lich* (D06-LAB 3;5), *ab-sicht-lich* (LEN 2;10), und präfigierten Bildungen wie *un-be-rechen-bar*⁴¹ (JAN 2;8), *un-vor-sicht-ig* (JAN 2;6) und *un-voll-ständ-ig* (D07-NIG 2;11).

Neologistische derivierte Adjektive kommen bei Jan in zwei Bildungen vor: mit 3;0 in **bahn-ig-es* und mit 4;10 in **ver-dämmel-t-e* (PP), bei Kathi in **teu-ig-er* ← *teuer*, bei den HSES-Kindern (2 Lem./2 Tok.) in **ein-ohr-ig* und **loch-ig-es* (D09-KAB 4;4). Lena bildet keine Neologismen.

⁴⁰ Irreguläre Stammmodifikation der Numerale *ein-*, etymologisch vgl. Kluge (2011, s. v. *eher*).

⁴¹ Vgl. *Rechen-schaft* vs. *rechn-en*: phonologische *e*-Schwa-Insertion im Auslaut und vor Konsonant.

8 Fehlerauswertung

8.1 Fehler ohne Fehler

Nachdem die spontansprachlichen Bildungen in dieser Arbeit untersucht werden, fließen auch umgangssprachliche Ausdrücke in die Analyse mit ein. Diese dem Standard nicht entsprechenden Lemmata werden als korrekt angesehen, bei der morphologischen Codierung mit ** markiert und in Folge nicht als Fehler ausgewertet. Hauptsächlich betrifft dies Verben, die die zweite und dritte Person Singular (3S, 2S) nicht nur durch Anfügung des Flexionsmorphems (-t, -st), sondern auch mit Umlaut bilden, der im umgangssprachlichen Ausdruck entfällt. Sowohl in CS als auch in CDS gibt es zahlreiche Beispiele wie jene in (7), (8) und (9):

(7) D05-FLB (4;8):

*CHI: aba [: aber] du **haltst** fuer mich **auf** .

%mor: CONJ|aber PRO|du **auf#V:07|halt-2S**** PREP|fuer PRO|ich-ACC .

(8) D12-JAS (4;9):

*CHI: und das # **fangt** auch mit k@l **an** .

%mor: CONJ|und PRO:dem|d-as **an#V:07|fang-3S**** ADV|auch PREP|mit lett|k .

(9) D06-LAB (4;8):

*PAR: dann **verrätst** dich .

%mor: ADV|dann **V:07|verrat-2S**** PRO:refl|du-ACC .

Die Verwendung eines Ablauts, wie in dem dialektalen Beispiel (10) dargestellt, ist dagegen nur in CDS und nur vereinzelt in Simplex-Verben zu finden:

(10) D09-KAB (4;4):

*PAR: owa [: aber] i(ch) **siech** [: seh] nix [: nichts] .

%mor: CONJ|aber PRO|ich **V:10|seh-1S**** PRO:indef|nichts .

Auch der Entfall des Partizipialpräfixes ge- gilt umgangssprachlich in vielen Fällen als korrekt, siehe Beispiel (11), und ist von jenen zu unterscheiden, die keinesfalls ge-los sein können, siehe Beispiele (12), (13) und (14):

(11) D28-MOH (3;0):

*CHI: +< hat da [: der] Digo **aufpasst** .

%mor: V:aux|hab-3S DET:art:def|d-er n:prop|Digo **auf#V:01|pass-PP**** .

8.2 Fehleranalyse

Tabelle 23 zeigt die Übersicht der Fehler in den Gesamtkorpora. Es werden die kindersprachlichen Fehlbildungen der Derivationen (Substantive, Verben, Adjektive) und Komposita (Substantive, Adjektive) für einen Gesamtüberblick aufgelistet.

Tabelle 23. Übersicht der Fehlerhäufigkeiten derivierter und zusammengesetzter Lexeme in Lemmas, Tokens und Prozenten in CS in allen Korpora im gesamten Zeitraum

Korpus	JAN	KAT	LEN	INPUT HSES	INPUT LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
N DERIV	155/1.108	37/68	36/183	183/829	106/600
Err. in N DERIV	2/2	2/3	3/3	5/5	5/8
% Err./N DERIV	1,29/0,18	5,41/4,41	8,33/1,64	2,73/0,60	4,72/0,83
V DERIV	388/1.715	99/217	132/433	556/1.737	409/1.342
Err. in V DERIV	29/48	10/11	29/48	30/39	49/71
% Err./V DERIV	7,47/2,80	10,10/5,07	21,97/11,09	5,40/2,25	11,98/5,29
ADJ DERIV	60/240	5/6	4/7	67/189	44/114
Err. in ADJ DERIV	0/0	1/1	0/0	2/2	5/5
% Err./ADJ DERIV	0,0/0,0	20,0/16,67	0,0/0,0	2,99/1,06	11,36/4,39
N KOMP	295/1.654	66/119	48/105	407/811	183/379
Err. in N KOMP	12/17	1/1	8/18	25/29	16/22
% Err./N KOMP	4,07/1,03	1,52/0,84	16,67/17,14	6,14/3,58	8,74/5,80
synth. KOMP	80/305	11/19	1/1	72/139	36/53
Err. in synth. KOMP	9/12	1/2	0/0	4/7	2/5
% Err./synth. KOMP	11,25/3,93	9,09/10,53	0,0/0,0	5,56/5,04	5,56/9,43
ADJ KOMP	14/16	1/1	0/0	13/22	5/8
Err. in ADJ KOMP	0/0	0/0	0/0	1/1	0/0
% Err./ADJ KOMP	0,0/0,0	0,0/0,00	0,0/0,0	7,69/4,55	0,0/0,0

Gesamt betrachtet haben die Kinder tendenziell am wenigsten Probleme mit Adjektivkomposita (ADJ KOMP), da Adjektivkomposita später als Substantive und Verben erworben werden und daher die Fähigkeit, Komposita zu bilden, schon weiter fortgeschritten ist. Die HSES-Kinder und Jan zeigen hier bereits mehrere Bildungen. Bei den HSES-Kindern, die in dieser Kategorie die prozentual größte Fehlerhäufigkeit in Relation zur Gesamtanzahl zusammengesetzter Adjektive aufweisen, handelt es sich um einen einzigen Fehler. Da Adjektivkomposita nach derivierten Adjektiven jene Kategorie mit den wenigsten Bildungen ist, ist die hohe prozentuale Fehlerrate bei den HSES-Kindern nicht repräsentativ. Die geringen

Lemma- und Tokenfrequenzen zeigen, dass die Kinder dieses Muster erst zu erwerben beginnen.

Nach Wortarten getrennt ergeben sich bei Jan und den INPUT-Kindern die höchsten Fehlerquoten in derivierten Verben in Lemmas und Tokens. Dabei ist anzumerken, dass es sich ebenfalls um die Kategorie mit den meisten Bildungen handelt. Im Bereich der Komposition überwiegen die Fehleranteile in synthetischen Komposita (synth. KOMP), als komplexeste Kategorie, bei allen Kindern in Tokens (für einen cross-linguistischen Vergleich siehe Dressler, Sommer-Lolei, et al. 2019). Jans fehleranfälliges Wortbildungsmuster in Lemmas ist über den gesamten Zeitraum gerechnet jenes der synthetischen Komposita, während die INPUT-Kinder mehr Fehler bei nominalen Komposita produzieren (siehe Tabelle 23). Wie bereits erwähnt, ist die prozentual hohe Fehlerrate bei den HSES-Kindern bei Adjektivkomposita dabei nicht repräsentativ, da es sich um eine einzelne fehlerhafte Bildung handelt.

Im direkten Vergleich der prozentualen Fehlerhäufigkeiten in Lemmas im INPUT-Korpus weisen die LSES-Kinder bei derivierten Substantiven (N DERIV) die geringste Fehlerdifferenz mit unter 2 % auf, während sich bei den derivierten Adjektiven und Verben die größten Differenzen zu den HSES-Kindern von durchschnittlich 7,5 % ergeben. Bei synthetischen Komposita finden sich keine Unterschiede in den Fehleranteilen in Lemmas zwischen den beiden Gruppen. In Tokens rangieren die Unterschiede der Fehleranteile zwischen 0,2 % (N DERIV) und 4,39 % (synth. KOMP).

Die LSES-Kinder sind im Bereich der morphologischen Komplexität weniger weit fortgeschritten und produzieren gegenüber den gleichaltrigen HSES-Kindern insgesamt eine niedrigere Anzahl an derivierten Lexemen und Komposita (siehe Entwicklungsverläufe in Abbildungen 10, 12, 25 und 27) weisen in ihren Bildungen allerdings mehr Fehler in Lemmas und Tokens auf (vgl. Tabelle 23). Dies zeigt sich in allen Kategorien, mit Ausnahme der Adjektivkomposita, wo die nur wenigen Lexeme (5 Lem./8 Tok.) fehlerfrei sind.

Im Folgenden werden die Fehlerhäufigkeiten bei derivierten Substantiven und Verben sowie bei synthetischen Komposita im gesamten Zeitraum in den beiden

Korpora JAN und INPUT auf Basis der verschiedenen Fehlercodes dargestellt, untersucht und anhand von Beispielen illustriert. Auf die wenigen Fehler in derivierten Adjektiven wird ebenfalls kurz eingegangen. Auf die Ausführung des Fehlers bei Adjektivkomposita wird verzichtet. Für eine genaue Beschreibung der Fehlercodierungen, wie sie in den folgenden Diagrammen Anwendung finden, siehe Anhang A2.2, Tabelle 43, S. 258.

8.2.1 Fehler bei synthetischen Komposita

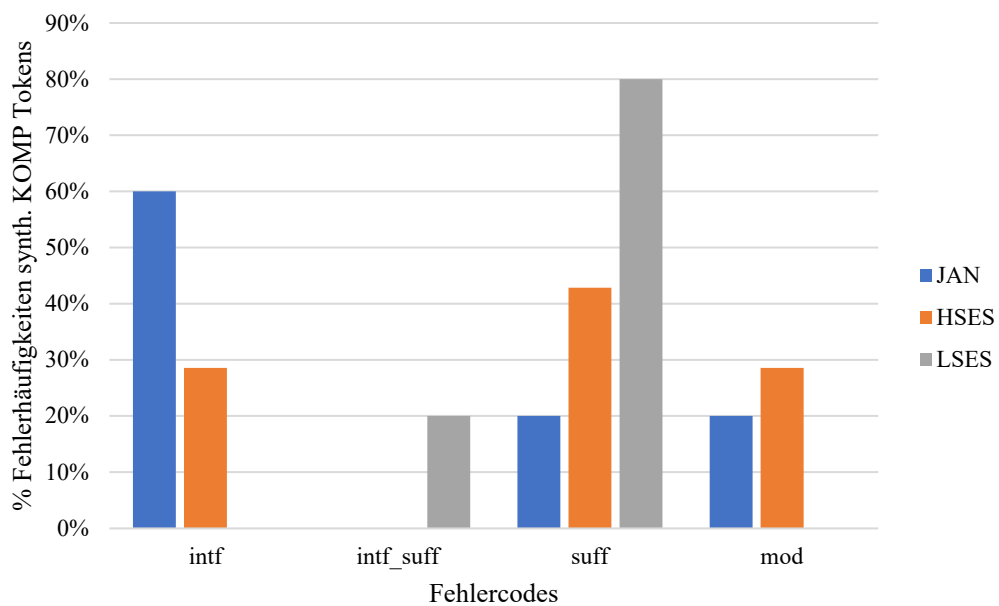


Abbildung 56. Fehlerhäufigkeiten je Fehlercode bei synthetischen Komposita in CS im JAN- und INPUT-Korpus⁴²

Alle Kinder produzieren, wie in Abbildung 56 dargestellt, bei synthetischen Komposita am häufigsten Fehler bei der Verwendung von Suffixen (suff) auch in Verbindung mit einem Interfixfehler bei den LSES-Kindern (intf_suff), sowie bei Interfixen (intf). Überdies kommt es bei synthetischen Komposita bei Jan und den HSES-Kindern zu Modifikation (mod). Darüber hinaus produzieren die Kinder Fehler bei Flexionssuffixen, die in diesem Zusammenhang nicht behandelt werden. Im Rahmen der Produktion synthetischer Komposita bilden die Kinder, falls nötig, immer den korrekten Umlaut. Für eine Übersicht der Fehlerhäufigkeiten in

⁴² Intf = Interfixfehler, intf_suff = Interfix- und Suffixfehler, suff = Suffixfehler, mod = Modifikation.

Lemmas, Tokens und Prozenten siehe Tabelle 23. Im Folgenden werden die Fehler mit Beispielen dargestellt.

Bei fehlerhaften Suffixen (suff) handelt es sich bei einem LSES-Kind um eine fehlerhafte Realisierung des Suffixes -er-ei, durch den Entfall des intervokalischen /r/ in *daumen+quetsch-e∅ei* (D24-RAG 4;7). Überdies kommt es bei den HSES-Kindern mehrfach zu Suffixsubstitutionen in synthetischen Komposita, wie in *super+kleb-(e)n*⁴³ statt *Superkleber* (D09-KAB 4;4), sowie zu einer Ersetzung des Diminutivsuffixes -erl durch -ler in *stab-ler+samml-er*, *stab-ler+samml-er* statt *Staberlsammler* (JAN 2;7), vermutlich hervorgerufen durch die Antizipation des Kopfes.

Interfixfehler (intf) bedeuten hauptsächlich den Entfall des Interfixes -n- wie z. B. in *apotheker-∅-verkäufer-in* (D04-SOS 4;9), *kassette-∅-aufnahme* (JAN 2;10) und *zwerge-∅-spiel* (JAN 2;9) oder bei einem LSES-Kind in Kombination mit dem Entfall des Derivationsuffixes -er in *sch(r)aube-∅-zieher-∅* statt *Schraubenzieher* (intf_suff, D08-FAJ 3;4). In wenigen Fällen wird ein Interfix hinzugefügt, wie in *spidy-s-spiel* (D27-JOZ 4;3) und in *gras-n-mäher* (JAN 1;9), wobei letzteres Beispiel auf eine Vermischung der beiden Substantive *Gras* und *Rasen* zurückzuführen sein dürfte.

Es ist daher feststellbar, dass den Kindern Interfixe und Suffixe bei der Bildung synthetischer Komposita bis 4;11 tendenziell Probleme bereiten, insofern als es häufig zu einem Interfixentfall des Interfixes -n- kommt, und Suffixe durch andere ersetzt werden.

⁴³ In der Äußerung (4;4): *CHI: wieso geht's [: geht es] nicht mitn [: mit dem] superklebn@m?

8.2.2 Fehler bei derivierten Substantiven

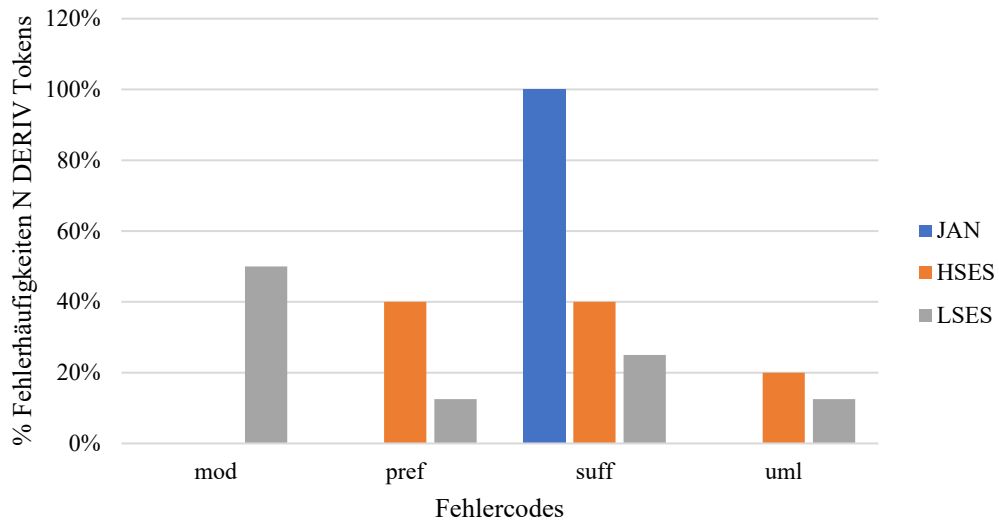


Abbildung 57. Fehlerhäufigkeiten je Fehlercode bei derivierten Substantiven in CS im JAN- und INPUT-Korpus

Die in Abbildung 57 dargestellten prozentualen Fehlerhäufigkeiten bei derivierten Substantiven zeigen die Verteilung der Fehler je Fehlercode in Relation zur Gesamtfehleranzahl. Für die absoluten Fehlerhäufigkeiten in Lemmas und Tokens siehe Tabelle 23. Die LSES-Kinder produzieren im Vergleich zu den HSES-Kindern und Jan mehr unterschiedliche Fehler. Es handelt sich dabei um Modifikationen (mod) des Wortstamms der Basis in der Bildung lüg-er statt Lügner (D24-RAG 4;7).

Im Vergleich zu den synthetischen Komposita (vgl. Kap. 8.2.1) kommt es in nominalen Derivationen zu einem Entfall des Umlauts (uml), wie in dem Kollektivum ge-busch statt Gebüsch (D10-LER 4;7) und in knoch-el statt Knöchel (D03-NIP 4;5).

Von den Suffixfehlern (suff) sind zur Gänze Diminutivsuffixe betroffen. Dabei handelt es sich entweder um einen Entfall des Suffixes, wie in zuck- \emptyset ⁴⁴ statt Zuckerl (D07-NIG 4;5), um den Entfall des auslautenden /n/ in tür-che \emptyset ⁴⁵ statt Türchen (D10-LER 3;4), wobei es sich hier auch um die Tilgung des vermeintlichen -n-Pluralmorphems handeln kann zur Bildung des Singulars, um einen Entfall des

⁴⁴ In der Äußerung (4;5): *CHI: dann darf ich ein zuck@m.

⁴⁵ In der Äußerung (3;4): *CHI: und da is(t) ein [/] ein [/] ein tuerche@m.

anlautenden /d/ in männ-erl statt Mannderl (JAN 2;10), in stern-erl statt Sternderl (D09-KAB 3;4), oder um Kontamination in schwein-chi statt Schweindi (D21-PAS 4;6), wo die beiden Suffixe -chen und -di in Konkurrenz zueinander stehen. Nur bei Jan kommt es zu der Ersetzung des Diminutivsuffixes -erl durch -chen in straß-chen statt Straßerl (JAN 4;1).

Eine Präfixersetzung oder -hinzufügung (pref) kann zu einem existenten Wort, wie in be-steck statt ver-steck (D17-TOB 3;0) führen, oder auch zu einem inexistenten Wort, einem Neologismus, wie in ge-rot statt rot (D15-LUN 4;3) und ver-schwemm-ung statt über-schwemm-ung (D04-SOS 4;5).

Vereinzelt kommt es zu einer Anbringung eines Derivationsuffixes an Simplizia, wie in holz-er statt Holz (D18-ALM 3;0), (ein) Mensch-e statt Mensch (D21-PAS 3;6), auf Basis einer möglichen denominalen -er-Suffigierung (vgl. Tät-er ← Tat), was zeigt, dass einige Kinder dieses Muster produktiv anwenden.

8.2.3 Fehler bei abgeleiteten Verben

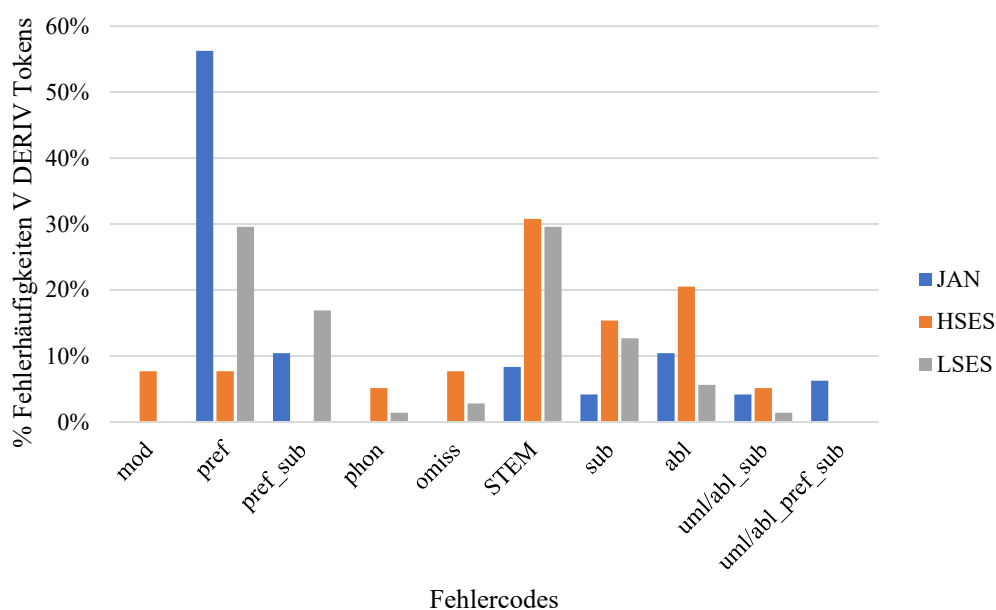


Abbildung 58. Fehlerhäufigkeiten je Fehlercode bei derivierten Verben in CS im JAN- und INPUT-Korpus⁴⁶

⁴⁶ Mod = Modifikation, pref = Präfixfehler, pref_sub = Präfixfehler und Substitution, phon = phonologischer Fehler, ommiss = *omission*, Entfall, STEM = Nur Wortstamm, sub = Substitution, abl = Ablautfehler, uml/abl_sub = Umlaut-/Ablautfehler und Substitution, uml/abl_pref_sub = Umlaut-/Ablautfehler, Präfixfehler und Substitution.

Die prozentualen Fehlerhäufigkeiten bei abgeleiteten Verben sind in Abbildung 58 abgebildet. Dabei wird die Fehlerverteilung nach Fehlercode im Vergleich zur Gesamtfehleranzahl dargestellt. Die Fehleranfälligkeit im Bereich derivierter Verben ist bei Präfixen (pref) bei Jan und den LSES-Kindern am höchsten, bei letzterer Gruppe zeigt sich diese gleich hoch mit der Verwendung des komplexen Wortstamms (STEM), was sich auch bei den HSES-Kindern als häufiges Fehlermuster zeigt. Sowohl die INPUT-Kinder als auch Jan bilden bei abgeleiteten Verben zahlreiche Umlaut- und Ablautfehler (uml/abl) und Ersetzungen (sub) auch in Kombination (uml/abl_sub). Bei Jan kommen Umlaut- und Ablautfehler auch zusammen mit Präfixfehlern und Ersetzungen vor (uml/abl_pref_sub). Zu Modifikationen (mod) kommt es ausschließlich bei den HSES-Kindern. Phonologische Abweichungen (phon), fehlende Endungen bei Perfektpartizipien (omiss) und die Kombination zuvor genannter Fehlercodes (pref_sub) kommen nur in geringem Ausmaß vor.

Von den verschiedenen Fehlern sind in der überwiegenden Mehrheit Partizipialbildungen betroffen, da es einerseits zu einem gänzlichen Entfall der Endung (-(-e)t, -en, omiss) oder des Partizipialpräfixes (pref) ge- kommen kann, oder zu der fehlerhaften Verwendung einer Endung (sub). Im Folgenden wird auf die verschiedenen Fehler eingegangen und diese anhand von Beispielen veranschaulicht.

Im Gegensatz zu den bereits erwähnten umgangssprachlich korrekten ge-losen Partizipialbildungen, vgl. Beispiel (11), sind die folgenden Bildungen fehlerhaft, wenn das Partizipialpräfix entfällt:

(12) D26-JUB (4;9):

*CHI: der [?] hat da viele spiele **runterladen@m** .

%mor: PRO:dem|d-er V:aux|hab-3S ADV|da DET:qn|viel-e N:03:n|spiel-PL
runter#V:08|lad-PP*m .

(13) D05-FLB (4;8):

*CHI: Lamibam [?] is(t) runtafalln [: **runterfallen@m**] .

%mor: n:prop|Lamibam V:aux|sein-3S runter#V:07|fall-PP*m .

(14) D08-FAJ (3;4)

*CHI: jetzt(t) **zumacht@m** .

%mor: ADV|jetzt **zu#V:01|mach-PP*m** .

Dabei kommt es neben dem Entfall des Präfixes häufig auch zu einer Substitution der Endung (pref_sub), wobei die bei starken Verben nötige Endung -en durch die produktive Endung -t schwacher Verben ersetzt wird, vgl. Beispiele (15) und (16):

(15) D24-RAG (3;0):

*CHI: ha i@g [: ich] ha@p **runterfallt@m** [=! zu PA2] .
%mor: INTERJ|ha PRO|ich phon|ha **runter#V:07|fall-PP*m** .

(16) D08-FAJ (3;4):

*CHI: suma@c titind@g [: christkind] **tommt@m** hab(e) .
%mor: chi|suma N:07:n|christ+kind **V:X|bekomm-PP*m** V:aux|hab-1S .

In einigen Fällen wird das zuvor genannte Fehlermuster um eine fehlende Ablautung (uml/abl_pref_sub), wie in *runter#-∅-spring-t* statt *runtergesprungen* (JAN 2;1) oder eine inkorrekte Umlautung in *rerunte-∅-läuf-t* statt *heruntergelaufen* (JAN 2;5) ergänzt. In beiden Beispielen werden unproduktive Bildungsmuster starker Verben durch das produktive Muster schwacher Verben ersetzt.

Allerdings gibt es auch PP-Bildungen, wo es zu einem Entfall (omiss) oder einer Substitution (sub) der Endung -en unter Beibehaltung des Partizipialpräfixes kommt, wie in *runter#ge-fall-∅* statt *runtergefallen* (D13-JOP 3;0), *rein#ge-geb-∅* statt *reingegeben* (D31-AND 3;7), *ab#ge-riss-∅* statt *abgerissen* (D13-JOP 3;3), *auf#ge-blas-t* statt *aufgeblasen* (JAN 2;6), wieder unter Verwendung des produktiven Bildungsmusters schwacher Verben, oder mit zusätzlich fehlender Ab- oder Umlautung (uml/abl_sub) in:

(17) D03-NIP (3;3):

*CHI: mama ich hab(e) meinen son@g [: schon] **reingesieht@m** !
%mor: N:01:f|mama PRO|ich V:aux|hab-1S PRO:poss|mein-en ADV|schon
rein#V:06|schieb-PP*m !

(18) D05-FLB (4;4):

*CHI: da hab(e) ich was reingeschreibt@m .
%mor: ADV|da V:aux|hab-1S PRO|ich PRO:indef|w-as **rein#V:04|schreib-PP*m** .

Dass manche Kinder ge- auch in Simplizia als Präfix wahrnehmen, beweist folgendes Beispiel, in dem das vermeintliche Präfix entfällt:

(19) D06-LAB (3;5):

*CHI: wir habe@m beide **wonnen@m** .
%mor: PRO|wir V:aux|hab-1P*m PRO:qn|beid-e **V:14|gewinn-PP*m** .

Das Kind lässt ge in gewinnen generell entfallen, wie in \emptyset winn-e ‚gewinn-1S‘, \emptyset winn-st ‚gewinn-2S‘, mit und ohne PP-Suffix: \emptyset wonn- \emptyset , \emptyset wonn-en ‚gewinn-PP‘, wie auch ver in: \emptyset lor-(e)n ‚verlier- PP‘. Es ist unwahrscheinlich, dass es sich bei diesen Bildungen um rein phonologische Aphäresen handelt, da Aphäresen in diesem Alter nur mehr sehr selten vorkommen (vgl. Rainer 2010: 172) und das Kind ge und ver sonst nicht entfallen lässt. Zu einem tatsächlichen Entfall des Präfixes zer- kommt es in einem abgeleiteten Verb:

(20) D07-NIG (4;2):

CHI: das die [] heu **kleinert@m** .

%mor: PRO:rel[d-as DET:art:def[d-ie [*] N:sgt:n]heu **V:01|zerkleiner*m-3S** .

Das LSES-Kind D17-TOB (3;0) ersetzt konsequent das Präfix ver- durch bi- in dem Lexem verstecken: bi-steck-e, bi-steck-st, bi-steck-t, bi-steck(e)n, bi-steck ‚versteck.STEM‘.

Alle Kinder verwenden bei Partikelverben häufiger nur den Wortstamm (STEM), wie in herum#häng- \emptyset (D29-SAK 3;5), her#komm- \emptyset (D13-JOP 3;3), an#schau- \emptyset (JAN 2;2) und produzieren bei diesem verbalen Bildungsmuster überdies fehlende oder inkorrekte Ablaute, wie in den Perfektpartizipien rauf#ge-peck-t statt raufgepickt (D06-LAB 3;4) und ver-brenn-t statt verbrannt (D23-SOK 3;3), wobei letzteres Beispiel auch die umgangssprachlich verwendete Form des Partizips darstellt.

Bei den HSES-Kindern kommt es zu einer phonologischen (phon) Dissimilation des Laterals in kling-e statt klingl-e ‚klingl-1S‘ (D16-LEP 3;1) und zu Wortstammmodifikationen (mod) in reg-t statt regn-et ‚regn-3S‘ (D29-SAK 3;3) und in dem Partikelverb über#hab-st statt über#ha-st ‚über#hab-2S‘ (D20-LUD 4;6).

Generell lässt sich feststellen, dass die Kinder bis 4;11 bei komplexen Verben gerne den Wortstamm verwenden und die meisten Probleme bei der Bildung von Perfektpartizipien starker Verben haben. Dabei haben sie Schwierigkeiten mit dem Partizipialpräfix ge-, mit der korrekten Endung und mit einem etwaigen Ablaut.

8.2.4 Fehler bei derivierten Adjektiven

Aufgrund der geringen Anzahl derivierter Adjektive in CS bis zu einem Alter von 4;11 (vgl. Tabelle 23) kommt es verhältnismäßig ebenfalls nur zu einigen wenigen Fehlern.

Vereinzelte kommen es zu Tilgungen am Wortanfang, wie in wachsenen statt erwachsenen (D18-ALM 3;0), durch den Entfall des Präfixes er-, oder zur Auslassung von ge in dem präfigierten Adjektiv ur-fähr-lig-e statt urgefährliche (D19-VAS 4;5), durch die Kontamination der beiden Suffixe -ig und -lich.

Zumeist handelt es sich um Hinzufügungen, wie etwa um die Einfügung von -ern- in *ge-spitz-ern-e statt gespitzte (D10-LER 4;4), die Hinzufügung von -i zur Verniedlichung von fertig: *fertigi (D23-SOK 3;3), oder die zusätzliche Verwendung des Suffixes -ig in der Superlativbildung *aller-aller-kürz-ig-ste statt allerallerkürzeste (D03-NIP 3;3), sowie um -ig in dem Komparativ *teu-ig-er statt teurer (KAT 2;8).

9 Elizitationsexperiment

Um festzustellen, ob ein Kind ein bestimmtes Derivationsmuster erworben und ein höheres Sprachbewusstsein entwickelt hat, hat die Verfasserin ein Experiment zur Elizitation nominaler deverbaler Derivationen entwickelt. Dabei wird das produktive Anwenden der Ableitungssuffixe -er und -ung getestet, inklusive der movierten Form -er-in. Alle zu elizitierenden Derivationen zählen zu den wortartverändernden Ableitungen. Im Detail handelt es sich um folgende Nominalsuffixe und Kategorien: -er und -er-in für Nomina Agentis, -er für Nomina Instrumenti und -ung für Nomina Acti und/oder Nomina Actionis.

Der Fokus auf deverbale nominale Suffigierung beruht darauf, dass bestimmte Wortbildungsmuster, wie die deverbale Ableitung mit dem Suffix -er, bereits sehr früh in der Kindersprache auftreten, während die meist abstrakten -ung-Derivationen erst wesentlich später im Spracherwerb produktiv in Erscheinung treten. Das Experiment soll daher zeigen, inwieweit Kinder im Alter von durchschnittlich 8 Jahren diese Suffigierungen im direkten Vergleich spontan meistern können. Überdies sollte die für das Experiment nötige Durchführungsdauer 15 Minuten nicht übersteigen, weshalb eine Einschränkung ohnehin nötig war.

9.1 Forschungsziel und Forschungsfragen

Dieses Experiment ist Teil einer Folgestudie zum INPUT-Projekt und soll die weitere sprachliche und morphologische Entwicklung der durchschnittlich 8-jährigen Kinder im Bereich deverbaler nominaler Ableitungen feststellen.

Primär geht es um die Beantwortung folgender Forschungsfragen: Wie konsequent wenden die Kinder im Alter von 8 Jahren die getesteten Ableitungsmuster an? Zeigen sich Unterschiede, die sich auf den sozioökonomischen Status (SES) zurückführen lassen? Welche Fehler sind die häufigsten? Lassen sich Korrelationen feststellen zu den vier Jahre zuvor erhobenen Spontansprachdaten oder zu Inputfrequenzen und Medienkonsum (YISF: YouTube, Internetnutzung, Smartphone, Fernsehen; HBR: Hörbuch und Radio; Büchervorlesen)?

9.2 Methode

Bei dem off-line Experiment handelt es sich um einen Satzvervollständigungstest zur Elizitation deverbaler nominaler Ableitungen, die den semantischen Wortbildungskategorien Nomina Agentis, Nomina Instrumenti und Nomina Acti/Actionis zugeordnet werden können (siehe dazu Tabelle 26). Es erfolgt keine visuelle Präsentation der beschriebenen Situationen oder Stimuli, sondern ausschließlich die mündliche Anleitung durch die das Experiment durchführende Person.

Die Experimentatorinnen wurden darauf hingewiesen, die Testsätze deutlich vorzulesen und die in den Experimentalblättern kursiv geschriebenen Wörter zu betonen. Außerdem sollten sie jede Antwort der Kinder schriftlich festhalten. Der vorgegebene Ablauf und die Zeitvorgabe von maximal 12 Minuten wurden von den Experimentatorinnen genau eingehalten. Reaktionszeiten konnten, wie meist bei off-line Experimenten, keine gemessen werden.

Das Elizitationsexperiment wurde mit dem Einverständnis der Eltern auf Video und Audio aufgenommen.

Pilotexperiment

Dem eigentlichen Test ging ein Pilotexperiment voraus, um die Eignung der Stimuli und Targets zu überprüfen und in weiterer Folge gegebenenfalls anzupassen. Die Pilotstudie, die im Mai 2017 mit 5 Schulkindern (2 Kinder im Alter von 7 Jahren und 3 Kinder von 8 Jahren = Durchschnittsalter 7;6) in der Schule, einem Vorschulkind (6 Jahre) und einem Kindergartenkind (4 Jahre), jeweils zuhause durchgeführt wurde, umfasste 25 der 36 Stimuli (siehe Kap. 9.4) und dauerte pro Teilnehmer unter fünf Minuten. 6 der 7 Kinder vervollständigten bereits beim zweiten Übungssitem (Spielerin) den Satz selbständig, was deutlich zeigt, dass die Bildung der movierten Form im Bereich der Nomina Agentis im Regelfall für Kinder jeden Alters problemlos möglich ist. Tabelle 24 zeigt die Teilnehmer des Pilotexperiments und gibt eine übersichtliche Darstellung der korrekten, semikorrekten und inkorrekten Derivationen, sowie etwaiger fehlender Antworten und die individuelle Fehleranzahl.

Tabelle 24. Übersicht der Teilnehmer des Pilotexperimentes und zusammenfassende Ergebnisse

Kürzel	Alter	korrr.	semikorr.	inkorr.	autokorr.	keine Antw.	Fehleranzahl
ADA (w/H)	8	19	1	5	0	0	6
NOR (w/H)	7	18	1	3	0	3	4
REO (m/H)	8	18	2	5	0	0	7
LOU (m/H)	8	17	1	6	0	1	7
MEI (w/H)	7	10	1	5	0	8	7
NIK (m/L)	4	6	2	12	0	5	14
EMM (w/L)	6	11	1	8	0	5	10
Gesamt		99	9	44	0	22	55

Die Auswertung des Pilotexperimentes hat ergeben, dass die Anzahl der korrekten Derivationen mit zunehmendem Alter steigt, während die Fehlerhäufigkeit oder die Tendenz keine Antwort zu geben deutlich sinkt (vgl. Tabelle 24). Außerdem kam es zu keinen spontanen Selbstkorrekturen (autokorr.).

Im Rahmen der Pilotstudie haben sich folgende für das Experiment wesentlichen Erkenntnisse ergeben: a) Die Ergebnisse (Antworten) waren gleich gut, unabhängig davon, ob das Stimulusverb finit oder infinit war, die Sprachverarbeitung blieb daher von der Flexion des Verbs unberührt, b) Bei Nichtbeantwortung bzw. Nichtreaktion des Kindes war eine Wiederholung der Frage ebenso wenig erfolgversprechend. Daher werden Sätze, auf die keine Reaktion erfolgt, im tatsächlichen Experiment nicht wiederholt, c) Die Kinder achten darauf, ob die Experimentatorin etwas mitschreibt. Setzt diese beispielsweise zu viele Haken für korrekte Antworten und schreibt dann bei einer falschen Antwort plötzlich etwas mit, kann dies dazu führen, dass die Kinder unsicher werden, weil sie das Gefühl haben, etwas falsch gemacht zu haben. Deshalb müssen die Experimentatorinnen alle Antworten im Verlauf des Tests mitschreiben, um eine etwaige Verunsicherung der Kinder zu vermeiden.

9.3 Teilnehmer

An dem Experiment nahmen insgesamt 26 Kinder, zwischen 8 und 9 Jahren, mit unterschiedlichem sozioökonomischen und sprachlichen Hintergrund, teil. Alle Kinder sind in ihrer körperlichen, sprachlichen und geistigen Entwicklung unauffällig. Es handelt sich um 16 Kinder mit L1 Österreichisches Standarddeutsch

im Alter von 7;9 bis 8;11 (Durchschnittsalter 8;3), sowie um 10 bilinguale Kinder, die mit Erstsprache Türkisch und Zweitsprache Österreichisches Standarddeutsch in Wien aufwachsen, im Alter von 7;10 bis 9;4 (Durchschnittsalter 8;9). Im Durchschnitt sind die bilingualen Kinder zum Zeitpunkt des Experiments 6 Monate älter. Alle in Tabelle 25 gelisteten Kinder und Teilnehmer des Elizitations-experiments waren bereits Teilnehmer des INPUT-Projekts und sind daher mit jenen in Tabelle 3 (für L1 Deutsch) identisch. Es handelt sich um 11 Mädchen und 5 Buben mit L1 Deutsch, und 8 Mädchen und 2 Buben mit L2 Deutsch/L1 Türkisch, nach dem sozioökonomischen Status um 17 HSES-Kinder (L1 Deutsch: 7 w & 4 m, L2 Deutsch: 5 w & 1 m) und 9 LSES-Kinder (L1 Deutsch: 4 w & 1 m, L2 Deutsch: 3 w & 1 m). In der überwiegenden Mehrheit sind die Teilnehmer des Experiments weiblichen Geschlechts (Gesamt 19 w vs. 7 m, L1 Deutsch: 11 w vs. 5 m).

Tabelle 25. Teilnehmer Derivationsexperiment

Kind (L1 Deutsch)	Alter	Kind (L2 Deutsch)	Alter
D02-SIJ (m/H)	8;1	T13-SUB (w/H)	9;1
D03-NIP (m/H)	8;0	T19-MEE (m/H)	9;3
D04-SOS (w/H)	8;2	T20-CAE (w/H)	9;1
D07-NIG (w/H)	7;9	T22-EDK (w/H)	8;8
D09-KAB (w/H)	7;9	T25-ELS (w/H)	8;2
D13-JOP (m/H)	8;7	T29-LOD (w/H)	8;1
D20-LUD (w/H)	7;9	T01-MOK (m/L)	9;4
D23-SOK (w/H)	8;6	T02-TUS (w/L)	8;6
D28-MOH (m/H)	8;5	T07-ZEO (w/L)	9;0
D29-SAK (w/H)	8;7	T10-BEO (w/L)	7;10
D31-AND (w/H)	8;3		
D06-LAB (w/L)	8;2		
D12-JAS (w/L)	7;9		
D18-ALM (w/L)	8;7		
D19-VAS (w/L)	8;7		
D21-PAS (m/L)	8;11		

Da es in dieser Arbeit nicht um einen Vergleich zwischen L1- und L2-Erwerb geht, schließt die Verfasserin die 10 bilingualen Kinder von ihrer Analyse aus. Allerdings wäre auch im Fall der bilingualen Kinder eine Gegenüberstellung der Ergebnisse des Elizitationsexperiments zu den zuvor erhobenen spontansprachlichen Daten, Inputmenge und Medienkonsum mit einem zusätzlichen Fokus auf SES-Unterschiede wünschenswert.

9.4 Stimuli

Bei den Stimuli handelt es sich um frequente deutsche Verben, die in CDS vorkommen und den Kindern daher meist bekannt sein sollten. Diese Verben sollen durch Anfügung eines Derivationsuffixes nominal abgeleitet werden. Sämtliche Stimuli wurden an den Wortschatz von Kindern bis 10 Jahre angepasst, nach ihrem Umlautungsverhalten ausbalanciert und nach ihrem Vorkommen in unseren CDS-Langzeitkorpora der Arbeitsgruppe für Komparative Psycholinguistik der Universität Wien ausgewählt. Die Zielwörter (Targets) sind deverbale nominale Ableitungen und zählen demzufolge zu den wortartverändernden Derivaten. Für jede semantische Wortbildungskategorie (Nomina Agentis [NAG], Nomina Instrumenti [NIN] und Nomina Acti/Actionis [NAC]) kommen jeweils 12 Stimuli zum Einsatz. In Summe ergibt dies 36 Testsätze. Tabelle 26 gibt eine Übersicht über die Stimuli und Targets, die semantische Kategorie, der das Derivat zugeordnet ist und ob im Zuge der Derivation eine Umlautung des Stammvokals nötig ist. Zu einer Umlautung (+) kommt es z. B. in *lauf-* → *Läufer*, nicht aber (-) in *üb-* → *Übung*. 23 Stimuli beinhalten einen umlautfähigen Vokal (a, au, e, o), davon jedoch 18 Stimuli, deren Derivate nicht umgelaute werden. Bei einer umzulautenden Verbbasis (*lauf-*) entfällt in der Umgangssprache in der zweiten und dritten Person Singular der Umlaut: **lauf-st* ‚lauf-2S‘ und **lauf-t* ‚lauf-3S‘, nicht jedoch in dem Zielwort *Läuf-er* (NAG). Von den 36 verbalen Basen sind 29 einfach und 7 komplex: 6 enthalten ein untrennbares Präfix (in Tabelle 26 grau hinterlegt): *ver-*, *ent-*, oder *be-* und eine enthält das trennbare Präfix (= Partikel, in Tabelle 26 blau hinterlegt): *ab#*). Zwei Verbbasen sind pseudo-deriviert (*radier-*, *verletz-*).

In der Gruppe der Nomina Agentis muss die Hälfte der Zielwörter (6 der 12 Items) im Zuge der Suffigierung zusätzlich umgelaute bzw. in einem Fall (*sing-*) abgelautet werden. Die andere Hälfte der Zielwörter wird ausschließlich durch die Anfügung des jeweiligen Ableitungsmorphems gebildet. Im Bereich der Nomina Instrumenti, als auch der Nomina Acti/Actionis kommt es bei der Derivation zu keiner Veränderung bzw. Um- oder Ablautung des Stammvokals.

Tabelle 26. Übersicht der Stimuli und Targets nach Ableitungssuffix, Um-/Ablaut, Basis und semantischer Wortbildungskategorie

Ableitungssuffix	Um-/Ablaut	Basis → Kategorie	Stimulus	Target
-er	ohne (-)	V → NAG	sammeln	Sammler
	ohne (-)		arbeiten	Arbeiter
	ohne (-)		zaubern	Zauberer
	mit (+)		laufen	Läufer
	mit (+)		jagen	Jäger
	mit (+)		backen	Bäcker
-er-in	ohne (-)	V → NAG	schwimmen	Schwimmerin
	ohne (-)		lehren	Lehrerin
	ohne (-)		malen	Malerin
	mit (+)		verkaufen	Verkäuferin
	mit (+)		tanzen	Tänzerin
	mit (+)		singen	Sängerin
-er	ohne (-)	V → NIN	bohren	Bohrer
	ohne (-)		öffnen	Öffner
	ohne (-)		mähen	Mäher
	ohne (-)		mixen	Mixer
	ohne (-)		radieren	Radierer
	ohne (-)		kleben	Kleber
	ohne (-)		lochen	Locher
	ohne (-)		abstreifen	Abstreifer
	ohne (-)		wecken	Wecker
	ohne (-)		toasten	Toaster
	ohne (-)		entsaften	Entsafter
	ohne (-)		rechnen	Rechner
	-ung		ohne (-)	V → NAC
ohne (-)		retten	Rettung	
ohne (-)		landen	Landung	
ohne (-)		verkühlen	Verkühlung	
ohne (-)		begleiten	Begleitung	
ohne (-)		verletzen	Verletzung	
ohne (-)		atmen	Atmung	
ohne (-)		befreien	Befreiung	
ohne (-)		verbessern	Verbesserung	
ohne (-)		impfen	Impfung	
ohne (-)		prüfen	Prüfung	
ohne (-)		üben	Übung	

9.5 Testdurchführung

Die Teilnehmer hören den ersten Satz (auditive, rezeptive Sprachverarbeitung) und sollen den darauffolgenden, zweiten Satz mit der korrekten Ableitung vervollständigen (elizitierte Sprachproduktion). Wie bereits erwähnt, zielt der Test auf das produktive Anwenden bereits erworbener morphologischer Regeln und Strukturen ab. Es erfolgt keine zusätzliche visuelle Präsentation der Testsätze bzw. Satzinhalte.

Das Experiment wurde mit einem Audioaufnahmegerät aufgezeichnet und die Zielwörter, also der Output der Kinder, überdies auf einem Experimentalblatt dokumentiert (siehe Anhang A10.1 Abbildungen 86 und 87). Das Elizitationsexperiment wurde mit den L1-Kindern von drei Experimentatorinnen bei den Kindern zuhause durchgeführt. Jede Antwort musste von der Experimentatorin mitgeschrieben werden, um eine Verunsicherung der Kinder zu vermeiden.

Ablauf des Experiments

Zunächst wird das Kind herzlich begrüßt und zu dem Spiel eingeladen. Das Experiment wird in einem weiteren Schritt mit 4 Übungsbeispielen (A–D), die in Tabelle 27 präsentiert werden, begonnen (je Wortbildungskategorie und Suffix jeweils ein Item). Die Übungssätze können dabei so oft wiederholt werden, bis sich die Experimentatorin sicher ist, dass die Vorgehensweise verstanden wurde. Danach beginnt der eigentliche Testdurchlauf, der pro Teilnehmer, inkl. Übungsitems, in etwa 10 Minuten dauert.

Tabelle 27. Übungssätze zur Vorbereitung auf das Elizitationsexperiment

Item	Einleitender Satz	Zu vervollständigender Satz	Zielwort
A	Der Mann möchte gerne fahren.	Er ist ein toller ...?	Fahr-er
B	Die Frau spielt sehr gerne.	Sie ist eine gute ...?	Spiel-er-in
C	Der Mann möchte den Bleistift spitzen.	Dazu braucht er einen ...?	Spitz-er
D	Die tanzende Frau dreht sich.	Sie macht eine schwungvolle ...?	Dreh-ung

Die darauffolgenden 36 Testsätze werden mit jedem Teilnehmer in derselben Reihenfolge durchgenommen (siehe Anhang A10.1 Abbildungen 86 und 87), dabei wartet die Experimentatorin nach den Sätzen in etwa 10 Sekunden auf eine Antwort

des Kindes. Erfolgt keine Antwort, setzt die Experimentatorin den Test mit dem nächsten Satz fort. Nach der letzten Antwort bedankt sich die Experimentatorin bei dem Kind für sein großartiges Mitmachen.

9.6 Datenauswertung

Die erhobenen Daten wurden anhand der Experimentalblätter (siehe Anhang A10.1 Abbildungen 86 und 87) und der Audioaufnahmen in ein Exceldatenblatt übertragen. Dabei wurden die einzelnen Teilnehmer einer Fallnummer zugeordnet und mit dem identischen, individuellen Kürzel des INPUT-Projekts anonymisiert, um eine exakte Zuordnung der Testdaten zu den Daten der Spontansprache zu gewährleisten. Neben der durchführenden Experimentatorin und dem Datum der Testdurchführung wurden ebenso das Geschlecht des Kindes, sein Alter, seine Erstsprache und seine SES-Zugehörigkeit erfasst.

Jede Antwort der Kinder wurde gemäß dem Schema in Anhang A10.2 Tabelle 63 eingetragen und akkurat erfasst, um eine Auswertung mittels Pivot-Tabellen in Excel und statistische Berechnungen in SPSS zu ermöglichen.

Ausgewertet wurden sämtliche Antworten der Kinder hinsichtlich folgender Faktoren:

- a) Antwort vs. keine Antwort
- b) korrekte vs. semikorrekte vs. inkorrekte Antwort
- c) korrektes vs. inkorrektes Suffix
- d) korrekter vs. inkorrekt Um-/Ablaut
- e) korrekte vs. inkorrekte Basis
- f) Selbstkorrektur
- g) Sozioökonomischer Status, Geschlecht

9.7 Fehlercodierung

Inkorrekten und semikorrekten Derivationen wurden bestimmte Fehlercodes zugeordnet, um zu sehen, welche Fehler die Kinder bei den Ableitungen spontan produzieren.

Tabelle 28 listet die entsprechenden Fehlercodierungen und -beschreibungen mit Beispielen auf, die für die Analyse der Fehler entscheidend sind. Dabei wurde in besonderem Ausmaß auf eine ausreichend detaillierte Codierung von Fehlern in Bezug auf die verwendete Ableitungsbasis (Base), das Derivationsuffix (Suff) und die Verwendung des Umlauts (Uml) geachtet.

Tabelle 28. Fehlercodierung für semikorrekte und inkorrekte Derivationen mit Beispielen

Fehlercode	Fehlerbeschreibung	Beispiel Basis → Output
0	kein Fehler, Zielwort korrekt abgeleitet	üb- → Übung
1	anderes Wort, keine Ableitung, kein Kompositum, semantischer Zusammenhang gegeben	arbeit- → Mann
2	anderes Wort, keine Ableitung, kein Kompositum, kein semantischer Zusammenhang gegeben	öffn- → Kasten
3	Kompositum statt Ableitung (+/-Base, -Deriv, +Komp)	kleb- → Kleb+e+band (-e = Interfix)
4	synthetisches Kompositum (+/-Base, +/-Suff, +Deriv, +Komp)	mäh- → Rasen +mäher entsaft- → Zitronen+ presse
5	korrektes Suffix, korrekte Ableitungsbasis, Umlaut fehlt (+Base, +Suff, -Uml) Ablaut fehlt (+Base, +Suff, -Abl)	lauf- → Laufer sing- → Singerin
6	inkorrektes Suffix, korrekte Ableitungsbasis (+Base, -Suff)	begleit- → Begleiter verkauf- → Verkäufern
7	korrektes Suffix, korrekte Ableitungsbasis, Umlaut zuviel (+Base, +Suff, +Uml)	loch- → Löcher
8	inkorrektes Suffix, korrekte Ableitungsbasis, Umlaut zuviel (+Base, -Suff, +Uml)	loch- → Löchung
9	korrektes Suffix, falsche Ableitungsbasis korrekter Umlaut (-Base, +Suff, +Uml)	wander- → Steigung verkauf- → Käuferin
10	inkorrektes Suffix, falsche Ableitungsbasis (-Base, -Suff)	verletz- → Wunde
11	PP, Verb oder Adverb statt Substantiv	üb- → geübt
12	Konversion, implizite Derivation	land- → Flug
13	inkorrektes Suffix, korrekte Ableitungsbasis, Umlaut fehlt (+Base, -Suff, -Uml)	lauf- → Lauferin
14	Insertion, Einfügung	land- → Landerung
15	korrektes Suffix, falsche Ableitungsbasis, inkorrekt Umlaut (-Base, +Suff, -Uml)	verkauf- → Käuferin
16	Pluralbildung an korrekt abgeleitetem Zielwort	verbess- → Verbesserungen

9.8 Ergebnisse des Elizitationsexperiments

Im folgenden Kapitel werden die Daten des Elizitationsexperiments umfassend analysiert und dokumentiert. Nachdem sich die Größe der Gruppen stark unterscheidet (11 HSES- und 5 LSES-Teilnehmer) werden in der grafischen Darstellung der Gruppenvergleiche Mittelwerte errechnet, um diesem Ungleichgewicht Rechnung zu tragen. In Kap. 10.1 werden die Ergebnisse mit den zeitlichen Angaben der Elterninterviews zu Medienkonsum, Büchervorlesen und Inputmenge korreliert.

Die in das Excel Sheet übertragenen Daten wurden zunächst in SPSS importiert. Zur Durchführung der Gruppenvergleiche und Berechnungen war es notwendig weitere Variablen zu ergänzen, die zusätzlich zu den einzelnen Frequenzen die Gesamthäufigkeiten, etwa pro Wortbildungskategorie, enthielten.

Bei der Auswertung des Experimentes zeigte sich eine semikorrekte Kategorie neben den rein korrekten und inkorrekten Kategorien als wesentlich, da es doch zumindest einmal pro Teilnehmer, mit Ausnahme des Kindes D31-AND, zu einer solch halbrichtigen Form kommt (siehe Tabelle 29). Als semikorrekt gelten alle gebildeten Formen, die entweder die richtige Basis, das korrekte Suffix oder beides beinhalten, jedoch in letzterem Fall entweder zusätzlich in ein Kompositum eingebunden sind, z. B. Rasen+**mäh-er**, **Rettt-ung-s**+aktion, oder einen inkorrekten Um-/Ablaut oder eine Einfügung aufweisen (Lecher statt Locher, Begleit-**er-ung** statt Begleitung). Da derartige Formen weder als vollständig inkorrekt noch als korrekt angesehen werden können, ist eine semikorrekte Kategorie die einzig sinnvolle Lösung, um diesem Problem adäquat zu entsprechen.

In Tabelle 29 wird ein erster Gesamtüberblick zu den Teilnehmern und den korrekten, semikorrekten, inkorrekten und fehlenden Antworten gegeben, sowie die Fehleranzahl und die Information, ob eine spontane Selbstkorrektur (autokorr.) stattgefunden hat. Als Selbstkorrektur gelten spontane, also direkt nach der ersten gegebenen Antwort automatisch vom Kind produzierte korrektive Zweitantworten. In einem solchen Fall wurde die zweite Antwort des Kindes als Antwort erfasst und ausgewertet. Die beobachteten spontanen Selbstkorrekturen führten dabei im zweiten Anlauf a) zu einer korrekten Antwort wie in Ansage → Prüfung (korrektes

Zielwort, D03-NIP), Verkäufer → Verkäuferin (korrektes Zielwort, D23-SOK) und Rasenmäher → Mäher (korrektes Zielwort, D31-AND), b) zu einer wiederholt inkorrekten Antwort wie in Sohn → Tochter (Zielwort: Wecker, D09-KAB) und Lehrerin → Fehler (Zielwort: Verbesserung, D04-SOS), sowie c) zu Verschlimmerungen wie in Kleber (korrektes Zielwort) → Tixo (D31-AND). 4 der 9 spontanen Selbstkorrekturen führten zum korrekten Zielwort, in nur einem Fall kam es zu einer Verschlimmerung, das heißt vom korrekten Zielwort zu einem inkorrekten Wort. In allen Fällen ist ein semantischer Zusammenhang der geäußerten Selbstkorrekturen zum Zielwort gegeben.

Tabelle 29. Überblick über die Ergebnisse des Elizitationsexperiments je Teilnehmer

Kürzel	Alter	korr.	semikorr.	inkorr.	autokorr.	keine Antw.	Fehleranzahl
D02-SIJ (m/H)	8;1	28	4	4	0	0	8
D03-NIP (m/H)	8;0	26	2	8	1	0	10
D04-SOS (w/H)	8;2	30	1	5	1	0	6
D06-LAB (w/L)	8;2	26	3	4	0	3	7
D07-NIG (w/H)	7;9	26	3	7	0	0	10
D09-KAB (w/H)	7;9	32	1	3	1	0	4
D12-JAS (w/L)	7;9	23	2	11	0	0	13
D13-JOP (m/H)	8;7	23	4	9	0	0	13
D18-ALM (w/L)	8;7	25	2	5	3	4	7
D19-VAS (w/L)	8;7	26	3	4	0	3	8
D20-LUD (w/H)	7;9	28	2	6	0	0	8
D21-PAS (m/L)	8;11	20	8	8	0	0	16
D23-SOK (w/H)	8;6	27	3	1	1	5	4
D28-MOH (m/H)	8;5	31	1	4	0	0	5
D29-SAK (w/H)	8;7	34	2	0	0	0	2
D31-AND (w/H)	8;3	30	0	6	2	0	7
Gesamt		435	41	85	9	15	128

9.8.1 Korrekte und semikorrekte Ableitungen

Die Antworten der Kinder wurden in Bezug auf die Häufigkeiten korrekter und semikorrekter Derivationen untersucht. Dabei wurde im Bereich der Nomina Agentis zwischen Targets mit und ohne Um-/Ablaut unterschieden. Diese Differenzierung entfällt bei Nomina Instrumenti und Nomina Acti/Actionis, da hier keine Um-/Ablautung zur Bildung nötig ist (vgl. dazu Tabelle 26). Für eine Übersicht der Häufigkeiten pro Teilnehmer nach SES-Status und

Wortbildungskategorie des Zielwortes siehe Tabelle 30 für korrekte Derivationen und in Anhang A10.3 Tabelle 64 für eine Übersicht der semikorrekten Ableitungen.

Tabelle 30. Individuelle und Gesamthäufigkeiten der korrekten Derivationen im Elizitationsexperiment in den Wortbildungskategorien nach SES-Zugehörigkeit

Kind	NAG		NAG		NIN	NAC	Gesamt
	-er	-er/U	-er-in	-er-in/U	-er	-ung	
D06-LAB	3	3	3	3	8	6	26
D12-JAS	2	3	3	3	8	4	23
LSES D18-ALM	3	3	3	3	6	7	25
D19-VAS	3	3	3	3	6	8	26
D21-PAS	2	3	2	2	5	6	20
<i>LSES Ges.^a</i>	<i>13/15</i>	<i>15/15</i>	<i>14/15</i>	<i>14/15</i>	<i>33/60</i>	<i>31/60</i>	<i>120/180</i>
D02-SIJ	3	3	3	3	9	7	28
D03-NIP	3	3	3	3	8	6	26
D04-SOS	3	3	3	3	10	8	30
D07-NIG	3	3	3	3	5	9	26
D09-KAB	3	2	3	3	9	12	32
HSES D13-JOP	2	3	2	3	7	6	23
D20-LUD	3	3	3	3	9	7	28
D23-SOK	3	3	3	3	8	7	27
D28-MOH	3	3	3	3	9	10	31
D29-SAK	3	3	3	3	10	12	34
D31-AND	3	3	3	3	8	10	30
<i>HSES Ges.^a</i>	<i>32/33</i>	<i>32/33</i>	<i>32/33</i>	<i>33/33</i>	<i>92/132</i>	<i>94/132</i>	<i>315/396</i>
Gesamt^a	45/48	47/48	46/48	47/48	125/192	125/192	435/576

a. Die Zwischensummen nach SES-Status geben die richtigen Antworten unter den gesamt möglichen Antworten an.

Es zeigt sich, dass alle Kinder Ableitungen der Wortbildungskategorie Nomina Agentis, unabhängig von Suffix und Um-/Ablaut, problemlos bilden. Beinahe alle Beispielsätze werden korrekt vervollständigt. Nur vereinzelt treten semikorrekte Ableitungen auf, wie beispielsweise aufgrund eines fehlenden Umlauts in Lauf-er, durch die Einbettung des korrekten Derivats in ein Kompositum als Zweitglied wie in Dinge-n+samml-er und vermehrt durch die Derivation einer inkorrekten, aber ebenfalls einfachen Basis wie in Lern-er-in statt Lehr-er-in (vgl. umgangssprachlich lernen in der Bedeutung von lehren) oder Helf-er statt Samml-er. In nur einem

einzigem Fall wird denominal abgeleitet in Schül-er-in statt dem semantisch nahen Antonym Lehr-er-in (was auch in Sprechfehlern von Erwachsenen häufig vorkommt, vgl. Fromkin 2013).

Bildungen von Nomina Instrumenti auf -er und Nomina Acti/Actionis auf -ung sind zwar in der Gesamtanzahl korrekter Ableitungen identisch, doch produzieren die Kinder wesentlich mehr semikorrekte Antworten im Bereich der Nomina Instrumenti (vgl. Abbildung 59). Hauptsächlich handelt es sich hier um komplexe Bildungen, die das korrekt abgeleitete Zielwort als Zweitglied eines Kompositums aufweisen des Typs Rasen+mäh-er, Dose-n+öffn-er, Tasche-n+rechn-er und Fuß+ab#streif-er. Ungeachtet der größeren Komplexität synthetischer Komposita sind speziell die ersten drei Nomina Instrumenti den Kindern durch ihren Input geläufiger und stehen damit in der Testsituation schneller zur Verfügung. Nur selten kommt es zum Entfall des Präfixes vom Typ *Saft-er statt Ent-saft-er und *Streif-er statt Ab#streif-er. Semikorrekte Bildungen bei Nomina Acti/Actionis weisen alle das korrekte Suffix auf. Sie werden a) als Erstglied bei Komposita verwendet wie in **Rett-ung**-s+aktion oder **Üb-ung**-s+sache, enthalten b) die Einfügung -er-, z. B. Land-er-ung, Be-geleit-er-ung, oder c) eine inkorrekte Ableitungsbasis wie in End-ung statt Be-geleit-ung, Rett-ung statt bedeutungsähnlichem Be-frei-ung und Be-lohn-ung statt Impf-ung. In letzterem Beispiel bezieht sich das komplexere Wortbildungsprodukt pragmatisch darauf, was nach der Impfung durch den Arzt passiert, wenn nämlich die Kinder eine Belohnung bekommen, weil sie so brav waren (vgl. letzter Testsatz in Anhang A10.1 Abbildung 87). Einmal kommt es bei den semikorrekten Bildungen bei Nomina Acti/Actionis zu einem Präfixentfall in Kühl-ung statt Ver-kühl-ung.

Abbildung 59 stellt die Mittelwerte der Häufigkeiten korrekter und semikorrekter Ableitungen in Relation zur Gesamtheit der Bildungen nach Wortbildungskategorie und SES-Zugehörigkeit dar.

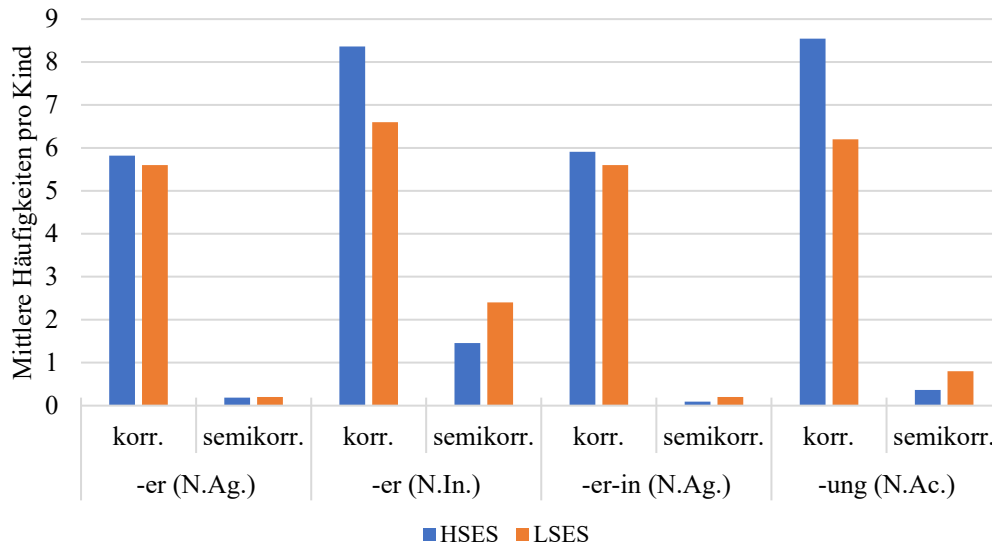


Abbildung 59. Häufigkeiten (Mittelwerte) der korrekten und semikorrekten Ableitungen im Elizitationsexperiment bei HSES und LSES

Bei einem Vergleich der errechneten Mittelwerte der beiden Gruppen (HSES und LSES) wird ersichtlich, dass die HSES-Kinder in allen Kategorien mehr korrekte und weniger semikorrekte Derivationen bilden. Die beiden Gruppen verhalten sich in der Wortbildungskategorie der Nomina Agentis in etwa gleich, während sich in den Bildungen von Nomina Instrumenti und Nomina Acti/Actionis größere Unterschiede zeigen. Abbildung 60 zeigt die Häufigkeiten (Mittelwerte) verschiedener Wortbildungsmuster, die bei den HSES- und LSES-Kindern bei semikorrekten Bildungen auftreten nach Wortbildungskategorie.

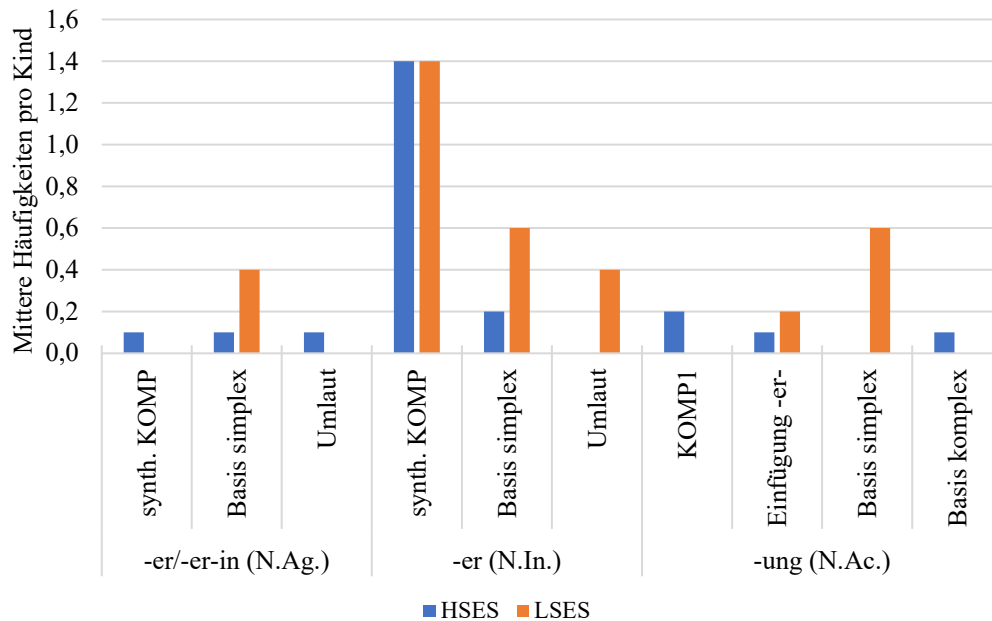


Abbildung 60. Häufigkeiten (Mittelwerte) der Wortbildungsmuster in semikorrekten Derivationen bei HSES und LSES je Wortbildungskategorie

Bei genauer Betrachtung der verwendeten Wortbildungsmuster bei semikorrekten Ableitungen gemessen an der Gesamtheit der Kinder mit semikorrekten Bildungen, ist die hohe Verwendungsrate synthetischer Komposita bei Nomina Instrumenti auffallend, die das korrekte Zielwort enthalten, sowie die generelle Tendenz der LSES-Kinder, einfache Basen zur Ableitung in allen Wortbildungskategorien zu präferieren. Bei den Nomina Agentis handelt es sich dabei um einen Austausch der ohnehin einfachen Basis durch eine andere, während es sich in den Kategorien der Nomina Instrumenti und Nomina Acti/Actionis um die Ersetzung einer komplexen Basis durch eine einfache Basis handelt. Am häufigsten wird die komplexe Ableitungsbasis *ent-saft-* ersetzt, gefolgt von *be-frei-*. Dabei setzen die Kinder geläufigere Wörter ein, die dem Zielwort semantisch nahestehen, etwa *Mix-er* statt *Ent-saft-er* und *Rett-ung* statt *Be-frei-ung*. Im Bereich der Nomina Acti/Actionis produzieren die HSES-Kinder Komposita mit einer Ableitung im Erstglied (KOMP1), wobei es sich bei diesen Derivationen, wie dies auch bei den synthetischen Komposita der Fall ist, um das jeweilige Zielwort handelt. Zusätzliche oder fehlende Um-/Umlautung bei sonst korrekter Ableitung, sowie die Ersetzung einer einfachen durch eine komplexe Basis kommen äußerst selten vor, ebenso wie die Einfügung von *-er-*, die sich ausschließlich bei Bildungen von Nomina Acti/Actionis findet.

9.8.2 Fehleranalyse

Im folgenden Kapitel sollen die Fehler, die die Kinder im Bereich der nominalen Ableitungen produziert haben, genau analysiert werden. Dabei werden die Fehler nach Wortbildungskategorie und SES-Zugehörigkeit unter Bezugnahme auf den Fehlercode untersucht. Ebenso wird auf die fehlenden Antworten eingegangen.

Abbildung 61 zeigt die prozentuale Verteilung der Fehlerhäufigkeiten und der fehlenden Antworten getrennt nach Wortbildungskategorie und Suffix in Relation zur Gesamtanzahl an fehlerhaften Bildungen und keinen Antworten.

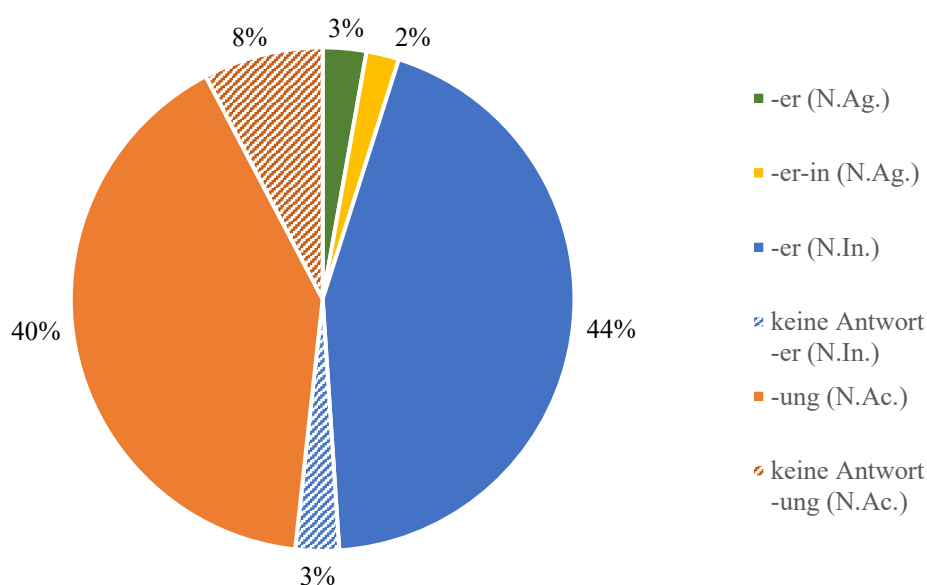


Abbildung 61. Prozentuale Fehlerhäufigkeiten und fehlende Antworten in den einzelnen Wortbildungskategorien gemessen an der Gesamtanzahl

Auf den ersten Blick werden die wesentlich höheren Fehlerquoten in Nomina Instrumenti (N.In.) und Nomina Acti/Actionis (N.Ac.) mit zusammen 84,61 % offensichtlich. Die schraffierten Flächen stellen zusätzlich die fehlenden Antworten in diesen beiden Wortbildungskategorien dar (siehe dazu auch Kap. 9.8.2.1), die sich in Summe auf 10,49 % belaufen. Die Kategorie der Nomina Agentis (N.Ag.) weist im Vergleich lediglich 2,8 % (-er) und 2,1 % (-er-in) Fehler auf. Die Tendenz, keine Antwort von einem Teilnehmer zu erhalten, ist bei Bildungen auf -ung am größten, gefolgt von -er-Ableitungen zur Bildung von Nomina Instrumenti, die überdies das größte Fehlerpotenzial für die Kinder bergen.

Wie sich die Fehlerhäufigkeiten semikorrekter und inkorrektter Ableitungen nach SES-Zugehörigkeit je Wortbildungskategorie verteilen, ist in Abbildung 62 dargestellt. Um jedoch keinen falschen Eindruck von der Höhe der Fehlerquote in den einzelnen Kategorien zu vermitteln, wurden die Mittelwerte der fehlenden Antworten zusätzlich angegeben, um ein gesamtes Bild der Fehlerhäufigkeiten und den gänzlich fehlenden Zielwörtern zu erhalten.

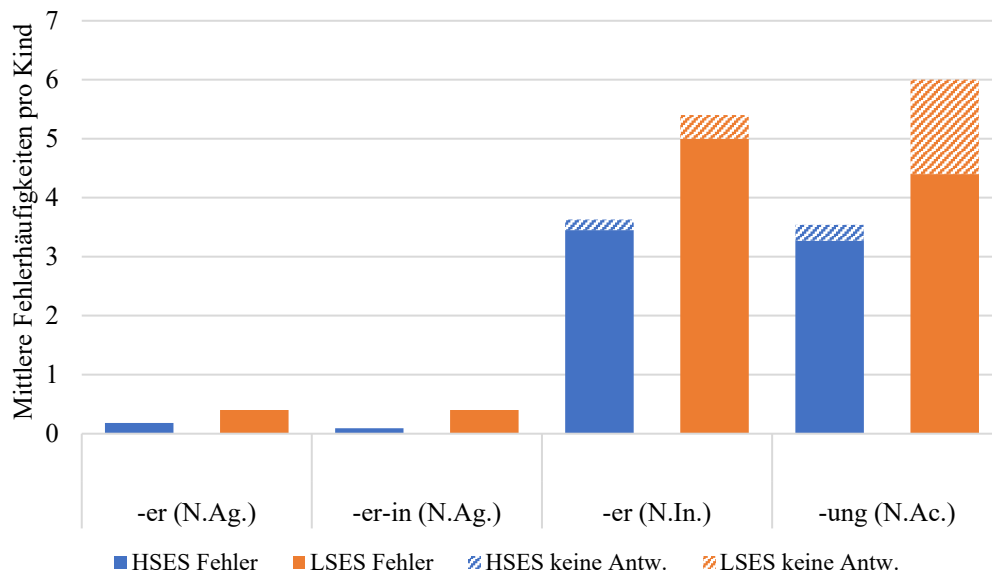


Abbildung 62. Mittelwerte der Fehlerhäufigkeiten semikorrekter und inkorrektter Bildungen, sowie der fehlenden Antworten (schraffiert) im Elizitationsexperiment bei HSES und LSES je Wortbildungskategorie

Anhand des Mittelwertvergleichs in Abbildung 62 lässt sich feststellen, dass die HSES-Kinder im Elizitationsexperiment tendenziell häufiger antworten (schraffiert dargestellt) und dabei weniger Fehler produzieren. Sie weisen nicht nur mehr korrekte und weniger semikorrekte Bildungen (vgl. Kap. 9.8.1, Abbildung 59), sondern auch insgesamt gerechnet eine niedrigere Fehlerquote im Vergleich zu den gleichaltrigen LSES-Kindern auf. Die Wortbildungskategorie der Nomina Agentis präsentiert sich in beiden Gruppen relativ fehlerfrei, wohingegen Nomina Instrumenti und Nomina Acti/Actionis anfälliger für fehlerhafte Bildungen bzw. Modifikationen sind, oder gar zu keiner Antwort führen.

Es ist hier in jedem Fall zu beachten, dass die Fehlergruppe der Nomina Instrumenti im Vergleich zu den anderen Wortbildungskategorien verhältnismäßig viele synthetische Komposita aufweist, die das korrekte Zielwort beinhalten, jedoch

zusätzlich durch das Erstglied näher bestimmt werden, was bei den Nomina Acti/Actionis nicht der Fall ist. In allen Kategorien sind komplexe Verben als Ableitungsbasen in der Unterzahl (1/12 bei NAG, 2/12 bei NIN, 5/12 bei NAC). Simplex-Verben bilden in allen Wortbildungskategorien die hauptsächlichen Ableitungsbasen im Elizitationsexperiment.

9.8.2.1 Keine Antwort

Bei jedem Test, der mit Kindern durchgeführt wird, kann es aus verschiedenen Gründen dazu kommen, dass das eine oder andere Kind keine Antwort gibt. Häufige Gründe können im Allgemeinen etwa eine anfängliche Schüchternheit oder Befangenheit des Kindes sein, die sich im Verlauf der Testung verringern sollte, sowie emotionale Befindlichkeiten (Zorn, Traurigkeit, Übermut), eine neue bzw. unbekannte Umgebung, die Testsituation mit einem:r unbekanntem Experimentator:in, oder bei älteren Kindern, Unsicherheiten, Erfolgsdruck bzw. die Angst davor, etwas falsch zu machen oder schlicht auf keine passende Antwort zu kommen. Von insgesamt 576 möglichen Antworten (36 Testsätze, 16 Kinder) wurde nur 15-mal keine Antwort gegeben, was bedeutet, dass 97,4 % aller Testsätze beantwortet wurden und die Ausfallquote bei lediglich 2,6 % liegt. Mit 8 % ist die Nichtantwortrate in der Kategorie der Nomina Acti/Actionis am höchsten, gefolgt von 3 % in Nomina Instrumenti. Bei den Nomina Agentis wurden sämtliche Testsätze von allen Kindern beantwortet (vgl. Abbildung 61).

Fehlende Antworten sind bei den LSES-Kindern häufiger als bei den HSES-Kindern. Alle Kinder, die manche Sätze nicht vervollständigt haben, sind weiblichen Geschlechts⁴⁷ und haben jeweils 3 oder mehr fehlende Antworten. In Summe handelt es sich um 3 LSES-Mädchen (D06-LAB, D18-ALM, D19-VAS), mit zusammen 10 fehlenden Antworten und 1 HSES-Mädchen (D23-SOK), die 5-mal nicht geantwortet hat (vgl. dazu Tabelle 29 und Abbildung 61). Tendenziell gaben die Kinder an, die Antwort nicht zu wissen: „Weiß nicht.“ oder keine Ahnung zu haben, was darauf hindeutet, dass diese Kinder versuchen, die Wörter aus ihrem Lexikon abzurufen statt sie aktiv von dem gegebenen Stimulusverb abzuleiten. Bei den 15 fehlenden Antworten handelt es sich hauptsächlich um fehlende Bildungen

⁴⁷ Dies lässt sich darauf zurückführen, dass 11 Teilnehmer des Experiments weiblich und weniger als die Hälfte (5) männlich waren.

mittels des Suffixes -ung und in vier Fällen um Nomina Instrumenti auf -er. Von den Ableitungsbasen, der nicht beantworteten Items sind 8 komplex, davon 7 mit untrennbarem Präfix: be-gleit-, ver-besser-, ent-saft-, 1 mit trennbarer Partikel: ab#streif-, und 7 einfach: rechn-, wander-, atm-, land-.

9.8.2.2 Verteilung der Fehler nach Fehlercode

Wie bereits allgemein in Kapitel 9.8.2 aufgegriffen, wird in diesem Abschnitt die Verteilung der einzelnen Fehlerhäufigkeiten je Kind und Fehlercode, mit einem Fokus auf den sozioökonomischen Status, näher betrachtet.

Tabelle 31 enthält die individuellen Fehlerfrequenzen je Kind und Fehlercode, sowie die Summen nach SES-Status und die jeweiligen Gesamtsummen der Teilnehmer.

Tabelle 31. Gesamte und individuelle Fehlerhäufigkeiten je Fehlercode bei HSES und LSES

Kind	Fehlercodierung													Gesamt
	1	2	3	4	5	6	7	9	10	11	12	14	16	
D06-LAB				2				1	2	1	1			7
D12-JAS	4	1						1	1	4	1	1		13
LSES D18-ALM	1		2	2						2				7
D19-VAS	1		2	2			1				1		1	8
D21-PAS			4	1		1		6	1		2	1		16
<i>Summe LSES</i>	<i>6</i>	<i>1</i>	<i>8</i>	<i>7</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>8</i>	<i>9</i>	<i>3</i>	<i>4</i>	<i>1</i>	<i>1</i>	<i>51</i>
D02-SIJ				2				2	3	1				8
D03-NIP	1		2	2						4	1			10
D04-SOS			1	1						3	1			6
D07-NIG		1	3	2				1	2	1				10
D09-KAB	1		2		1									4
HSES D13-JOP	2	1	3	4				1	1	1				13
D20-LUD	1		2	2						2	1			8
D23-SOK				2						1		1		4
D28-MOH	1			1						2	1			5
D29-SAK				2										2
D31-AND	1		3							1	1		1	7
<i>Summe HSES</i>	<i>7</i>	<i>2</i>	<i>16</i>	<i>18</i>	<i>1</i>	<i>0</i>	<i>0</i>	<i>4</i>	<i>19</i>	<i>8</i>	<i>0</i>	<i>1</i>	<i>1</i>	<i>77</i>
Gesamt	13	3	24	25	1	1	2	12	28	11	4	2	2	128

In der in dieser Arbeit auf die L1-Kinder eingeschränkten Datenauswahl des Experiments kommen die Fehlercodes 8, 13 und 15 nicht vor (vgl. Tabelle 31).

Die Fehler 10, 3 und 4 werden mit jeweils über 24 am häufigsten von den Kindern produziert, die Fehler 1 und 11 kommen dagegen bereits halb so häufig vor. Das frequenteste Fehlerbild ist daher die Verwendung einer anderen Ableitungsbasis und eines anderen Suffixes (vgl. dazu Übersicht der Fehlercodierungen in Tabelle 28). Die Kinder verwenden beispielsweise Krank-heit statt ver-kühl- → Verkühlung abzuleiten, oder Spritz-e statt impf- → Impf-ung. Hauptsächlich sind davon Zielwörter auf -ung betroffen. Anders verhält es sich mit Fehlercode 3, also der Bildung eines Kompositums statt einer Derivation. Hier werden hauptsächlich Nomina Instrumenti auf -er durch ein Kompositum ersetzt, z. B. Radier+gummi statt Radier-er, Schuh+matte statt Ab#streif-er und Bohr+maschine statt Bohr-er. Die Ableitungsbasis wird dabei in der Hälfte der Fälle zur Bildung verwendet.

Synthetische Komposita (Fehlercode 4) enthalten in der überwiegenden Mehrheit das Zielwort: Rasen+mäh-er, Fuß+ab#streif-er, Auto+fahr-er, Dosen+öffn-er, Taschen+rechn-er, Dingen+samml-er. In einem Fall bleibt die Ableitungsbasis bestehen: Impf+spritz-e statt Impf-ung und in zwei Fällen ist dies nicht der Fall: Zitronen+press-e statt Ent-saft-er, und Schuh+bürst-e statt Ab#streif-er. Die Verwendung synthetischer Komposita, die die korrekte Derivation als Zweitglied in ein Kompositum einbinden, zeigt, dass diese spezifischeren Ausdrücke in der Spontansprache gebräuchlicher und für die Kinder in dieser Situation leichter abrufbar sind.

Weniger häufig werden statt Ableitungen oder Komposita gänzlich andere Substantive oder Eigennamen verwendet (Fehlercode 1), die allerdings einen semantischen Zusammenhang zum Zielwort aufweisen: Mann statt Arbeit-er, Tixo statt Kleb-er, Besen statt Ab#streif-er. Sofern ein Kind die Derivation nicht bildet oder nicht bilden kann, wird das erste, dem Kind vertraute Wort abgerufen, das denselben semantischen Gehalt aufweist. Die Kinder verwenden also meist nicht irgendwelche Wörter, sondern versuchen, den zweiten Teilsatz adäquat zu beenden.

Von Fehlercode 11 sind zwei Derivationen betroffen, hauptsächlich üb- → Üb-ung, und auch, wenngleich in wesentlich geringerem Ausmaß, ver-besser- → Ver-

besser-ung. In dem Testsatz, der Übung elizitieren sollte, liegt es zum einen an der Konstruktion des zweiten Teilsatzes, zum anderen an der Präsentation durch die Experimentatorin. Im Experimentalblatt für die Experimentatorin wurde auf dieses mögliche Problem, ein Perfektpartizip an Stelle einer Ableitung zu elizitieren, im Testsatz 31 konkret hingewiesen (vgl. in Anhang A10.1 Abbildung 86). Das zeigt jedoch auch, dass die Kinder intuitiv reagieren, auch wenn sie das System des Tests bereits verstanden haben.

Unter Fehlercode 12 wurden morphologische und syntaktische Konversion sowie implizite Ableitung zusammengefasst. Im Zuge des Experimentes kommt es bei drei LSES-Kinder mindestens einmal zu einer derartigen Bildung. Dabei handelt es sich um drei implizite Derivationen: Ausflug statt Wanderung (D12-JAS), Flug statt Landung, Schloss statt Öffner (D21-PAS) und eine syntaktische Konversion, also um die Bildung eines nominalisierten Infinitivs: Abstreifen statt Abstreifer (D06-LAB).

9.8.2.3 Die Verwendung anderer Suffixe

Zur Bildung eines semi- oder inkorrekten Zielwortes kommen die folgenden Suffixe in absteigender Reihenfolge vor: -er, -e, -ung, -heit, -er-in, -el, -in, und -t. Tabelle 32 gibt einen Überblick über die Häufigkeiten der verwendeten Ableitungssuffixe in semi- und inkorrekten Derivationen bei HSES- und LSES-Kindern und enthält Beispiele aus dem Elizitationsexperiment.

Tabelle 32. Übersicht und Häufigkeiten der verwendeten Suffixe in semi- und inkorrekten Ableitungen bei HSES und LSES mit Beispielen

Suffix	HSES	LSES	Beispiel
-er	22	13	*Saft-er
-e	13	2	Zitronen+press-e
-ung	5	5	Be-lohn-ung
-heit	4	5	Frei-heit
-er-in	1	3	Schül-er-in
-el	2	0	Schlüss-el
-in	0	1	*Ver-käuf-in
-t	1	0	Fahr-t

9.8.2.4 Ableitungsbasen in semi- und inkorrekten Derivationen

In inkorrekten und semikorrekten Derivationen kann es zur Verwendung anderer als der vorgegebenen Ableitungsbasen kommen. Abbildung 63 zeigt die Verteilung der Basen in Mittelwerten bei HSES und LSES nach Wortbildungskategorie, wie sie von den Kindern zur Bildung eines Zielworts verwendet wurden.

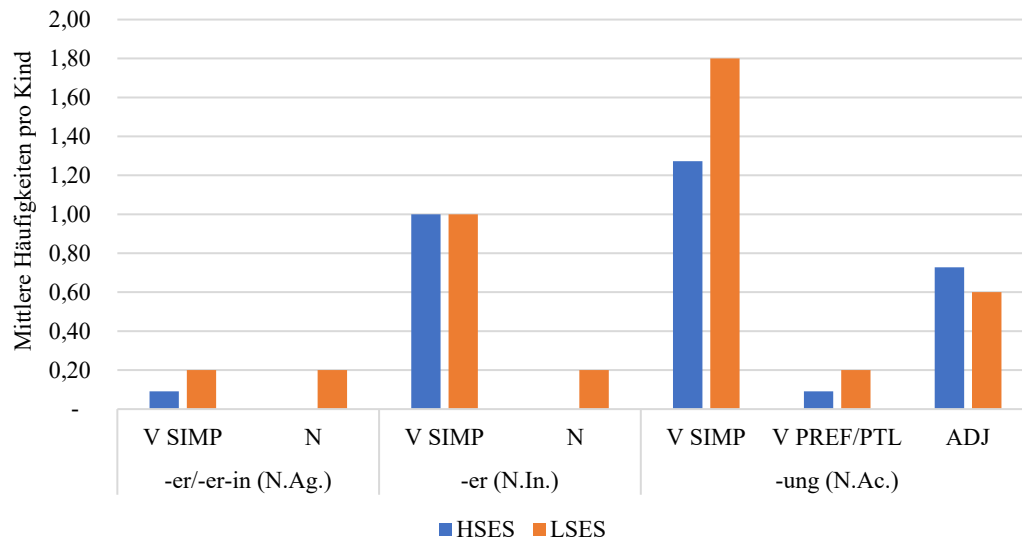


Abbildung 63. Verteilung (Mittelwerte) der verwendeten Ableitungsbasen in semi- und inkorrekten Derivationen bei HSES und LSES je Wortbildungskategorie

Unabhängig von ihrer SES-Zugehörigkeit verwenden die Kinder in semikorrekten und inkorrekten Antworten zumeist Simplex-Verben (V SIMP) als Ableitungsbasis in den Kategorien der Nomina Instrumenti und Nomina Acti/Actionis. 2 der 16 Kinder ziehen je einmal eine komplexe Verbbasis zur Bildung heran: be-lohn- (D02-SIJ) und aus#flieg- (D12-JAS). Adjektive als Basis für Derivationen kommen in den Antworten der Kinder in Verbindung mit -heit zur Bildung von Nomina Acti/Actionis vor: Krank-heit, Frei-heit. Substantive, wie in Schül-er-in (D21-PAS) oder *Saft-er (D06-LAB), dienen nur in diesen Einzelfällen als Ableitungsbasis.

Wie bereits erwähnt blieben annähernd gleich viele Testsätze mit komplexen und einfachen Basen in den Wortbildungskategorien Nomina Instrumenti und Nomina Acti/Actionis unbeantwortet (vgl. Kap. 9.8.2.1 und Abbildung 61) oder führten zu semi- oder inkorrekten Derivationen (vgl. Abbildung 62), wobei die Zahl der semikorrekten Bildungen in NIN wesentlich höher ist im Vergleich zu Nomina Acti/Actionis und umgekehrt in inkorrekten Ableitungen. Jeweils mehr als die

Hälfte der Stimuli sind Simplex-Verben (10/12 NIN, 7/12 NAC, vgl. Tabelle 26). Es kann demnach nicht allein an dem Komplexitätsgrad der Basis liegen, da die Kinder ähnlich viele korrekte Ableitungen in beiden Kategorien bildeten, obwohl in der Gruppe der Nomina Acti/Actionis häufiger von komplexen Verben abzuleiten ist. Dies weist darauf hin, dass die Komplexität der Basis zwar eine Rolle spielt, diese aber hauptsächlich in der Wortbildungskategorie der Nomina Acti/Actionis zum Tragen kommt, wenn also abstrakte statt konkrete Substantive gebildet werden sollen.

9.8.2.5 Die Krux mit dem Umlaut?

Gesamt betrachtet zeigen die Kinder einen sicheren Umgang mit dem Umlaut. In allen Testsätzen wird der Umlaut von den Kindern nur dreimal inkorrekt gebraucht. Einmal entfällt der Umlaut fälschlicherweise in **Lauf**-er (statt Läu**f**-er, D09-KAB). Dies dürfte auf die in der Umgangssprache umlautlosen finiten Verbformen *lauf-st ‚lauf-2S‘ und *lauf-t ‚lauf-3S‘ zurückzuführen sein. In einem Fall wird der Umlaut inkorrekt verwendet in L**ö**ch-er (statt Loch-er, D12-JAS) und bei demselben Testitem kommt es ebenfalls einmal zu einem Ablaut in Le**ch**-er (D19-VAS).

9.8.3 Statistische Auswertung der Ergebnisse des Experiments

Die statistischen Berechnungen wurden mit dem Softwarepaket IBM[®] SPSS[®] Statistics, Version 27, durchgeführt. Zu diesem Zweck wurden die Rohdaten des Experiments zunächst aus Excel importiert. In einem weiteren Schritt wurde jede Variable beschriftet, mit den korrekten Werten versehen und einem Messniveau zugeordnet. Für die Gruppenvergleiche wurden unabhängige T-Tests durchgeführt, um Gruppenunterschiede (SES, Geschlecht) statistisch festzustellen. Da sich die beiden Gruppengrößen voneinander unterscheiden, wurde mit Mittelwerten gerechnet. Die Gruppenvergleiche wurden aufgrund multiplen Testens ebenfalls einer Benjamini-Hochberg-Korrektur unterzogen.

9.8.3.1 Unabhängige T-Tests

Jene Ergebnisse des Elizitationsexperiments, die sich in den vorangehenden Untersuchungen als tendenziell unterschiedlich herausgestellt haben, werden an dieser Stelle statistisch ausgewertet. Dabei sollen mittels unabhängiger T-Tests

Gruppenunterschiede nach sozioökonomischem Status und Geschlecht festgestellt werden.

Gruppenvariable Sozioökonomischer Status

Da ein Vergleich auf Basis sozioökonomischer Faktoren für diese Arbeit besonders relevant ist, wurden die Antworten der Kinder auf Basis verschiedener Variablen statistisch ausgewertet. Da sich die beiden SES-Gruppen in der Wortbildungskategorie der Nomina Agentis annähernd gleich verhalten und aufgrund dessen keine statistisch signifikanten Unterschiede zu erwarten sind, werden diese Variablen nicht berechnet.

Tabelle 33. Mittelwertdifferenzen nach SES-Zugehörigkeit, Signifikanzen (*p*, 2-seitig) und Effektstärken (*r*) verschiedener Variablen des Elizitationsexperiments

Gruppenvergleiche SES

Variablen	LSES		HSES		t	df	<i>p</i>	Sig. ^a	<i>r</i> ^b
	M	SD	M	SD					
korr. Deriv.	24,00	2,550	28,64	3,139	-2,882	14	0,01206	sig. ^a	0,61
-er (NIN)	6,60	1,342	8,36	1,433	-2,323	14	0,03577	sig. ^a	0,53
-ung (NAC)	6,20	1,483	8,55	2,207	-2,145	14	0,04996	sig. ^a	0,50
korr. Basis	27,60	3,050	31,64	2,248	-2,989	14	0,00976	sig. ^a	0,62
korr. Suffix	27,60	1,673	30,73	2,494	-2,533	14	0,02391	sig. ^a	0,56
korr. Umlaut	33,60	1,817	35,45	1,508	-2,146	14	0,04985	sig. ^a	0,50

a. Signifikanz nach Benjamini-Hochberg-Korrektur (*p* < 0,05)

b. Effektstärke (*r* = 0,1 schwach, *r* = 0,3 mittel, *r* = 0,5 stark)

Die Gruppenvergleiche nach SES-Status ergeben statistisch signifikante Unterschiede zwischen HSES (*N* = 11) und LSES (*N* = 5) in allen getesteten Variablen (vgl. Tabelle 33) mit den stärksten Effekten in der Verwendung der korrekten Ableitungsbasis (*r* = 0,62) und der insgesamt korrekt gebildeten Derivationen (*r* = 0,61).

Insgesamt betrachtet bilden die LSES-Kinder durchschnittlich um -4,64 weniger korrekte Derivationen als die HSES-Kinder (95 %-KI[-8,09; -1,19], *p* = 0,01, *r* = 0,61). In den einzelnen Wortbildungskategorien bilden die LSES-Kinder durchschnittlich -1,76 weniger korrekte Nomina Instrumenti und -2,35 weniger korrekte Nomina Acti/Actionis im Vergleich zu den gleichaltrigen HSES-Kindern (-er NIN: 95 %-KI[-3,39; -0,14], *p* = 0,04, *r* = 0,53, -ung NAC: 95 %-KI[-4,69; -0,0], *p* = 0,05, *r* = 0,50). In Bezug auf die Verwendung der korrekten

Ableitungsbasis, des korrekten Suffixes und Umlauts zeigen sich ebenso signifikante Unterschiede bei einem Gruppenvergleich, die sich in der LSES-Gruppe jeweils durchschnittlich geringer zeigen: -4,04 korrekte Basis (95 %-KI [-6,93; -1,14], $p < 0,01$, $r = 0,62$), -3,13 korrektes Suffix (95 %-KI [-5,78; -0,48], $p = 0,02$, $r = 0,56$), -1,85 korrekter Umlaut (95 %-KI [-3,71; -0,0], $p = 0,05$, $r = 0,50$).

Es zeigen sich hingegen keine signifikanten Unterschiede in der mittleren Anzahl der fehlenden Antworten $t(14) = 1,77$; $p = 0,10$ und der Gesamtfehleranzahl $t(14) = 1,70$; $p = 0,11$. Ebenso wenig konnten signifikante SES-Unterschiede in Derivationen, die semikorrekt $t(14) = 1,61$; $p = 0,13$ oder inkorrekt waren $t(14) = 1,02$; $p = 0,32$ festgestellt werden.

Gruppenvariable Geschlecht

Um herauszufinden, ob sich neben SES-spezifischen Unterschieden auch signifikante Unterschiede in Abhängigkeit des Geschlechts in Bezug auf die gleichen Variablen wie in Tabelle 33 zeigen, wurden ebenfalls unabhängige T-Tests durchgeführt, deren Ergebnisse in Tabelle 34 übersichtlich zusammengefasst werden.

Tabelle 34. Mittelwertdifferenzen nach Geschlecht und Signifikanzen (p , 2-seitig) verschiedener Variablen des Elizitationsexperiments

Gruppenvergleiche Geschlecht								
Variablen	männlich		weiblich		t	df	p	Sig. ^a
	M	SD	M	SD				
korr. Deriv.	25,60	4,278	27,91	3,2697	-1,1936	14	0,25247	nicht sig. ^a
-er (NIN)	7,60	1,673	7,91	1,6404	-0,34734	14	0,73350	nicht sig. ^a
-ung (NAC)	7,00	1,732	8,18	2,4421	-0,96866	14	0,34915	nicht sig. ^a
korr. Basis	29,40	3,050	30,82	3,1565	-0,84104	14	0,41446	nicht sig. ^a
korr. Suffix	29,40	2,408	29,91	2,8794	-0,34285	14	0,73680	nicht sig. ^a
korr. Umlaut	36,00	0,000	34,36	1,9633	2,764328	10	0,01998	nicht sig. ^a

a. Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

Die Ergebnisse der T-Tests offenbaren keine statistisch signifikanten Unterschiede nach der Benjamini-Hochberg-Korrektur aufgrund dessen, ob die Teilnehmer männlichen ($N = 5$) oder weiblichen Geschlechts ($N = 11$) sind. Sie unterscheiden sich weder in Relation zur Gesamtheit korrekter Antworten, noch zur Verwendung

der korrekten Basis, des korrekten Suffixes oder des korrekten Umlauts. Des Weiteren zeigen sich keine signifikanten geschlechtsspezifischen Unterschiede in semikorrekten $t(14) = 2,01; p = 0,07$ oder inkorrekten Derivationen $t(14) = 1,23; p = 0,24$, sowie in der Anzahl der fehlenden Antworten $t(10) = -2,30; p = 0,04$ und Fehlerhäufigkeiten $t(14) = 1,89; p = 0,08$. Die Effektstärke r wurde nicht berechnet, da keine signifikanten Unterschiede nach Geschlecht gefunden wurden.

9.9 Abschließende Betrachtungen und Zusammenfassung

Nachdem es sich um ein Elizitationsexperiment handelt, das mündlich durchgeführt wird, können keine Reaktionszeiten gemessen werden. Überdies wurde es bei den Kindern zuhause durchgeführt, wie dies bereits bei den Aufnahmen der Spontansprache (vgl. Kap. 4.1.2) und anderen Testungen der Fall war.

Es hat sich gezeigt, dass sich das vertraute Umfeld günstig auf die Kinder auswirkt, im speziellen, wenn es sich um ein eher schüchternes, zurückhaltendes Kind handelt. Allerdings konnte festgestellt werden, dass die Kinder nicht allein durch das Mitschreiben der Experimentatorin verunsichert werden können, sondern ebenfalls durch die Anwesenheit der Hauptbezugsperson. Wenn sich die Hauptbezugsperson neben dem Kind im Raum befindet, so tendieren die Kinder in einer Testsituation dazu, häufig zu ihrer vertrauten Bezugsperson zu blicken, um sich zu vergewissern, ob sie alles richtig machen.

Außerdem scheint die Prosodie bei der Satzpräsentation durch die Experimentatorin eine Rolle zu spielen. Durch die rein auditive Präsentation verlangt der Test von den Teilnehmern keinerlei Lesekompetenz, was speziell für jüngere Kinder, die noch nicht so lange Lesen lernen, von Vorteil ist, ebenso wie für legasthene Kinder, sofern sie keine Förderung im akustischen Sinneswahrnehmungsbereich benötigen. Überdies werden durch diese Art der Testdurchführung metalinguistische Anforderungen weitgehend vermieden. Durch die konsequente Ableitung verbaler Basen wird das Testsystem schnell aufgefasst und soweit als möglich umgesetzt.

Vereinzelt kommt es im Verlauf des Experiments zu Rückfragen wie bei D06-LAB, Testsatz 34 (Atmung): „Ist das, wenn das Herz so schlägt?“ und Feststellungen, Testsatz 14 (Öffner): „Ich weiß nicht, was er öffnen möchte.“ oder D18-ALM,

Testsatz 24 (Entsafter): „Ich weiß nicht, wie das heißt.“ oder zu Erklärungsversuchen der Kinder, wie D31-AND bei der Vervollständigung des Testsatzes 28 (Verkühlung): „Verkühlung (.) oder Krankheit, das geht beides“.

Im Fall des Testsatzes 14 wurde bei der Satzkonstruktion bewusst auf die Angabe eines direkten Objekts wie z. B. Flasche oder Dose verzichtet, um möglichst kein synthetisches Kompositum zu elizitieren. Dennoch waren die Kinder bemüht bestimmte Substantive so spezifisch wie möglich zu benennen. Es ist davon auszugehen, dass synthetische Komposita, deren Kopf die Ableitungen *-rechn-er*, *-öffn-er* oder insbesondere *-mäh-er* enthalten im Sprachgebrauch häufiger vorkommen und den Kindern somit geläufiger sind, weshalb diese komplexen Wortbildungen in der Testsituation leichter als das Derivat für sich alleine (*Rechn-er*, *Mäh-er*, *Öffn-er*) abrufbar sind (vgl. Kap. 9.8.1, S. 203).

In Summe zeigen die Kinder mit 8 Jahren eine große Sicherheit bei den Nomina Agentis, sowohl die Verwendung der Suffixe (*-er*, *-er-in*) betreffend, als auch des Umlauts. Mehr Variation zeigt sich bei Nomina Instrumenti und Nomina Acti/Actionis.

Die Ergebnisse der T-Tests der Gruppenvergleiche nach sozioökonomischem Status offenbaren in Bezug auf korrekte Bildungen insgesamt, aber auch bezüglich der Verwendung der korrekten Ableitungsbasis, des korrekten Suffixes und des korrekten Umlauts Defizite der LSES-Gruppe bei einem Durchschnittsalter von 8;3 Jahren im Bereich nominaler Ableitungen in den Wortbildungskategorien der Nomina Instrumenti und Nomina Acti/Actionis. Es hat sich herausgestellt, dass LSES-Kinder im Vergleich zu den gleichaltrigen HSES-Kindern signifikant weniger korrekte Bildungen produzieren. Überdies konnte festgestellt werden, dass die HSES-Kinder häufiger antworten und auch weniger Fehler produzieren. Die von beiden Gruppen am häufigsten produzierten Fehler sind zum einen die Bildung eines anderen Derivats mittels einer anderen Ableitungsbasis und eines anderen Suffixes, zum anderen die Bildung von Komposita oder synthetischen Komposita (siehe Kap. 9.8.2).

Gruppenvergleiche nach Geschlecht der Kinder führten hingegen zu keinen statistisch signifikanten Unterschieden. Die Mädchen und Buben des Experiments verhalten sich in allen untersuchten Kategorien annähernd gleich.

Es konnte festgestellt werden, dass die Menge des Büchervorlesens einerseits mit der Anzahl der korrekten Derivationen insgesamt, als auch mit korrekten Ableitungen auf -er in Nomina Instrumenti stark korreliert, andererseits ebenfalls mit der isolierten Verwendung der korrekten Basis, und des korrekten Suffixes. Kinder, denen im Alter von 4 Jahren mehr vorgelesen wird, produzieren im Alter von 8 Jahren nicht nur mehr korrekte Derivationen insgesamt und in Nomina Instrumenti im Test, sondern verwenden auch häufiger die korrekten Suffixe und Ableitungsbasen zur Bildung von Derivationen (siehe dazu Kap. 10.1.5). Signifikante Zusammenhänge zwischen der Menge des elterlichen Inputs, des Fernsehkonsums, der Nutzung Neuer Medien, Hörbücher und Radio konnten hingegen nicht festgestellt werden.

10 Korrelationen

Der Erwerb der Morphologie wird häufig in Abhängigkeit zur Inputmenge (Lemma- und Tokenfrequenzen) betrachtet. In diesem Kapitel soll der Input durch die Hauptbezugsperson, gemessen in Minuten pro Woche im Alter von 4 Jahren, einerseits den Ergebnissen der Spontansprachdaten zum gleichen Zeitpunkt, andererseits den Ergebnissen des Derivationstests, der im durchschnittlichen Alter von 8;3 Jahren durchgeführt wurde, gegenübergestellt werden.

Neben der Inputmenge sollen noch weitere Variablen als mögliche Einflussfaktoren auf Zusammenhänge untersucht werden. Die immer intensivere Nutzung von Medien aller Art, insbesondere audiovisueller Unterhaltung durch Internetnutzung, YouTube, Smartphone und Fernsehen (YISF), aber auch akustische Alternativen wie Hörbücher und Radio (HBR), sowie das traditionelle Büchervorlesen, ebenfalls gemessen in Minuten pro Woche, kann sich auf den Spracherwerb auswirken.

In die Analyse der Daten in Bezug auf die Zusammenhänge zum Derivationsexperiment fließen die Daten der 16 Teilnehmer ein, in Relation zu den Spontansprachdaten werden alle 29 Kinder des INPUT-Projekts untersucht. Von den drei Kindern Jan, Kathi und Lena, die monatlich über mindestens drei Jahre aufgenommen wurden, liegen keine Daten zum zeitlichen Umfang der Inputmenge und des Medienkonsums in Minuten pro Woche vor, daher und wegen der geringen Stichprobengröße können diese nicht auf diese Weise ausgewertet werden.

10.1 Korrelationen Ergebnisse Derivationstest

Die Zusammenhänge ausgewählter Ergebnisse des Derivationsexperiments zum Medienkonsum: YouTube, Internetnutzung, Smartphone und Fernsehen (YISF), Hörbücher und Radio (HBR) und Büchervorlesen sowie der Inputmenge durch die Hauptbezugsperson, die im Alter von 4 Jahren in Elterninterviews erhoben wurden (siehe Kap. 4.2), sollen mittels Korrelationen festgestellt werden. Dabei werden die Zeitangaben des Medienkonsums und Inputs in Minuten pro Woche, den Anzahlen der Ergebnisse des Tests gegenübergestellt, um zu sehen, ob sich signifikante Zusammenhänge über diesen Zeitraum zeigen.

Die Analyse der Daten offenbarte, dass einige Variablen nicht normalverteilt waren. Aus diesem Grund wurde eine Spearman-Korrelation durchgeführt, um den Zusammenhang zwischen den individuellen Variablen zu zeigen. Jede der vier Variablen (INPUT, YISF, HBR und Büchervorlesen) befindet sich in den Korrelationstabellen (35, 36, 37 und 38) an der ersten Position, wo ihre Relation zu den anderen Variablen präsentiert wird und dadurch einfach ablesbar ist. Die Daten aller 16 Kinder, die Teilnehmer des Derivationsexperiments waren, sind in diese Korrelationen integriert, eine Unterscheidung nach SES-Zugehörigkeit wurde nicht getroffen.

In den folgenden Spearman-Korrelationstabellen (35, 36, 37 und 38) bedeutet ein * eine signifikante Korrelation auf dem Niveau von 0,05 (2-seitig) und ** bedeutet ein Signifikanzniveau von 0,01 (2-seitig). Die errechneten Korrelationen wurden aufgrund multipler Vergleiche Benjamini-Hochberg-korrigiert. Variablen, die nach der Korrektur auf einem Signifikanzniveau von 0,05 oder darunter immer noch signifikant sind, sind durch ein + markiert, ein zusätzliches ^a weist auf einen hoch signifikanten Zusammenhang hin auf einem Niveau von 0,01 oder darunter. Variablen, die danach nicht mehr signifikant korrelieren, werden im Folgenden nicht behandelt, wesentliche Trends jedoch aufgezeigt.

Im Anschluss an die Tabellen werden zunächst die primären Zusammenhänge erläutert und danach die Korrelationen zu den Variablen INPUT, YISF, HBR und Büchervorlesen beschrieben.

Tabelle 35. Korrelationen Ergebnisse des Derivationsexperiments – INPUT

	INPUT Min./Wo.	korr. Deriv. -er (NIN)	korr. Deriv. -ung (NAC)	keine Antw.	korr. Suffix	korr. Umlaut	korr. Basis	Fehleranzahl	korr. Deriv.	semikorr. Deriv.	inkorr. Deriv.
INPUT Min./Wo.	1	0,05	-0,09	0,25	0,09	-0,03	-0,04	-0,23	0,06	0,26	-0,27
korr. Deriv. -er (NIN)	0,05	1	0,41	-0,34	,763**+ ^a	0,27	,713**+ ^a	-,635**+ ^a	,814**+ ^a	-,541*	-0,49
korr. Deriv. -ung (NAC)	-0,09	0,41	1	-0,15	,787**+ ^a	0,14	,874**+ ^a	-,709**+ ^a	,831**+ ^a	-,548**+	-,622**+
keine Antw.	0,25	-0,34	-0,15	1	-0,23	-,872**+ ^a	-0,25	-0,27	-0,24	0,23	-0,41
korr. Suffix	0,09	,763**+ ^a	,787**+ ^a	,787**+ ^a	1	0,24	,912**+ ^a	-,764**+ ^a	,942**+ ^a	-0,38	-,770**+ ^a
korr. Umlaut	-0,03	0,27	0,14	-,872**+ ^a	0,24	1	0,25	0,25	0,22	-0,09	0,36
korr. Basis	-0,04	,713**+ ^a	,874**+ ^a	-0,25	,912**+ ^a	0,25	1	-,744**+ ^a	,962**+ ^a	-,596**+	-,668**+ ^a
Fehleranzahl	-0,23	-,635**+ ^a	-,709**+ ^a	-0,27	-,764**+ ^a	0,25	-,744**+ ^a	1	-,822**+ ^a	,547**+	,876**+ ^a
korr. Deriv.	0,06	,814**+ ^a	,831**+ ^a	-0,24	,942**+ ^a	0,22	,962**+ ^a	-,822**+ ^a	1	-,641**+ ^a	-,696**+ ^a
semikorr. Deriv.	0,26	-,541*	-,548**+	0,23	-0,38	-0,09	-,596**+	,547**+	-,641**+ ^a	1	0,16
inkorr. Deriv.	-0,27	-0,49	-,622**+	-0,41	-,770**+ ^a	0,36	-,668**+ ^a	,876**+ ^a	-,696**+ ^a	0,16	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,01$)

Tabelle 36. Korrelationen Ergebnisse des Derivationsexperiments – YISF

	YISF Min./Wo.	korr. Deriv. -er (NIN)	korr. Deriv. -ung (NAC)	keine Antw.	korr. Suffix	korr. Umlaut	korr. Basis	Fehleranzahl	korr. Deriv.	semikorr. Deriv.	inkorr. Deriv.
YISF Min./Wo.	1										
korr. Deriv. -er (NIN)	-0,06	1									
korr. Deriv. -ung (NAC)	0,01	0,41	1								
keine Antw.	,601*	-0,34	-0,15	1							
korr. Suffix	0,02	,763**+ ^a	,787**+ ^a	-0,23	1						
korr. Umlaut	-0,599*	0,27	0,14	-0,23	0,24	1					
korr. Basis	-0,01	,713**+ ^a	,874**+ ^a	-0,25	0,25	0,25	1				
Fehleranzahl	-0,37	-0,34	-0,15	-0,27	0,25	0,25	0,25	1			
korr. Deriv.	-0,04	,814**+ ^a	,831**+ ^a	-0,24	,942**+ ^a	,962**+ ^a	-0,24	-0,24	1		
semikorr. Deriv.	0,10	-0,49	-0,22	-0,38	-0,09	-0,09	-0,641**+ ^a	-0,641**+ ^a	1	1	0,16
inkorr. Deriv.	-0,45	-0,49	-0,22	-0,41	0,36	0,36	-0,668**+ ^a	-0,696**+ ^a	-0,641**+ ^a	0,16	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,01$)

Tabelle 37. Korrelationen Ergebnisse des Derivationsexperiments – HBR

	HBR Min./Wo.	korr. Deriv. -er (NIN)	korr. Deriv. -ung (NAC)	keine Antw.	korr. Suffix	korr. Umlaut	korr. Basis	Fehleranzahl	korr. Deriv.	semikorr. Deriv.	inkorr. Deriv.
HBR Min./Wo.	1	-0,01	-0,18	0,46	-0,12	-,559*	-0,16	0,00	-0,16	0,06	-0,13
korr. Deriv. -er (NIN)	-0,01	1	0,41	-0,34	,763** ^a	0,27	,713** ^a	-,635** ^a	,814** ^a	-,541*	-0,49
korr. Deriv. -ung (NAC)	-0,18	0,41	1	-0,15	,787** ^a	0,14	,874** ^a	-,709** ^a	,831** ^a	-,548** ^a	-,622** ^a
keine Antw.	0,46	-0,34	-0,15	1	-0,23	-,872** ^a	-0,25	-0,27	-0,24	0,23	-0,41
korr. Suffix	-0,12	,763** ^a	,787** ^a	-,872** ^a	1	0,24	,912** ^a	-,764** ^a	,942** ^a	-0,38	-,770** ^a
korr. Umlaut	-,559*	0,27	0,14	-,872** ^a	0,24	1	0,25	0,25	0,22	-0,09	0,36
korr. Basis	-0,16	,713** ^a	,874** ^a	-,872** ^a	-,912** ^a	0,25	1	-,744** ^a	,962** ^a	-,596** ^a	-,668** ^a
Fehleranzahl	0,00	-,635** ^a	-,709** ^a	-0,25	,912** ^a	0,25	-,744** ^a	1	-,822** ^a	,547** ^a	,876** ^a
korr. Deriv.	-0,16	,814** ^a	,831** ^a	-0,24	,942** ^a	0,22	,962** ^a	-,822** ^a	1	-,641** ^a	-,696** ^a
semikorr. Deriv.	0,06	-,541*	-,548** ^a	0,23	-0,38	-0,09	-,596** ^a	,547** ^a	-,641** ^a	1	0,16
inkorr. Deriv.	-0,13	-0,49	-,622** ^a	-0,41	-,770** ^a	0,36	-,668** ^a	,876** ^a	-,696** ^a	0,16	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,01$)

Tabelle 38. Korrelationen Ergebnisse des Derivationsexperiments – Büchervorlesen

	Bücher Min./Wo.	korr. Deriv. -er (NIN)	korr. Deriv. -ung (NAC)	keine Antw.	korr. Suffix	korr. Umlaut	korr. Basis	Fehleranzahl	korr. Deriv.	semikorr. Deriv.	inkorr. Deriv.
Bücher Min./Wo.	1	,611*+	,526*	-0,41	,622**+ ^a	0,28	,588*+	-0,48	,661**+ ^a	-0,38	-0,32
korr. Deriv. -er (NIN)	,611*+	1	0,41	-0,34	,763**+ ^a	0,27	,713**+ ^a	-,635**+ ^a	,814**+ ^a	-,541*	-0,49
korr. Deriv. -ung (NAC)	,526*	0,41	1	-0,15	,787**+ ^a	0,14	,874**+ ^a	-,709**+ ^a	,831**+ ^a	-,548*+	-,622*+
keine Antw.	-0,41	-0,34	-0,15	1	-0,23	-,872**+ ^a	-0,25	-0,27	-0,24	0,23	-0,41
korr. Suffix	,622**+ ^a	,763**+ ^a	,787**+ ^a	-0,23	1	0,24	,912**+ ^a	-,764**+ ^a	,942**+ ^a	-0,38	-,770**+ ^a
korr. Umlaut	0,28	0,27	0,14	-,872**+ ^a	0,24	1	0,25	0,25	0,22	-0,09	0,36
korr. Basis	,588*+	,713**+ ^a	,874**+ ^a	-0,25	,912**+ ^a	0,25	1	-,744**+ ^a	,962**+ ^a	-,596*+	-,668**+ ^a
Fehleranzahl	-0,48	-,635**+ ^a	-,709**+ ^a	-0,27	-,764**+ ^a	0,25	-,744**+ ^a	1	-,822**+ ^a	,547*+	,876**+ ^a
korr. Deriv.	,661**+ ^a	,814**+ ^a	,831**+ ^a	-0,24	,942**+ ^a	0,22	,962**+ ^a	-,822**+ ^a	1	-,641**+ ^a	-,696**+ ^a
semikorr. Deriv.	-0,38	-,548*+	-,548*+	0,23	-0,38	-0,09	-,596*+	,547*+	-,641**+ ^a	1	0,16
inkorr. Deriv.	-0,32	-0,49	-,622*+	-0,41	-,770**+ ^a	0,36	-,668**+ ^a	,876**+ ^a	-,696**+ ^a	0,16	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,05$)

+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p < 0,01$)

10.1.1 Primäre Zusammenhänge der Ergebnisse

Die Gesamtanzahl der Antworten setzt sich aus allen korrekten, semikorrekten und inkorrekten Derivationen zusammen. Die Variablen weisen einen signifikanten Zusammenhang zu korrekten Derivationen auf, sowie einen starken, negativen Effekt: semikorr. Deriv. ($r_s = -0,641, p < 0,01$) und inkorr. Deriv. ($r_s = -0,696, p < 0,01$). Je mehr korrekte Derivationen, desto weniger semi- oder inkorrekte Derivationen werden gebildet. Ebenso zeigt sich eine erwartete starke, negative Korrelation korrekter Derivationen zur Fehleranzahl ($r_s = -0,822, p < 0,01$) und ein starker positiver Zusammenhang zu inkorr. Deriv. ($r_s = 0,876, p < 0,01$).

Die Summe der korrekten Derivationen (korr. Deriv.) enthält alle korrekten Ableitungen in allen Wortbildungskategorien. Die Subvariablen der korrekten Derivationen auf -er in Nomina Instrumenti ($r_s = 0,814, p < 0,01$) und auf -ung in Nomina Acti/Actionis ($r_s = 0,831, p < 0,01$) korrelieren signifikant und zeigen starke Effekte zum Gesamtscore, da sie in diesem enthalten sind.

Weitere signifikante Zusammenhänge bestehen zwischen der Anzahl der korrekten Derivationen und der Verwendung des korrekten Suffixes oder der korrekten Ableitungsbasis. Je mehr korrekte Derivationen gebildet wurden, desto häufiger wurde auch insgesamt gesehen (in allen Ableitungen) das korrekte Suffix ($r_s = 0,942, p < 0,01$) oder die korrekte Basis verwendet ($r_s = 0,962, p < 0,01$). Diese starken Effekte sind ebenfalls in Bezug auf die Subvariablen korr. Deriv. -er in Nomina Instrumenti (korr. Suffix: $r_s = 0,763, p < 0,01$, korr. Basis: $r_s = 0,713, p < 0,01$) und -ung in Nomina Acti/Actionis (korr. Suffix: $r_s = 0,787, p < 0,01$, korr. Basis: $r_s = 0,874, p < 0,01$) feststellbar.

Ebenso korreliert die Anzahl der korrekten Umlaute (korr. Umlaut) mit der Anzahl der fehlenden Antworten (keine Antw.) signifikant. Je größer die Anzahl der fehlenden Antworten, desto weniger korrekte Umlaute wurden gebildet. Anders formuliert, tendierten jene Kinder, die mehr korrekte Umlaute bildeten dazu, seltener nicht zu antworten, also häufiger zu antworten.

10.1.2 Inputmenge

Die in Tabelle 35 errechneten Korrelationen ergeben keinen signifikanten Zusammenhang und keine Effekte zwischen der Inputmenge in Minuten pro Woche im Alter von 4 Jahren zu den Variablen des Derivationsexperiments. Die Ergebnisse des Tests sind demnach unabhängig und stehen in keinem Zusammenhang zu der Menge des Inputs.

10.1.3 Neue Medien – YouTube, Internetnutzung, Smartphone, Fernsehen

Es können nach einer Benjamini-Hochberg-Korrektur keine signifikanten Korrelationen zwischen der Mediennutzung YISF und den Ergebnissen des Derivationstests festgestellt werden (vgl. Tabelle 36). Es ergeben sich jedoch davor gewisse Trends, die auf einen potenziellen Zusammenhang zu der Menge des Medienkonsums schließen lassen. Kinder, die im Alter von 4 Jahren mehr Neue Medien konsumieren, neigen dazu, im Experiment im Alter von 8 Jahren einerseits häufiger nicht zu antworten ($r_s = 0,601, p = 0,01$), andererseits seltener den korrekten Umlaut zu bilden ($r_s = -0,599, p = 0,01$). Diese starken Zusammenhänge sind jedoch nach der Korrektur nicht mehr signifikant (vgl. Tabelle 36).

10.1.4 Hörbücher und Radio

Der Konsum von Hörbüchern und Radioprogrammen korreliert nicht mit den Ergebnissen des Derivationsexperiments, wie in Tabelle 37 ersichtlich. Es kann kein statistisch signifikanter Zusammenhang festgestellt werden zwischen den Minuten pro Woche, die die Kinder mit Hörbüchern und Radio zubringen, und den Variablen des Tests.

10.1.5 Büchervorlesen

Im Gegensatz zu den anderen Medienvariablen zeigen sich im Bereich des Büchervorlesens signifikante Zusammenhänge und starke Effekte zu den Ergebnissen des Derivationstests (vgl. Tabelle 38). Kinder, denen mit 4 Jahren mehr vorgelesen wurde, produzieren signifikant mehr korrekte Ableitungen im Elizitationsexperiment im Alter von 8 Jahren ($r_s = 0,661, p < 0,01$). Dies zeigt sich ebenso an mehr korrekten Ableitungen auf -er in Nomina Instrumenti ($r_s = 0,611, p < 0,05$), was wenig überrascht, da es sich um eine Subvariable der Gesamtheit der korrekten Derivationen handelt. Allerdings korreliert eine höhere Vorleserate

ebenso mit der Anzahl der insgesamt im Test verwendeten korrekten Suffixe ($r_s = 0,622, p < 0,01$) und der korrekt verwendeten Basen ($r_s = 0,588, p < 0,05$). Ein Trend zeigt sich ebenso in Zusammenhang mit korrekten Derivationen auf -ung, wo vor der Benjamini-Hochberg-Korrektur ein starker Zusammenhang besteht ($r_s = 0,526, p = 0,04$).

Kinder, die im Alter von 4 Jahren mehr vorgelesen bekommen, verwenden im Alter von 8 Jahren nicht nur mehr korrekte Suffixe und Basen, sondern bilden insgesamt signifikant mehr korrekte Derivationen in diesem Experiment.

10.2 Korrelationen zu den Spontansprachdaten

In Kapitel 10.1 wurden die Ergebnisse des Derivationsexperiments, das im Alter von 8 Jahren durchgeführt wurde, der Inputmenge und der Mediennutzung im Alter von 4 Jahren gegenübergestellt. In diesem Kapitel sollen die spontansprachlichen Ergebnisse der Kinder in Bezug auf Derivationen und Komposita, die zu den mittleren Zeitpunkten 4;4 und 4;8 erhoben wurden, ebenfalls auf Zusammenhänge zum Medienkonsum (YISF: YouTube, Internetnutzung, Smartphone, Fernsehen; HBR: Hörbücher, Radio; Bücher: Büchervorlesen) und zu ihrer Inputmenge (INPUT) im gleichen Alterszeitraum überprüft werden. Der Medienkonsum und die Inputmenge wurden in Minuten pro Woche in Elterninterviews erhoben (siehe Kap. 4.2) und werden der Anzahl der Derivationen und Komposita gegenübergestellt, um etwaige signifikante Korrelationen zwischen den spontansprachlichen Bildungen der Kinder zu den zuvor erwähnten vier Variablen aufzuzeigen.

Wie bereits in Kap. 10.1 erwähnt, waren einige Variablen nicht normalverteilt, weshalb Spearman-Korrelationen berechnet wurden, um den Zusammenhang zwischen den individuellen Variablen zu berechnen. Dabei befinden sich in den Korrelationstabellen (39, 40 und 41) die vier Variablen (INPUT, YISF, HBR, Büchervorlesen) in dieser Reihenfolge an den ersten vier Positionen.

Es fließen die Daten aller 29 Kinder, die Teilnehmer des INPUT-Projekts waren, in die Korrelationen ein, wobei es keine Unterscheidung nach SES-Zugehörigkeit gibt. Berechnet wurden die Anzahlen der Derivationen und Komposita zu den Datenzeitpunkten 3 und 4 in den drei Wortklassen Substantiv, Verb und Adjektiv,

wobei jeweils die in CS häufigsten Kategorien für die Korrelationen ausgewählt wurden.

In den folgenden Korrelationstabellen (39, 40 und 41) markiert ein * eine signifikante Korrelation auf dem Niveau von 0,05 (2-seitig), ein ** ein Signifikanzniveau von 0,01 (2-seitig). Aufgrund multipler Vergleiche wurden die errechneten Korrelationen Benjamini-Hochberg-korrigiert. Variablen, die nach der Korrektur auf einem Niveau von 0,05 oder darunter nach wie vor signifikant sind, wurden mit einem + markiert, ein zusätzliches ^a weist auf einen hoch signifikanten Zusammenhang auf einem Niveau von 0,01 oder darunter hin. Korrelieren die Variablen nach der Korrektur nicht mehr signifikant, werden sie im Folgenden nicht behandelt, Trends jedoch aufgezeigt.

Die primären Zusammenhänge der Variablen werden im Anschluss an die Tabellen in Kap. 10.2.1 kurz erläutert. Auf die errechneten Korrelationen der Spontansprache zu den Variablen INPUT, YISF, HBR und Büchervorlesen wird in den Kap. 10.2.1–10.2.5 eingegangen, in Kap. 10.2.6 werden spontansprachliche Bildungen auf -er und -ung den Ergebnissen des Elizitationsexperiments gegenübergestellt und die statistisch signifikanten Zusammenhänge beschrieben.

Tabelle 39. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): Verbale Derivationen – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)

DP 3+4 CHI Lemmas	INPUT Min./Wo.	YISF Min./Wo.	HBR Min./Wo.	Bücher Min./Wo.	V DERIV	V PTL	V PREF	V KONV
INPUT	1	,413*	-,445*	0,233	-0,324	-0,274	-0,227	-0,201
YISF	,413*	1	0,007	-0,330	-0,362	-0,343	-0,344	-0,101
HBR	-,445*	0,007	1	-,454*	0,201	0,173	-0,038	0,076
Bücher	0,233	-0,330	-,454*	1	-0,006	-0,017	0,215	-0,084
V DERIV	-0,324	-0,362	0,201	-0,006	1	,957***+ ^a	,556***+ ^a	0,233
V PTL	-0,274	-0,343	0,173	-0,017	,957***+ ^a	1	0,359	0,090
V PREF	-0,227	-0,344	-0,038	0,215	,556***+ ^a	0,359	1	0,283
V KONV	-0,201	-0,101	0,076	-0,084	0,233	0,090	0,283	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,05$)

+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,01$)

Tabelle 40. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): derivierte und zusammengesetzte Adjektive – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)

DP 3+4 CHI Lemmas	INPUT Min./Wo.	YISF Min./Wo.	HBR Min./Wo.	Bücher Min./Wo.	ADJ DERIV	ADJ KOMP	ADJ PP	ADJ SUFF
INPUT	1	,413*	-,445*	0,233	-0,003	0,265	0,049	0,008
YISF	,413*	1	0,007	-0,330	-0,310	-0,153	-0,277	-0,262
HBR	-,445*	0,007	1	-,454*	-0,249	-0,267	0,075	-0,202
Bücher	0,233	-0,330	-,454*	1	0,291	0,296	-0,040	0,386
ADJ DERIV	-0,003	-0,310	-0,249	0,291	1	,384*	-0,120	,830***+ ^a
ADJ KOMP	0,265	-0,153	-0,267	0,296	,384*	1	0,052	0,211
ADJ PP	0,049	-0,277	0,075	-0,040	-0,120	0,052	1	-0,179
ADJ SUFF	0,008	-0,262	-0,202	0,386	,830***+ ^a	0,211	-0,179	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,05$)

+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,01$)

Tabelle 41. Korrelationen Ergebnisse der Spontansprachdaten in CS (DP 3+4): Nominale Derivationen und Komposita – INPUT, YISF, HBR, Büchervorlesen (Min./Wo.)

DP 3+4 CHI Lemmas	INPUT Min./Wo.	YISF Min./Wo.	HBR Min./Wo.	Bücher Min./Wo.	N DERIV Min./Wo.	N KOMP	N synth. KOMP	N SUFF	N KONV
INPUT Min./Wo.	1	,413*	-,445*	0,233	-,0,230	0,015	-,0,191	-,0,183	-,0,203
YISF Min./Wo.	,413*	1	0,007	-,0,330	-,501**+	-,383*	-,0,281	-,541**+ ^a	-,0,252
HBR Min./Wo.	-,445*	0,007	1	-,454*	0,026	-,0,343	0,007	-,0,004	0,162
Bücher Min./Wo.	0,233	-,0,330	-,454*	1	0,112	,398*	0,257	0,071	0,065
N DERIV	-,0,230	-,501**+	0,026	0,112	1	0,284	0,310	,745**+ ^a	,581**+ ^a
N KOMP	0,015	-,383*	-,0,343	,398*	0,284	1	,734**+ ^a	,535**+ ^a	-,0,256
N synth. KOMP	-,0,191	-,0,281	0,007	0,257	0,310	,734**+ ^a	1	,529**+ ^a	-,0,159
N SUFF	-,0,183	-,541**+ ^a	-,0,004	0,071	,745**+ ^a	,535**+ ^a	,529**+ ^a	1	0,026
N KONV	-,0,203	-,0,252	0,162	0,065	,581**+ ^a	-,0,256	-,0,159	0,026	1

+ Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,05$)

+^a Signifikanz nach Benjamini-Hochberg-Korrektur ($p \leq 0,01$)

10.2.1 Primäre Zusammenhänge

Die Variablen N DERIV, V DERIV und ADJ DERIV sind die Hauptvariablen, die alle anderen derivationalen Kategorien in sich vereinen. So ist die Variable ADJ SUFF in ADJ DERIV ebenso enthalten wie V PTL und V PREF in V DERIV und N SUFF und N KONV in N DERIV. Daher korrelieren die Subvariablen jeweils signifikant mit ihrer Hauptvariable, da sie in dieser enthalten ist: ADJ SUFF ($r_s = 0,830, p < 0,01$), V PTL ($r_s = 0,957, p < 0,01$), V PREF ($r_s = 0,556, p < 0,01$), N SUFF ($r_s = 0,745, p < 0,01$) und N KONV ($r_s = 0,581, p < 0,01$).

Überdies gibt es signifikante Zusammenhänge zwischen der Anzahl zusammengesetzter Substantive (N KOMP), synthetischer Komposita (N synth. KOMP) und suffigierter Substantive (N SUFF). Je mehr Bildungen es in einer der drei Kategorien gibt, desto mehr werden auch in den anderen beiden Kategorien gebildet. Beispielsweise steigt die Anzahl suffigierter Substantive an, wenn eine höhere Anzahl synthetischer Komposita oder Substantivkomposita gebildet wird und umgekehrt. Die Variablen weisen die folgenden signifikanten und starken Zusammenhänge auf: N KOMP zu synth. KOMP ($r_s = 0,734, p < 0,01$), N KOMP zu N SUFF ($r_s = 0,535, p < 0,01$), synth. KOMP zu N SUFF ($r_s = 0,529, p < 0,01$).

Ein Trend ergibt sich bei derivierten und zusammengesetzten Adjektiven. Die beiden Wortbildungsarten korrelieren moderat positiv, allerdings nach der Benjamini-Hochberg-Korrektur nicht mehr signifikant ($r_s = 0,384, p = 0,04$).

10.2.2 Inputmenge

Die in den Tabellen 39, 40 und 41 errechneten Korrelationen ergeben keine statistisch signifikanten Zusammenhänge zwischen der Inputmenge in Minuten pro Woche der Hauptbezugsperson zu den Variablen der Spontansprachergebnisse der Kinder. Es kann weder für Substantive noch für Verben oder Adjektive, unabhängig davon, ob diese deriviert oder zusammengesetzt sind, ein statistisch signifikanter Zusammenhang zur Inputmenge gezeigt werden. Im Alter von 4 Jahren stehen die Häufigkeiten gebildeter Derivationen und Komposita in CS in den untersuchten Wortklassen in keinem Zusammenhang zur Inputmenge im gleichen Alterszeitraum.

10.2.3 Neue Medien – YouTube, Internetnutzung, Smartphone, Fernsehen

Im Bereich der Neuen Medien sind signifikante Zusammenhänge und starke, negative Effekte zu der Anzahl derivierter Substantive und im Speziellen suffigierter Substantive feststellbar (vgl. Tabelle 41). Kinder, die mit 4 Jahren mehr Neue Medien konsumieren, bilden im gleichen Alter signifikant weniger nominale Ableitungen ($r_s = -0,501$, $p < 0,05$). Dieser Effekt zeigt sich insbesondere bei suffigierten Substantiven, wo jene Kinder mit einem höheren Medienkonsum signifikant weniger bilden ($r_s = -0,541$, $p < 0,01$). Nachdem die Variable N SUFF eine Subvariable von N DERIV darstellt, ist der Effekt, der sich bei der Hauptvariable zeigt, darauf zurückzuführen. Ein Trend zeigt sich ebenfalls in Richtung nominaler Komposita vor der Benjamini-Hochberg-Korrektur, wo sich ebenso ein negativer, wenngleich nicht signifikanter Zusammenhang zum Medienkonsum ergibt ($r_s = -0,383$, $p = 0,04$). Die Anzahl derivierter und zusammengesetzter Lexeme in den Wortklassen der Verben und Adjektive korrelieren statistisch nicht signifikant mit der Variable YISF (vgl. Tabellen 39, 40).

Kinder, die im Alter von 4 Jahren mehr Neue Medien konsumieren, sei es YouTube, Internet, Smartphone oder Fernsehen, bilden im gleichen Alter insgesamt signifikant weniger derivierte Substantive, insbesondere jedoch signifikant weniger suffigierte Substantive. Dieser Effekt kann überdies weder für derivierte Verben, noch für derivierte oder zusammengesetzte Adjektive festgestellt werden.

10.2.4 Hörbücher und Radio

Es können keine Korrelationen festgestellt werden zwischen dem Konsum von Hörbüchern und Radioprogrammen und der Anzahl an spontansprachlichen Bildungen von Derivationen und Komposita, wie in den Tabellen 39, 40 und 41 für die Wortklassen der Substantive, Verben und Adjektive ersichtlich. Es besteht daher kein Zusammenhang zwischen der Menge an auditivem Medienkonsum und der Menge an Wortbildungen im Alter von 4 Jahren.

10.2.5 Büchervorlesen

Nach der Benjamini-Hochberg-Korrektur können keine statistisch signifikanten Zusammenhänge zwischen der Zeit, die die Kinder Bücher vorgelesen bekommen,

und der Menge derivierter und zusammengesetzter Substantive und Adjektive, sowie derivierter Verben festgestellt werden (vgl. Tabellen 39, 40, 41). Die Lexeme, die die Kinder mit 4 Jahren bilden, stehen folglich in keinem Zusammenhang zu der Vorlesemenge im gleichen Alter. Vor der Korrektur zeigt sich einzig der Trend, dass die Anzahl nominaler Komposita mit der Zeit des Vorlesens moderat positiv korreliert, d. h. die Zahl der Substantivkomposita steigt mit zunehmender Vorlesedauer.

10.2.6 Elizitationsexperiment vs. Spontansprache

Um festzustellen, ob es einen Zusammenhang zwischen der Anzahl korrekter Ableitungen im Derivationsexperiment im Alter von 8 Jahren und den Häufigkeiten in Lemmas der spontanen Derivationen im Alter von 4 Jahren bei den 16 teilnehmenden Kindern gibt, wurde ebenfalls eine nichtparametrische Korrelation nach Spearman errechnet, mit dem Ergebnis, dass die Anzahl korrekter -er-Ableitungen im Experiment bei Nomina Agentis ($r_s = 0,726, p = 0,01$) und Nomina Instrumenti ($r_s = -0,506, p = 0,05$) stark positiv korrelierte. Demnach besteht eine statistisch signifikante Korrelation zwischen der Anzahl derivierter Substantive im Alter von 4 Jahren und der Anzahl korrekter -er-Ableitungen beim Test mit 8 Jahren. Je mehr derivierte Substantive die Kinder im Alter von 4 Jahren bildeten, desto mehr korrekte -er-Derivationen bildeten sie auch 4 Jahre später im Experiment. Bei einer Einschränkung auf -er-suffigierte Substantive im gleichen Zeitraum konnte hingegen kein Zusammenhang festgestellt werden, ebenso wenig wie für die Korrelationen zur Anzahl korrekter -ung-Ableitungen oder zur Gesamtanzahl korrekter Derivationen im Experiment.

11 Konklusion

Kinder entdecken rund um ihren zweiten Geburtstag die Möglichkeit der Komposition (und Zerlegung), sowie einige frequente Ableitungsmuster, dazu zählen insbesondere die Partikelverbbildung, nominale Suffigierung und morphologische Konversion, wie in den Studien von Korecky-Kröll et al. (2017) für Komposita und von Sommer-Lolei et al. (2021) für Derivationen festgestellt. Im Alter von 3 Jahren leiten die Kinder vor allem Verben und Substantive erfolgreich ab, bilden nominale Komposita und synthetische Komposita, sowie in wesentlich geringerem Umfang auch Adjektivkomposita und -derivationen.

Die Untersuchung der spontanen Äußerungen von HSES- und LSES-Kindern ab 3 bis knapp 5 Jahre zeigt, dass die morphologischen Strukturen bei Kindern mit durchschnittlich 3;1 Jahren bereits so weit entwickelt sind, dass sich bei einem Vergleich zweier Datenzeitpunkte (3;1–3;4 zu 4;4–4;8) keine signifikanten Veränderungen in den für CS wesentlichsten Wortklassen der Verben und Substantive feststellen lassen. Es zeigt sich zwar eine generelle Zunahme der Lemmafrequenzen in den Kategorien der derivierten Substantive, Verben und Adjektive und auch bei Substantivkomposita, synthetischen Komposita und Adjektivkomposita vom früheren zum späteren Zeitpunkt, signifikante Veränderungen sind jedoch nur in der sich noch entwickelnden Kategorie der derivierten Adjektive feststellbar.

Die Komposition ist auch im Alter zwischen 3 und 5 Jahren bei Substantiven die vorrangig genutzte Wortbildungsart im Deutschen in CS und CDS. Danach folgen derivierte Substantive und synthetische Komposita, die den kleinsten Teil in diesem Alterszeitraum einnehmen. Insgesamt überwiegen nominale Komposita in Bezug auf ihre Variation (= Lemmafrequenz). Die überwältigende Mehrheit der Verben wird mit Partikeln gebildet, während alle anderen Bildungsmittel in der frühen Kindersprache und in CDS nur eine Randposition einnehmen. Das wesentlichste Wortbildungsmuster zur Bildung von Adjektiven ist die Suffigierung, danach folgen Perfektpartizipien, die als Adjektive verwendet werden und Zusammensetzung.

Ob in der Kindersprache wortartbewahrende Derivationen vor wortartverändernden auftreten, hängt von der untersuchten Wortklasse ab. Die ersten Wortbildungsmuster, die in den drei longitudinal erhobenen Kindern Jan, Kathi und Lena zuerst auftreten, sind deverbale Präfigierungen (mit trennbarer Partikel) und Suffigierungen. Während die ersten Suffigierungen bei Adjektiven und Substantiven zu den wortartverändernden Derivationen zählen, modifizieren Partikel verbale Basen und sind somit wortartbewahrend. Dies entspricht der generell festgestellten Tendenz in CS und CDS, da die meisten abgeleiteten Substantive einen Wortartwechsel vollziehen, während der überwiegende Teil derivierter Verben seiner Wortart treu bleibt. Wird der monatliche Zeitraum der Emergenz und nicht das zuerst auftretende Lexem als Kriterium herangezogen, so lässt sich feststellen, dass die Kinder zur Bildung von Verben und Substantiven im gleichen Altersmonat sowohl denominal als auch deverbale ableiten und daher wortartbewahrende und -verändernde Derivationen rund um den zweiten Geburtstag bilden.

In Bezug auf die Erfüllung des Mini-Paradigma-Kriteriums im Zeitraum von 3;1 bis 4;11 bei Jan in Erweiterung zur Untersuchung von Sommer-Lolei et al. (2021: 127–130) werden vor allem im Bereich suffigierter Verben und Adjektive weitere Wortbildungsmuster in CS potenziell produktiv.

Das von Sommer-Lolei et al. (a.a.O.) bis 3;0 nicht behandelte Wortbildungsmuster der impliziten Ableitung wird im Alter von 2 Jahren potenziell produktiv. Bis 4;11 erfüllt Jan das Kriterium außerdem für das Adjektivsuffix *-lich* (4;6) und die verbalen Suffixe *-ier* (3;6) und *-el* (3;7), jedoch nicht für *-er*, das bis 4;11 weiterhin partiell erfüllt wird. Eine teilweise Erfüllung des Kriteriums ist überdies für das verbale Zirkumfix *ent-...-ig* (3;10), das Suffix *-isch* (4;0) und das Präfix *un-* (3;11) zur Ableitung von Adjektiven feststellbar. Jans Wortschatz erweitert sich demnach zwischen 3;0 und 4;11 vor allem im Bereich der suffigierten Verben und derivierten Adjektive.

Morphologisch komplexe Lexeme und Basen kommen in allen Spontansprachkorpora ab 3 Jahren vor. Komplexität kann jedoch nicht der alleinige und ausschlaggebende Faktor für oder gegen die Bildung eines Lexems sein, da die Kinder statt einfacher Derivationen (= ein Wortbildungsmuster) häufig synthetische Komposita

im Rahmen des durchgeführten Elizitationsexperiments bilden (bei Nomina Instrumenti), die sowohl den derivationellen Prozess als auch die Komposition beinhalten. Entweder beruht dies darauf, dass diese Bildungen in CDS häufiger vorkommen und den Kindern daher geläufiger sind (Rasenmäher, Dosenöffner, Taschenrechner) oder dass die Kinder es für nötig befinden, die Art eines Instruments näher zu spezifizieren. Generell antworten die HSES-Kinder im Experiment häufiger als die gleichaltrigen LSES-Kinder und produzieren dabei weniger Fehler. Da die Leistung in einem Experiment größer ist, als bei der spontanen Sprachproduktion, aufgrund eines notwendigen höheren Sprachbewusstseins (*language awareness*), deutet dies auf eine höhere sprachliche Fähigkeit im Sinne der Sprachentwicklungstheorie (*Theory of redescription*) von Karmiloff-Smith (1996). Es hat sich außerdem gezeigt, dass die Testvariable der Nomina Agentis bei Kindern im Alter von 8 Jahren eher vernachlässigbar ist und ein größeres Augenmerk auf die Bildung abstrakter Substantive gelegt werden sollte.

Hypothese 1

Hypothese 1 (H1) kann in Bezug auf den Einfluss des sozioökonomischen Status' der Familie auf die morphologische Entwicklungsgeschwindigkeit des Kindes bestätigt werden. Die statistischen Berechnungen der Mittelwertdifferenzen ergeben, dass Kinder mit höherem sozioökonomischen Status im Vergleich zu Kindern mit niedrigerem SES im Durchschnitt mehr Lemmas und Tokens in allen Wortbildungsmustern bilden. Dies konnte ebenfalls für die Hauptbezugspersonen festgestellt werden. Der Wortschatz der HSES-Kinder und -Hauptbezugspersonen ist in allen Wortklassen und Wortbildungsarten zu allen Datenzeitpunkten diverser und weist höhere Tokenfrequenzen auf als jener der LSES-Kinder und -Hauptbezugspersonen, was sich speziell bei morphologisch komplexeren Bildungen zeigt. Überdies vergrößern sich die Differenzen in Lemmas bei derivierten Substantiven und Verben im Erwerbsverlauf (bis 4;11) bei einem direkten Vergleich der HSES- vs. LSES-Kinder. Diese Differenzen zeigen sich ebenfalls noch im Alter von 8 Jahren beim Derivationstest. Die morphologische Komplexität der Lexeme betreffend, treten in beiden Gruppen komplexe Lexeme ab der ersten Aufnahme auf, dennoch zeigen sich höhere Lemma- und Tokenfrequenzen bei den HSES-Dyaden. Es ist jedoch festzuhalten, dass einfache Ableitungsbasen in CS und

CDS tendenziell bevorzugt werden. LSES-Kinder und ihre Hauptbezugspersonen verwenden überdies insgesamt weniger verschiedene Wortbildungsmuster, insbesondere weniger unterschiedliche Suffixe zur Bildung von Substantiven, weniger verschiedene Verbpartikel und weniger mehrgliedrige Komposita.

Im Elizitationsexperiment hat sich gezeigt, dass die HSES-Kinder im Vergleich zu ihren gleichaltrigen LSES-Peers durchschnittlich häufiger antworten, und ebenfalls mehr korrekte Antworten produzieren, insbesondere bei Nomina Instrumenti und Nomina Acti/Actionis. Überdies verwenden sie häufiger die korrekte Ableitungsbasis, den korrekten Umlaut und das korrekte Suffix und bilden weniger semi- und inkorrekte Formen.

Hypothese 2

Neben sprachtypologischen Aspekten nehmen zahlreiche weitere Faktoren Einfluss auf den Morphologieerwerb von Kindern, die ihre Erstsprache erwerben. Die Präferenz für morphologisch natürlichere Wortbildungsmuster zeigt sich in CS an dem frühen Auftreten der Partikelverbbildung, sowie -er-Suffigierung zur Ableitung von Substantiven, da diese Bildungsmuster viele der Natürlichkeitsparameter und Kriterien erfüllen, was im Folgenden erläutert wird.

In Bezug auf die Partikelverbbildung kann festgestellt werden, dass trennbare Präfixe (= Partikel) betont und daher prosodisch und positionell salient sind, sie stehen in einer Äußerung oft in satzfinaler Position und erfüllen daher das Kriterium der positionellen Salienz im Sinne des *Recency effect*, weiters haben sie in der Regel eine bestimmte uniforme Funktion und sind morphotaktisch und morphosemantisch transparent, da die Partikel die Bedeutung ihrer Basis weiter spezifiziert und zu keiner Opazität führt, vgl. z. B. *Ich halte etwas unter* (unter#halt-), das morphosemantisch transparenter ist im Vergleich zu dem präfigierten Verb in *Ich unterhalte jemanden* (unter-halt-). Überdies sind Partikel meist eine Silbe lang, das Wortbildungsprodukt ist meist binär und zusammen mit der Basis nicht länger als drei Silben. Ähnlich verhält es sich mit dem Suffix -er, abgesehen davon, dass es unbetont, polysem und prosodisch nicht salient ist. In Bezug auf diese und andere früh in CS vorkommenden und häufig verwendeten Wortbildungsmuster, wie etwa die nominale Komposition, gilt H2 als vollkommen bestätigt, da morphologisch

natürlichere Wortbildungsmuster in CS bevorzugt verwendet werden und sich früher und schneller entwickeln. Ebenso für die wenig frequenten Wortbildungsmuster der Zirkumfigierung und Modifikation (z. B. reiner Ablaut, wie in *Sprung*), die generell als wenig natürlich gelten und in CS nur in vereinzelten Bildungen vorkommen.

Was jedoch früh auftretende Konversionen bei Verben und Substantiven betrifft, kann die Hypothese nicht bestätigt werden. Konversionen sind aufgrund fehlender Ikonizität und Binarität, sowie vieler weiterer Faktoren der Natürlichkeitstheorie zufolge unnatürlich und sollten deshalb erst später in der Kindersprache auftreten. Es ist anzunehmen, dass das frühe Auftreten von Konversionen ihrer Einfachheit geschuldet ist. H2 kann jedoch in Bezug auf die Verwendungshäufigkeit zumindest teilweise bestätigt werden, weil Konversionen in geringeren Tokenfrequenzen verwendet werden als natürlichere Wortbildungsmuster.

Hypothese 3

Hypothese 3, die Zusammenhänge zwischen der Menge des Inputs und/oder des Medienkonsums und der Sprachproduktion der Kinder in der Spontansprache und im Derivationsexperiment erwartet, wird im Folgenden überprüft.

Bei den Spontansprachdaten konnten sowohl bei HSES als auch bei LSES in Bezug auf die Lemma- und Tokenfrequenzen derivierter und zusammengesetzter Lexeme zwischen den Kindern und ihren Hauptbezugspersonen signifikante Unterschiede zwischen CS und CDS festgestellt werden (vgl. Kap. 5.8.2.1). Die berechneten Korrelationen, die die Menge des Inputs in Minuten pro Woche in Relation zu den Ergebnissen des Derivationstests (vgl. Kap. 10.1.2) und der spontan gebildeten Substantive, Verben und Adjektive setzten (vgl. Kap. 10.2.2), zeigten keinen signifikanten Zusammenhang. Der erwartete Effekt, dass die Menge des Inputs mit der Sprachproduktion der Kinder korrelieren würde, konnte demnach für diese Daten nicht festgestellt werden, es zeigten sich auch keine Trends.

Es hat sich bereits in einer Untersuchung von Blaschitz (2019: 98) mit denselben Kindern des INPUT-Projekts gezeigt, dass Kinder, denen mit 4 Jahren mehr vorgelesen wurde, im gleichen Alter gute Ergebnisse im passiven Wortschatztest des PPVTs (*Peabody Picture Vocabulary Test*) erzielten und auch in ihrer

Erzählweise bei der Wiedergabe der *Frogstory*⁴⁸ kompetenter waren. Auch bei dem in dieser Arbeit durchgeführten Elizitationsexperiment wurden statistisch signifikante Zusammenhänge zwischen der Vorlesemenge im Alter von 4 Jahren und der Anzahl korrekter Basen, Suffixe und Derivationen im Alter von 8 Jahren festgestellt. Kinder, denen im Alter von 4 Jahren mehr vorgelesen wurde, bildeten im Alter von 8 Jahren nicht nur signifikant mehr korrekte Derivationen im Elizitationsexperiment, sondern verwendeten ebenfalls mehr korrekte Suffixe und Basen (vgl. Kap. 10.1.5). Statistisch signifikante Zusammenhänge zwischen der Menge des Büchervorlesens in Minuten pro Woche und der Anzahl spontansprachlicher Derivationen und Komposita bei Substantiven, Verben und Adjektiven konnten hingegen nicht festgestellt werden (vgl. Kap. 10.2.5).

Überdies wirkte sich der geringere Fernsehkonsum im Alter von 4 Jahren positiv auf die Lesebereitschaft im Alter von 8 Jahren aus. Wie Blaschitz (2019: 98) feststellen konnte, lasen jene Kinder mit 8 häufiger, die mit 4 weniger ferngesehen haben. Eine Korrelation mit dem Fernsehkonsum konnte für die Auswertungen der Ergebnisse des Elizitationstests nicht festgestellt werden (vgl. Kap. 10.1.3). Die Untersuchung der Anzahl spontansprachlicher Derivationen und Komposita hingegen ergab einen statistisch signifikanten Zusammenhang zwischen der Menge des Medienkonsums, der neben anderen auch Fernsehen enthält, und der Bildung derivierter, vor allem suffigierter Substantive. Kinder, die mit 4 Jahren mehr Neue Medien konsumieren, also auch mehr fernsehen, bilden signifikant weniger derivierte, und insbesondere suffigierte Substantive im gleichen Alter (vgl. Kap. 10.2.3).

H3 wird in Bezug auf einen Zusammenhang zwischen der Inputmenge durch die Hauptbezugsperson und der Kindersprache sowohl für die Spontansprache als auch für die Testergebnisse verworfen. Ebenso konnte kein Effekt des Hörbuch- und Radiokonsums auf CS nachgewiesen werden. Allerdings wurden einerseits negative Auswirkungen der Menge des Medienkonsums (YISF) auf das spontansprachliche Verhalten der Kinder im Alter von 4 Jahren festgestellt, da Kinder mit mehr Fernsehkonsum weniger derivierte, im Speziellen weniger

⁴⁸ Die so bezeichnete *Frogstory* (Mayer 1969) ist eine Bildgeschichte, die in der Sprachwissenschaft häufig genutzt wird, um die Erzählkompetenz eines Sprechers oder einer Sprecherin festzustellen. In diesem Fall von 4-jährigen Kindern.

suffigierte Substantive spontan bildeten, sowie andererseits positive Effekte des Büchervorlesens (mit 4 Jahren) auf die Ergebnisse des Derivationstests (mit 8 Jahren), da jene Kinder, denen mehr vorgelesen wurde, nicht nur mehr korrekte Derivationen beim Test bildeten, sondern auch häufiger das korrekte Suffix oder die korrekte Basis zur Bildung verwendeten. Es zeigten sich auch die Trends, dass a) Kinder, die mit 4 Jahren mehr Neue Medien konsumieren (YISF) im Elizitationsexperiment im Alter von 8 Jahren dazu tendieren, häufiger nicht zu antworten und seltener den korrekten Umlaut zu bilden, und b) eine längere Vorlesedauer (mit 4 Jahren) eine höhere Anzahl nominaler Komposita in CS begünstigt. Insofern kann H3 teilweise bestätigt werden.

12 Ausblick

Wie in dieser Arbeit gezeigt werden konnte, gibt es viele wesentliche Faktoren, die im Zuge des Spracherwerbs Einfluss nehmen auf die morphologische Entwicklung von jungen Kindern. Es gilt insbesondere, jene Variablen zu fördern, die dazu beitragen können, dass Eltern oder andere Hauptbezugspersonen ihre Kinder optimal auf ihrer sprachlichen Entwicklungsreise begleiten und unterstützen können. Diese Untersuchung lässt erkennen, dass ein abwechslungsreich gestalteter Input im Sinne höherer Lemmafrequenzen auch zu mehr Varietät im Output der Kinder führt und in direktem Zusammenhang zu diesem steht. Eine zusätzliche Erweiterung dieser Studie in Bezug auf mögliches *fine-tuning* der Hauptbezugsperson im Erwerbsverlauf, auf das im Rahmen dieser Arbeit nicht eingegangen werden konnte, wäre in jedem Fall anzustreben (vgl. *fine-tuning* bei Komposita Korecky-Kröll et al. 2017: 24 f.; bei synthetischen Komposita Dressler, Sommer-Lolei, et al. 2019: 421). Überdies wirken sich insbesondere Büchervorlesen und Medienkonsum, wie Fernsehen und Onlinespiele (YouTube, Smartphone, Internet) in frühen Jahren auf die sprachliche Entwicklung aus. Bücher vorgelesen zu bekommen, fördert nicht nur die Bindung zur Hauptbezugsperson und die optischen und akustischen Sinneswahrnehmungen, was insbesondere auch für legasthene Kinder vorteilhaft ist, sondern regt zudem die Fantasie an und legt wichtige Grundsteine für die lexikalische Weiterentwicklung der Kinder. Übermäßiges Sich-selbst-überlassen-sein durch zunehmenden Medienkonsum

(YouTube, Internet, Smartphone, Fernsehen) wirkt sich hingegen nachteilig auf die genannten Bereiche aus.

Des Weiteren sollten Kinder, die aus Familien mit niedrigerem sozioökonomischen Status kommen, vor allem im Bereich nominaler und verbaler Ableitungen bereits früh in ihrer sprachlichen Entwicklung gezielt gefördert werden, sodass diese etwaige Differenzen zu Kindern mit höherem sozioökonomischen Status mit der Zeit aufholen können. Ein frühzeitig durchgeführter Derivationstest, der auf das entsprechende Alter der Kinder angepasst ist, dient dazu den Entwicklungsstand im Bereich der Derivationsmorphologie festzustellen und kann auf diese Weise dazu beitragen jene Bildungsmuster zu erkennen, die einem Kind tendenziell mehr Probleme bereiten. Insbesondere ab dem Eintritt in den Kindergarten sollte das explizite Üben abgeleiteter Wörter im Rahmen von spielerischen Interaktionen in den Tagesablauf integriert werden, um die Sprachentwicklung der Kinder auf diesem Gebiet zu fördern.

In jedem Fall wäre eine zusätzliche Untersuchung der an die Kinder gerichteten Sprache der Kindergartenpädagoginnen interessant (siehe dazu Anhang 11 Tabelle 65) und inwieweit sich ein früher oder später Eintritt in den Kindergarten auf die sprachliche Entwicklung der HSES- und vor allem der LSES-Kinder auswirkt.

Überdies ist ein Vergleich der L1- und L2-Kinder anzustreben, um die Testergebnisse und Spontansprachdaten gegenüberstellend analysieren zu können. Die Ergebnisse können Aufschluss darüber geben, welche sprachtypologischen Eigenschaften oder grammatischen Merkmale des Deutschen L2-lernenden Kindern Probleme bereiten. Dadurch wäre es möglich mehrsprachige Kinder von Beginn an speziell zu fördern und in ihrem Spracherwerb zu begleiten, damit sie die Sprache früher und sicherer beherrschen, als es sonst der Fall wäre. Für türkisch-deutsche Kinder, wie jene des INPUT-Projekts, könnte beispielsweise ein gezieltes Umlauttraining förderlich sein.

Die Analyse der Spontansprachdaten ist, wie in dieser Dissertation geschehen, in jedem Fall durch experimentelle Untersuchungen zu ergänzen. Dies war im Rahmen des INPUT-Projekts durch zahlreiche Wortschatz- und Grammatiktests bereits der Fall. Nur auf diesem Weg wird die Kindersprache aus verschiedenen

Blickwinkeln betrachtet. Außerdem wäre die zusätzliche Messung von Reaktionszeiten wünschenswert, um beispielsweise den Einfluss von Frequenz oder Vertrautheit (Familiarität) bestimmter Wortbildungsmuster sichtbar zu machen. Ein weiterer Vorteil zusätzlicher experimenteller Forschungen besteht darin, dass neben der in den Korpora untersuchten Sprachproduktion auch das Sprachverständnis miteinbezogen werden kann.

Die Trends, die sich bei den Korrelationen zu den Ergebnissen des Elizitationsexperiments ergeben haben, sollten mit einer weiteren Testreihe und einer größeren Teilnehmeranzahl wiederholt werden, um zu sehen, ob sich diese bestätigen.

Auf Basis der vorliegenden Untersuchungen ist ein Vergleich der Ergebnisse dieser Arbeit zum Spracherwerb des Deutschen mit jenen Resultaten, die in anderen Sprachen durchgeführt wurden (vgl. Dressler et al. 2017; Mattes et al. 2021), anzustreben. Ein cross-linguistischer Vergleich des Erwerbs abgeleiteter Adjektive in attributiver und prädikativer Verwendung im Deutschen und Schwedischen ist bereits in Vorbereitung. Ebenso ist eine Untersuchung der Flexionsfehler bei Substantivkomposita und derivierten Substantiven und Verben anzustreben, um zu sehen, was sich unter welchen Bedingungen schneller in der Kindersprache entwickelt.

13 Literaturverzeichnis

- Aitchison, Jean (1994) *Words in the Mind. An Introduction to the Mental Lexicon*. (2. Ed.). Oxford, Cambridge MA: Blackwell.
- Bartke, Susanne (1998) *Experimentelle Studien zur Flexion und Wortbildung. Pluralmorphologie und lexikalische Komposition im unauffälligen Spracherwerb und im Dysgrammatismus*. Tübingen: Niemeyer.
- Barz, Irmhild (2009) Die Wortbildung. In *Die Grammatik, Band 4, 8. überarbeitete Aufl.*, Duden [Hrsg.], 634–762. Mannheim, Zürich: Dudenverlag.
- Bates, Elizabeth & MacWhinney, Brian (1987) Competition, Variation and Language Learning. In *Mechanisms of language acquisition. Proceedings of the 20th Carnegie-Mellon Symposium on Cognition*, Brian MacWhinney [Hrsg.], 157–193. Hillsdale NJ: Lawrence Erlbaum.
- Behrens, Heike (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>
- Berman, Ruth Aronson (2004) Between emergence and mastery: The long developmental route of language acquisition. In *Language Development across Childhood and Adolescence*, Ruth Aronson Berman [Hrsg.], 9–34. Amsterdam: John Benjamins.
- Bittner, Dagmar, Dressler, Wolfgang U. & Kilani-Schoch, Marianne [Hrsg.] (2003) *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. [Studies on Language Acquisition 21]. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110899832>
- Blaschitz, Verena (2019) Eine longitudinale Perspektive auf Lesen und Erzählen als Teil der sprachlichen Bildung. *ÖDaF-Mitteilungen* 35(1+2): 85–103.
- Brzić, Katharina (2007) *Das geheime Leben der Sprachen: Gesprochene und verschwiegene Sprachen und ihr Einfluss auf den Spracherwerb in der Migration*. [Internationale Hochschulschriften]. Münster [u.a.]: Waxmann.
- Bybee, Joan L. (1995) Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455. <https://doi.org/10.1080/01690969508407111>
- Bybee, Joan L. (2006) *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, Joan L. (2010) *Language, Usage and Cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Chomsky, Noam (1965) *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

- Chomsky, Noam (1986) *Knowledge of language: Its nature, origin and use*. New York NY [u.a.]: Praeger.
- Chomsky, Noam (2002) *On Nature and Language*. Cambridge [u.a.]: Cambridge University Press.
- Chomsky, Noam & Lasnik, Howard (1993) The Theory of Principles and Parameters. In *Syntax: Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research* (Vol. 1). [Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 9.1], Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann [Hrsg.], 506–569. Berlin, New York: De Gruyter Mouton.
- Clahsen, Harald, Marcus, Gary, Bartke, Susanne & Wiese, Richard (1996) Compounding and inflection in German child language. *Yearbook of Morphology 1995*: 115–142.
- Clahsen, Harald, Rothweiler, Monika, Woest, Andreas & Marcus, Gary F. (1992) Regular and irregular inflection in the acquisition of German noun plurals. *Cognition 45*(3): 225–255. [https://doi.org/10.1016/0010-0277\(92\)90018-D](https://doi.org/10.1016/0010-0277(92)90018-D)
- Clark, Eve V. (2009) *First language acquisition*. (2. Ed.). Cambridge [u.a.]: Cambridge Univ. Press.
- Cordes, Anne-Kristin (2014) *The role of frequency in children's learning of morphological constructions*. [Language in performance 48]. Tübingen: Narr Verlag.
- Cornell, Edward H. & Bergstrom, Linda I. (1983) Serial-position effects in infants' recognition memory. *Memory & Cognition 11*: 494–499. <https://doi.org/10.3758/BF03196986>
- Czinger, Christine, Korecky-Kröll, Katharina, Uzunkaya-Sharma, Kumru & Dressler, Wolfgang U. (2015) Wie beeinflusst der sozioökonomische Status den Erwerb der Erst- und Zweitsprache? Wortschatzerwerb und Geschwindigkeit im NP/DP-Erwerb bei Kindergartenkindern im türkisch-deutschen Kontrast. In *Deutsche Grammatik in Kontakt. Deutsch als Zweitsprache in Schule und Unterricht* (Vol. 64), Klaus-Michael Köpcke & Arne Ziegler [Hrsg.], 207–240. Berlin [u.a.]: De Gruyter. <https://doi.org/10.1515/9783110367171-010>
- Don, Jan (2005) On Conversion, Relisting and Zero-derivation. A comment on Rochelle Lieber: English word-formation processes. *SKASE Journal of theoretical linguistics 2*(2): 2–16.
- Donegan, Patricia J. & Stampe, David (1979) The study of Natural Phonology. In *Current Approaches to Phonological Theory*, Daniel A. Dinnsen [Hrsg.], 126–173. Bloomington: Indiana University Press.
- Dressler, Wolfgang U. (1995) Wealth and poverty of functional analyses, with special reference to functional deficiencies. In *Form and Function in*

- Language. Proceedings of the 1st Rasmus Rask Colloquium*, Sharon Millar & Jacob Mey [Hrsg.], 11–39. Odense: Odense University Press.
- Dressler, Wolfgang U. (2005) Word-formation in natural morphology. In *Handbook of word-formation*, Pavol Štekauer & Rochelle Lieber [Hrsg.], 267–284. Dordrecht: Springer.
- Dressler, Wolfgang U. (2006) Introduction: Natural Morphology. *Folia Linguistica* 40(1–2): 1–6. <https://doi.org/10.1515/flin.40.1-2.1>
- Dressler, Wolfgang U. (2007) Introduction. In *Typological perspectives on the acquisition of noun and verb morphology*. [Antwerp Papers in Linguistics 112], Sabine Laaha & Steven Gillis [Hrsg.], 3–9. Antwerpen: Antwerp University. <https://homepage.univie.ac.at/sabine.laaha/links/apil112.pdf>
- Dressler, Wolfgang U. (2009) Natural Phonology as part of Natural Linguistics. *Poznań Studies in Contemporary Linguistics* 45(1): 33–42. <https://doi.org/10.2478/v10010-009-0003-9>
- Dressler, Wolfgang U., Christofidou, Anastasia, Gagarina, Natalia, Korecky-Kröll, Katharina & Kilani-Schoch, Marianne (2019) Morphological blind-alley developments as a theoretical challenge to both usage-based and nativist acquisition models. *Italian Journal of Linguistics* 31(2): 107–140. <https://doi.org/10.26346/1120-2726-138>
- Dressler, Wolfgang U. & Karpf, Annemarie (1995) The theoretical relevance of pre- and protomorphology in language acquisition. *Yearbook of Morphology 1994*: 99–122.
- Dressler, Wolfgang U., Ketrez, F. Nihan & Kilani-Schoch, Marianne [Hrsg.] (2017) *Nominal Compound Acquisition*. Amsterdam: John Benjamins.
- Dressler, Wolfgang U. & Kilani-Schoch, Marianne (2016) Natural Morphology. In *The Cambridge Handbook of Morphology*, Andrew Hippisley & Gregory T. Stump [Hrsg.], 356–389. Cambridge: Cambridge Univ. Press. <https://doi.org/10.1017/9781139814720.014>
- Dressler, Wolfgang U., Kilani-Schoch, Marianne & Klampfer, Sabine (2003) How does a child detect morphology? Evidence from production. In *Morphological Structure in Language Processing*. [Trends in Linguistics. Studies and Monographs 151], R. Harald Baayen & Robert Schreuder [Hrsg.], 391–425. Berlin & New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110910186.391>
- Dressler, Wolfgang U. & Korecky-Kröll, Katharina (2015) Evaluative morphology and language acquisition. In *Edinburgh Handbook of Evaluative Morphology*, Nicola Grandi & Lívia Körtvélyessy [Hrsg.], 134–141. Edinburgh: Edinburgh Univ. Press. <https://doi.org/10.3366/j.ctt1g09zqx>
- Dressler, Wolfgang U., Lettner, Laura E. & Korecky-Kröll, Katharina (2010) First language acquisition of compounds with special emphasis on early German child language. In *Cross-Disciplinary Issues in Compounding*, Sergio Scalise

- & Irene Vogel [Hrsg.], 323–344. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.311.24dre>
- Dressler, Wolfgang U., Lettner, Laura E. & Korecky-Kröll, Katharina (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], Ferenc Kiefer, Mária Ladány & Péter Siptár [Hrsg.], 237–264. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.322.11dre>
- Dressler, Wolfgang U., Mayerthaler, Willi, Panagl, Oswald & Wurzel, Wolfgang Ullrich (1987) *Leitmotifs in Natural Morphology*. [Studies in Language Companion Series 10]. Amsterdam: John Benjamins Publishing Company.
- Dressler, Wolfgang U., Sommer-Lolei, Sabine, Korecky-Kröll, Katharina, Argus, Reili, Dabašinskienė, Ineta, Kamandulytė-Merfeldienė, Laura, Ijäs, Johanna Johansen, Kazakovskaya, Victoria V., Laalo, Klaus & Thomadaki, Evangelia (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>
- Dressler, Wolfgang U., Stephany, Ursula, Aksu-Koç, Ayhan & Gillis, Steven (2007) Discussion and Conclusion. In *Typological perspectives on the acquisition of noun and verb morphology*. [Antwerp Papers in Linguistics 112], Sabine Laaha & Steven Gillis [Hrsg.], 67–71. Antwerpen: Antwerp University.
- Eisenberg, Peter (2013) *Das Wort. Grundriss der deutschen Grammatik, Bd. 1, 4., aktualisierte und überarb. Aufl.* Stuttgart [u.a.]: Metzler.
- Elman, Jeffrey L., Bates, Elizabeth A., Johnson, Mark H., Karmiloff-Smith, Annette, Parisi, Domenico & Plunkett, Kim (1996) *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge MA [u.a.]: MIT Press.
- Elsen, Hilke (1999) *Ansätze zu einer funktionalistisch-kognitiven Grammatik. Konsequenzen aus Regularitäten des Erstspracherwerbs*. [Linguistische Arbeiten 403]. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110915976>
- Elsen, Hilke (2014) *Grundzüge der Morphologie des Deutschen, 2. aktualisierte Aufl.* [De Gruyter Studium]. Berlin, Boston: De Gruyter.
- Erben, Johannes (2006) *Einführung in die deutsche Wortbildungslehre. 5., durchges. und erg. Aufl.* Berlin: Schmidt.
- Eschenlohr, Stefanie (1999) *Vom Nomen zum Verb. Konversion, Präfigierung und Rückbildung im Deutschen*. [Germanistische Linguistik Monographien 3]. Hildesheim [u. a.]: Olms.
- Eysenck, Michael W. & Keane, Mark T. (2010) *Cognitive Psychology. A Student's Handbook*. (6. Ed.). Hove, New York: Psychology Press.

- Fleischer, Wolfgang & Barz, Irmhild [Hrsg.] (2012) *Wortbildung der deutschen Gegenwartssprache* (4. Auflage). [De-Gruyter-Studium]. Berlin, Boston: De Gruyter.
- Fromkin, Victoria A. [Hrsg.] (2013) *Speech Errors as Linguistic Evidence*. [Janua Linguarum. Series Maior 77]. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110888423>
- Gaeta, Livio (2010) Synthetic compounds. With special reference to German. In *Cross-disciplinary Issues in Compounding*. [Current Issues in Linguistic Theory 311], Sergio Scalise & Irene Vogel [Hrsg.], 219–236. Amsterdam: John Benjamins.
- Ganzeboom, Harry B. G. & Treiman, Donald J. (1996) Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research* 25: 201–239.
- Gathercole, Virginia C. Mueller & Thomas, Enlli Môn (2007) Factors contributing to language transmission in bilingual families: The core study--Adult Interviews. In *Language Transmission in Bilingual Families in Wales*, Virginia C. Mueller Gathercole [Hrsg.], 59–181: Welsh Language Board.
- Gillis, Steven (2003) A case study of the early acquisition of verbs in Dutch. In *Development of verb inflection in first language acquisition. A cross-linguistic perspective*. [SOLA Studies on Language Acquisition 21], Dagmar Bittner, Wolfgang U. Dressler & Marianne Kilani-Schoch [Hrsg.], 171–203. Berlin, Boston: De Gruyter, Inc.
- Glück, Helmut [Hrsg.] (2010) *Metzler Lexikon Sprache* (4., aktualisierte und überarbeitete Auflage). Stuttgart, Weimar: Metzler.
- Harris, Margaret (1992) *Language Experience and Early Language Development. From Input to Uptake*. [Essays in Developmental Psychology]. Hove, Hillsdale NJ: Erlbaum.
- Hart, Betty & Risley, Todd R. (1995) *Meaningful differences in the everyday experience of young American children*. Baltimore MD: Paul H. Brookes Publishing Co.
- Haspelmath, Martin (1996) Word-class Changing Inflection and Morphological Theory. In *Yearbook of Morphology 1995*, Geert Booij & Jaap van Marle [Hrsg.], 43–66. Dordrecht: Kluwer.
- Hoff, Erika (2002) Causes and Consequences of SES-Related Differences in Parent-to-Child Speech. In *Socioeconomic Status, Parenting, and Child Development*. [Monographs in Parenting], Marc H. Bornstein & Robert H. Bradley [Hrsg.], 147–160. Mahwah NJ: Erlbaum.
- Jakobson, Roman (1971) *Selected Writings*. (Vol. II Word and Language). The Hague: De Gruyter Mouton. <https://doi.org/10.1515/9783110873269>

- Kalédaitė, Violeta & Savickienė, Ineta (2007) The role of a child's gender in language acquisition. *Estonian Papers in Applied Linguistics* 3: 285–297. <https://doi.org/10.5128/ERYa3.19>
- Karmiloff-Smith, Annette (1996) *Beyond modularity: A developmental perspective on cognitive science*. (1st MIT Press pbk. Ed.). [Learning, development, and conceptual change]. Cambridge, Mass.: MIT Press.
- Karpf, Annemarie (1990) *Selbstorganisationsprozesse in der sprachlichen Ontogenese: Erst- und Fremdsprachen*. Tübingen: Narr.
- Kauschke, Christina (2012) *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. [Germanistische Arbeitshefte 45]. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110283891>
- Kauschke, Christina & Klann-Delius, Gisela (2007) Characteristics of maternal input in relation to vocabulary development in children learning German. In *Frequency Effects in Language Acquisition. Defining the Limits of Frequency as an Explanatory Concept*. [Studies on Language Acquisition SOLA 32], Insa Gültzow & Natalia Gagarina [Hrsg.], 181–203. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110977905.181>
- Kelić, Maja & Dressler, Wolfgang U. (2019) The development of morphotactic and phonotactic word-initial consonant clusters in Croatian first-language acquisition. *Suvremena Lingvistika* 45(88): 179–200. <https://doi.org/10.22210/suvlin.2019.088.03>
- Kilani-Schoch, Marianne & Dressler, Wolfgang U. (2005) *Morphologie naturelle et flexion du verbe français* (Natürliche Morphologie und Flexion des französischen Verbs). Tübingen: Narr.
- Kluge, Friedrich (2011) *Etymologisches Wörterbuch der deutschen Sprache*. Bearbeitet von Elmar Seebold. (Vol. 25). Berlin, Boston: De Gruyter.
- Korecky, Paul C. (2015) CLANTOCSV [Computer Software].
- Korecky-Kröll, Katharina (2011) Der Erwerb der Nominalmorphologie bei zwei Wiener Kindern: Eine Untersuchung im Rahmen der Natürlichkeitstheorie. Dissertation, Universität Wien.
- Korecky-Kröll, Katharina & Dressler, Wolfgang U. (2007) Diminutives and Hypocoristics in Austrian German (AG). In *The acquisition of diminutives: a cross-linguistic perspective*, Ineta Savickienė & Wolfgang U. Dressler [Hrsg.], 207–230. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.43.09kor>
- Korecky-Kröll, Katharina & Dressler, Wolfgang U. (2015) Acquisition of German Adjective Inflection and Semantics by two Austrian Children. In *Semantics and Morphology of Early Adjectives in First Language Acquisition*, Elena Tribushinina, Maria D. Voeikova & Sabrina Noccetti [Hrsg.], 23–52. Newcastle upon Tyne: Cambridge Scholars Publishing.

- Korecky-Kröll, Katharina, Sommer-Lolei, Sabine & Dressler, Wolfgang U. (2017) Emergence and early development of German compounds. In *Nominal Compound Acquisition*. [Language Acquisition & Language Disorders 61], Wolfgang U.; Dressler, F. Nihan; Ketz & Marianne Kilani-Schoch [Hrsg.], 19–37. Amsterdam: Benjamins. <https://doi.org/10.1075/lald.61>
- Kühnhold, Ingeburg, Putzer, Oskar & Wellmann, Hans (1978) *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Dritter Hauptteil: Das Adjektiv*. Düsseldorf: Schwann.
- Kühnhold, Ingeburg & Wellmann, Hans (1973) *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache. Hauptteil 1: Das Verb*. Düsseldorf: Schwann.
- Laaha, Sabine (2004) Développement précoce de la morphologie verbale: une étude comparative sur l'acquisition de l'allemand autrichien et du français (Frühe Entwicklung der Verbmorphologie: eine vergleichende Studie zum Erwerb des österreichischen Deutsch und Französischen). Dissertation, Universität Wien.
- Laaha, Sabine & Bassano, Dominique (2013) On the role of input for children's early production of bare infinitives in German and French. Frequency, informativeness, salience. *Language, Interaction and Acquisition* 4(1): 70–90. <https://doi.org/10.1075/lia.4.1.04laa>
- Laaha, Sabine & Gillis, Steven [Hrsg.] (2007) *Typological perspectives on the acquisition of noun and verb morphology*. [Antwerp Papers in Linguistics 112]. Antwerpen: Antwerp University.
- Laaha, Sabine, Gillis, Steven, Kilani-Schoch, Marianne, Korecky-Kröll, Katharina, Xanthos, Aris & Dressler, Wolfgang U. (2007) Weakly inflecting languages: French, Dutch, and German. In *Typological perspectives on the acquisition of noun and verb morphology*. [Antwerp Papers in Linguistics 112], Sabine Laaha & Steven Gillis [Hrsg.], 21–33. Antwerpen: Antwerp University.
- Laaha, Sabine & Korecky-Kröll, Katharina (2009) Kodierung und Analyse von Kindersprache mit CHILDES. Beitrag präsentiert bei der 1. Jahrestagung des Zentrums Sprachwissenschaften, Bild- und Tondokumentation, Wien.
- Lettner, Laura E. (2008) Der Erwerb von Nominalkomposition und Diminutivbildung durch ein Wiener Kind. Diplomarbeit, Universität Wien.
- Lettner, Laura E., Korecky-Kröll, Katharina & Dressler, Wolfgang U. (2011) Charakteristika von deutschen Nominalkomposita in der protomorphologischen Phase des Erstspracherwerbs. In *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch: Perspektiven – Analysen – Anwendungen*, Hilke Elsen & Sascha Michel [Hrsg.], 191–208. Stuttgart: Ibidem-Verlag.

- Lieven, Elena (2010) Input and first language acquisition: Evaluating the role of frequency. *Lingua* 120(11): 2546–2556. <https://doi.org/10.1016/j.lingua.2010.06.005>
- Lieven, Elena (2014) First language development: A usage-based perspective on past and current research. *Journal of Child Language* 41: 48–63. <https://doi.org/10.1017/S0305000914000282>
- Longobardi, Emiddia, Rossi-Arnaud, Clelia, Spataro, Pietro, Putnick, Diane L. & Bornstein, Marc H. (2015) Children's acquisition of nouns and verbs in Italian: contrasting the roles of frequency and positional salience in maternal language. *Journal of Child Language* 42(1): 95–121. <https://doi.org/10.1017/S0305000913000597>
- MacWhinney, Brian (2000) *The CHILDES Project: Tools for Analyzing Talk*. Mahwah NJ: Lawrence Erlbaum.
- Manova, Stela & Dressler, Wolfgang U. (2005) The Morphological Technique of Conversion in the Inflecting-fusional Type. In *Approaches to Conversion / Zero-Derivation*, Laurie Bauer & Salvador Valera [Hrsg.], 67–101. Münster [u. a.]: Waxmann.
- Marchman, Virginia A. & Bates, Elizabeth (1994) Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language* 21(2): 339–366. <https://doi.org/10.1017/S0305000900009302>
- Mattes, Veronika (2018) Derivationsmorphologie und Wortarten im Erstspracherwerb des Deutschen. Mit einem Fokus auf dem Schuleintrittsalter. Habilitationsschrift, Universität Graz.
- Mattes, Veronika (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>
- Mattes, Veronika, Sommer-Lolei, Sabine, Korecky-Kröll, Katharina & Dressler, Wolfgang U. [Hrsg.] (2021) *The Acquisition of Derivational Morphology. A cross-linguistic perspective*. [Language Acquisition & Language Disorders 66]. Amsterdam, Philadelphia: John Benjamins. <https://doi.org/10.1075/lald.66>
- Mayer, Mercer (1969) *Frog, where are you?* New York: Dial Press.
- Mayerthaler, Willi (1981) *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
- Meibauer, Jörg (1995a) Neugebildete -er Derivate im Spracherwerb. Ergebnisse einer Langzeitstudie. *Sprache und Kognition* 14: 138–160.
- Meibauer, Jörg (1995b) Wortbildung und Kognition. Überlegungen zum deutschen -er Suffix. *Deutsche Sprache* 23: 97–123.
- Meibauer, Jörg (1999) Über Nomen-Verb-Beziehungen im frühen Wortbildungserwerb (On noun-verb relationships in early acquisition of word

- formation). In *Das Lexikon im Spracherwerb*, Jörg Meibauer & Monika Rothweiler [Hrsg.], 184–207. Tübingen: Francke.
- Meibauer, Jörg (2001) Sprachwandel und Spracherwerb – Eine Skizze am Fall der Wortbildung. In *Sprachgeschichte, Dialektologie, Onomastik, Volkskunde: Beiträge zum Kolloquium am 3./4. Dezember 1999 an der Johannes Gutenberg-Universität Mainz; Wolfgang Kleiber zum 70. Geburtstag*, Rudolf Bentzinger [Hrsg.], 177–155. Stuttgart: Steiner.
- Meibauer, Jörg, Guttfropf, Anja & Scherer, Carmen (2004) Dynamic aspects of German -er-nominals: a probe into the interrelation of language change and language acquisition. *Linguistics* 42(1): 155–193.
- Miestamo, Matti, Sinnemäki, Kaius & Karlsson, Fred [Hrsg.] (2008) *Language Complexity. Typology, contact, change*. [Studies in Language Companion Series 94]. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.94>
- Neugebauer-Kostenblut, Hanna (1914) Sprachliche Eigenbildungen meines Sohnes (My son's linguistic inventions). *Zeitschrift für Kinderforschung* 19: 174–181, 242–246, 362–370.
- OECD (1999) *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries*. OECD. 1999 Edition.
- Olsen, Susan (2017) Synthetic compounds from a lexicalist perspective. *Journal of Word Formation* 1(1): 17–45.
- Pan, Barbara A., Rowe, Meredith L., Singer, Judith D. & Snow, Catherine E. (2005) Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development* 76(4): 763–782. <https://doi.org/10.1111/1467-8624.00498-i1>
- Peirce, Charles S. (1965) *Collected Papers*. Cambridge: Harvard University Press.
- Peters, Ann M. (2001) Filler syllables: what is their status in emerging grammar? *Journal of Child Language* 28(1): 229–242.
- Rainer, Franz (2010) *Carmens Erwerb der deutschen Wortbildung* (Carmen's acquisition of German word formation). Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Ravid, Dorit, Dressler, Wolfgang U., Nir-Sagiv, Bracha, Korecky-Kröll, Katharina, Souman, Agnita, Rehfeldt, Katja, Laaha, Sabine, Bertl, Johannes, Basbøll, Hans & Gillis, Steven (2008) Core morphology in child directed speech. Crosslinguistic corpus analyses of noun plurals. In *Corpora in Language Acquisition Research: History, Methods, Perspectives*. [Trends in Language Acquisition Research 6], Heike Behrens [Hrsg.], 25–60. Amsterdam: John Benjamins.
- Roeper, Tom (2007) What frequency can do and what it can't. In *Frequency Effects in Language Acquisition. Defining the Limits of Frequency as an Explanatory Concept*. [Studies on Language Acquisition SOLA 32], Insa Gülzow & Natalia

- Gagarina [Hrsg.], 23–47. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110977905.23>
- Scheele, Anna F., Leseman, Paul P. M. & Mayo, Aziza Y. (2010) The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics* 31(1): 117–140. <https://doi.org/10.1017/S0142716409990191>
- Schipke, Christine S. & Kauschke, Christina (2011) Early word formation in German language acquisition: A study on word formation growth during the second and third years. *First Language* 31: 82. <https://doi.org/10.1177/0142723709359240>
- Schulz, Petra, Wymann, Karin & Penner, Zvi (2001) The early acquisition of verb meaning in German by normally developing and language impaired children. *Brain and Language* 77: 407–418.
- Serra i Raventós, Miquel (2014) Interfaces between language components: Bootstrapping in sensitive periods after reaching a critical mass of usage. *Anuario de Psicología/The UB Journal of Psychology* 44(1): 85–93.
- Smolka, Eva, Libben, Gary & Dressler, Wolfgang 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>
- Smolka, Eva, Preller, Katrin H. & Eulitz, Carsten (2014) ‚Verstehen‘ (‚understand‘) primes ‚stehen‘ (‚stand‘): Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *Journal of Memory and Language* 72: 16–36.
- Sommer-Lolei, Sabine & Dressler, Wolfgang U. (akzeptiert) First-language acquisition of agent and instrument nouns in German as (non-)prototypical derivational categories. In *Derivation and Inflection*, Martina Werner & Wolfgang U. Dressler [Hrsg.]. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Sommer-Lolei, Sabine, Mattes, Veronika, Korecky-Kröll, Katharina & Dressler, Wolfgang U. (2021) Early phases of development of German derivational morphology. In *The Acquisition of Derivational Morphology. A cross-linguistic perspective*. [Language Acquisition & Language Disorders 66], Veronika Mattes, Sabine Sommer-Lolei, Katharina Korecky-Kröll & Wolfgang U. Dressler [Hrsg.], 109–140. Amsterdam, Philadelphia: John Benjamins. <https://doi.org/10.1075/lald.66.05som>
- Stampe, David (1969) The acquisition of phonetic representation. *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*: 443–453.
- Stern, Clara & Stern, William (1928) *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung* (Children’s language: a psychological and

- language theoretical investigation). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Szagun, Gisela (2007) Langsam gleich gestört? Variabilität und Normalität im frühen Spracherwerb. *Forum Logopädie* 3(21): 20–25.
- Szagun, Gisela, Steinbrink, Claudia, Franik, Melanie & Stumper, Barbara (2006) Development of vocabulary and grammar in young German-speaking children assessed with a German language development inventory. *First Language* 26(3): 259–280. <https://doi.org/10.1177/0142723706056475>
- Tomasello, Michael (2003) *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Tribushinina, Elena, Voeikova Maria, D. & Noccetti, Sabrina [Hrsg.] (2015) *Semantics and Morphology of Early Adjectives in First Language Acquisition*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Trubetzkoy, Nikolaj S. (1939) *Grundzüge der Phonologie*. [Travaux du Cercle Linguistique de Prague 7]. Prague.
- Unsworth, Sharon (2013) Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition* 16(1): 86–110. <https://doi.org/10.1017/S1366728912000284>
- Valera, Salvador (2015) Conversion. In *Word-Formation. An International Handbook of the Languages of Europe* (Vol. 1). [Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 40.1], Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer [Hrsg.], 322–339. Berlin, Boston MA: De Gruyter Mouton. <https://doi.org/10.1515/9783110246254-019>
- Van der Schuit, Margje, Segers, Eliane, Van Balkom, Hans & Verhoeven, Ludo (2011) How cognitive factors affect language development in children with intellectual disabilities. *Research in Developmental Disabilities* 32(5): 1884–1894. <https://doi.org/10.1016/j.ridd.2011.03.015>
- Wellmann, Hans (1975) *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Hauptteil 2: Das Substantiv* (German word formation: Types and tendencies in present-day language. Main part 2: The noun). Düsseldorf: Schwann.
- Werner, Martina (2017) Zur Entwicklung der synthetischen Komposition in der Geschichte des Deutschen. *Journal of Word Formation* 1(1): 73–92.
- Wijnen, Frank, Kempen, Masja & Gillis, Steven (2001) Root infinitives in Dutch early child language: An effect of input? *Journal of Child Language* 28(3): 629–660. <https://doi.org/10.1017/S0305000901004809>
- Zurek, Wojciech H. [Hrsg.] (1990) *Complexity, entropy, and the physics of information*. Reading, MA: Addison-Wesley.

Anhang 1 Transliterations- und Codierungsbeispiele

A1.1 Auszug aus einem Transkript (CHAT)

JAN (3;10.08a):

- *JAN: ### jetzt hab(e) ich das licht eing(e)schaltet damit der fahrer sieht welche bausteine das sind .
*MUT: # ah ja .
*MUT: ja das muss er schon sehen .
*JAN: ja .
*MUT: is(t) es sonst zu dunkel ?
*JAN: ja .
*MUT: mhm .
*JAN: weil im &au +//.
*JAN: weil sonst is(t) es zu dunkel in seinem auto .

A1.2 Beispiel morphologische Codierung (MOR)

JAN (3;10.08a):

- *JAN: ### jetzt hab(e) ich das licht eing(e)schaltet damit der fahrer sieht welche bausteine das sind .
%mor: ADV|jetzt V:aux|hab-1S PRO|ich DET:art:def|d-as N:07:n|licht V:01|ein#schalt-PP CONJ|damit DET:art:def|d-er N:05:m|fahrer V:10|seh-3S DET:wh|welch-e N:03:m|bau+stein-PL PRO:dem|d-as V:S|sein-3P .
*MUT: # ah ja .
%mor: INTERJ|ah CO:ass|ja .
*MUT: ja das muss er schon sehen .
%mor: CO:ass|ja PRO:dem|d-as V:mod|muess-3S PRO|er ADV|schon V:10|seh-INF .
*JAN: ja .
%mor: CO:ass|ja .
*MUT: is(t) es sonst zu dunkel ?
%mor: V:S|sein-3S PRO|es ADV|sonst ADV|zu ADV|dunkel ?
*JAN: ja .
%mor: CO:ass|ja .
*MUT: mhm .
%mor: CO:ass|mhm .
*JAN: weil im &au +//.
%mor: CONJ|weil PREP|in~DET:art:def|d-em +//.
*JAN: weil sonst is(t) es zu dunkel in seinem auto .
%mor: CONJ|weil ADV|sonst V:S|sein-3S PRO|es ADV|zu ADV|dunkel PREP|in DET:pro:poss|sein-em N:01:n|auto .

A1.3 Codierungsbeispiel Zitat (CIT)

Beispiel für ein Zitat (Liedtext): D12-JAS (3;1):

- *PA2: # Haensel und Gret(e)l verlieben sich im wald [=! singt]. [+ cit]
%cit: n:prop|Haensel CONJ|und n:prop|Gretel V:07|verlauf-PRET:3P PRO:refl|sie-PL-ACC PREP|in~DET:art:def|d-em N:08:m|wald.

A1.4 Codierungsbeispiel Imitation (IMI)

Beispiel für eine direkte, unmittelbar auf die Äußerung der Mutter folgende Imitation: D04-SOS (4;9):

*PAR: jetzt spielma [: spielen wir] ja was andares [: anderes].
 %mor: ADV|jetzt V:01|spiel-1P PRO|wir ADV|ja PRO:indef|w-as ADJ|ander-es .
 *CHI: jetzt spielma [: spielen wir] ja was anderes. [+ imi]
 %imi: ADV|jetzt V:01|spiel-1P PRO|wir ADV|ja PRO:indef|w-as ADJ|ander-es .

Anhang 2 Codierungsschemata

A2.1 Basiscodierungen der Spontansprachdaten

Nach dem Import der Daten mittels der CLANTOCSV Software war es notwendig wichtige Basiscodierungen in Microsoft[®] Excel zu ergänzen. In dem Codierungsschema in Tabelle 42 werden diese relevanten Basiscodierungen beschrieben und mit Beispielen aus den Korpora ergänzt.

Tabelle 42. Übersicht über die wesentlichen Basiscodierungen der Spontansprachdaten

Basiscodierungen			
Spaltenbezeichnung	Beschreibung	Beispiel / möglicher Eintrag	
		INPUT	Jan, Kat, Len
Child_name	Name oder Kürzel des Kindes	D02-SIJ	JAN / LEN
	Eindeutige Identifikation des Korpus	D18-ALM	KAT
Child_SES	Sozioökonomischer Status des Kindes	high / low	high
Child_sex	Geschlecht des Kindes	m = männlich / w = weiblich	
Child_age	Alter des Kindes zum Aufnahmezeitpunkt (JMMTT)	30812	21024
Age_chi	Alter des Kindes zum Aufnahmezeitpunkt (JMM) für die monatliche Auswertung von JAN, KAT, LEN	n. a.	210
Age_chi_yrs	Alter des Kindes zum Aufnahmezeitpunkt (J) für die jährliche Auswertung von JAN, KAT, LEN	n. a.	2
Recording	Nummer der Aufnahme (JAN, KAT, LEN)	1 / 2 / 3 / 4	1 ... 114
	Erhebungszeitpunkt (INPUT): 1 = erste bis 4 = vierte Erhebung, für die Auswertung des INPUT-Korpus nach Datenpunkt		
Person	aus CHILDES importierte Abkürzung des Sprechers	*CHI	*KAT
	*CHI = <i>child</i> , Zielkind, *PAR = <i>parent</i> (= HBP) *JAN, *KAT, *LEN = Jan, Kathi, Lena, *MUT = Mutter andere Sprecher als die hier gelisteten werden nicht ausgewertet	*PAR	*MUT
mor_imi_cit_ads	aus CHILDES importierte Codierungszeile	mor / imi / cit / ads	
	mor = morphologische Codierungszeile, imi = Imitationen cit = Zitate, ads = <i>adult-directed</i> , pds = <i>pet-directed</i> zum Ausschluss der Codierungen CIT, IMI, ADS, PDS	pds	n. a.
Utterance_cha	aus CHILDES importierte Äußerung des Sprechers für die Disambiguierung eines Wortes, z. B. das Schloss (Gebäude) vs. das Schloss (Instrument zum Verschließen) für tatsächliche Realisation vs. Standardisierung (Aufahmeræt vs. Aufnahmegerät)	und da is(t) ein rad abgebroch(e)n.	

B a s i s c o d i e r u n g e n		
Spaltenbezeichnung	Beschreibung	Beispiel / möglicher Eintrag
		INPUT Jan, Kat, Len
Utterance_mor_imi_cit	aus CHILDES importierte, codierte Äußerung des Sprechers	CONJ und ADV da V:aux sein-3S DET:art:indef ein N:08:n rad V:11 ab#brech-PP .
Lemma	aus CHILDES importiertes Lemma für genaue Lemmazählung	ab#brech
Wordform	aus CHILDES importierte Wortform für Worttokenzählung	ab#brech-3S
Word_class	aus CHILDES importierte Wortklasse zur Auswahl der Wortklasse	N = Substantiv V = Verb / ADJ = Adjektiv
Word_class_specified	aus CHILDES importierte Subkategorie der Wortklasse zum Ausschluss bestimmter Kategorien, wie Hilfsverben (aux), Modalverben (mod) und Eigennamen (prop)	aux / mod / prop

A2.2 Zusatzcodierungen der Spontansprachdaten

Neben den essenziellen Basiscodierungen wurde ein weiteres, umfassendes Codierungsschema für Substantive, Verben und Adjektive entwickelt, das eine genaue Zuordnung jedes Lemmas ermöglicht, wie in Tabelle 43 dargestellt.

Tabelle 43. Übersicht über die erforderlichen Zusatzcodierungen für die Analyse der Derivationen und Komposita, sowie für die Fehlerauswertung, mit Beispielen

Z u s a t z k o d i e r u n g e n f ü r D e r i v a t i o n e n u n d K o m p o s i t a			
Spalte	Beschreibung	Kategorie	Beispiel
Word_category	Zuordnung eines Wortes in eine Kategorie, zur einfachen Auswahl im Zuge der Auswertung Als pseudo-deriviert gelten Wörter, die keine Basis haben, aber aufgrund ihrer formalen Struktur abgeleitet sein könnten	0 = Simplex 1 = pseudo-deriviertes Wort 2 = deriviertes Wort (Derivat) 3 = Kompositum (KOMP) 4 = synthetisches Kompositum	Katze, bauen, alt Fahne, gewinnen, fertig Fahr-er, be-halten, witz-ig Haus+tür, hell+grün Tennis+spiel-er
V_type	Zuordnung eines Verblemmas zu einer Kategorie	SIMP = Simplex PTLV = Partikelverb (trennbar) PREFV = Prefixverb (untrennbar) ZIRKV = zirkumfigiertes Verb SUFFV = suffigiertes Verb KONV mor = Konversion morph. IMPLIZIT = implizite Derivation Stamm(vokal)änderung	laufen, sehen, springen an#sehen, auf#springen be-suchen, ver-laufen be-sänft-ig-en telefon-ier-en angeln, fischen legen, füttern, stellen
N_Comp	Unterteilung der Substantivkomposita in drei Kategorien	1 = KOMP Ableitung im Erstglied 2 = KOMP Ableitung im Letztglied 3 = Kompositum ohne Ableitung	Flug+zeug Dosen+öffn-er Apfel+kern

Zusatzkodierungen für Derivationen und Komposita			
Spalte	Beschreibung	Kategorie	Beispiel
N_Comp_type	Genauere Angabe der Komponenten eines Substantivkompositums	N = Substantiv V = Verb(stamm) A = Adjektiv P = Präposition K = Kurzwort E = Eigenname C = Cranberrymorphem NUM = Numerale NEO = Neologismus O = Onomatopöie	NN: Saft+glas VN: Kratz+baum AN: Bunt+stift PN: Unter+hose KN: Bio+müll EN: Barbie+bett CN: Brom+beere NUMN: Ein+horn NEON: Schwutz+zuckerl ON: Mieke+katze
A_Comp_type	Genauere Angabe der Komponenten eines Adjektivkompositums	A = Adjektiv N = Substantiv NUM = Numerale	AA: kurz+ärmelig, rosa+rot NA: baby+leicht NUMA: sechs+eckig
WBCat_txt	Wortbildungskategorie des Lemmas (Substantive)	Nag = Nomen agentis Nin = Nomen instrumenti Nac = Nomen actionis Nres = Nomen acti (Resultat) Nloc = Nomen loci Nqu = Nomen qualitatis kollektiv = Kollektiva NUM = Numerale other = andere	Apothek-er-in Schrauben+zieh-er Stoß, Ein#sturz Verletz-ung, Aufkleb-er Buch+handl-ung, Bäck-er-ei Stärk-e Mensch-heit, Ge-büsch Eins-er, Acht-er Geburt-s+tag
Derivation_type_txt	Kategorisierung der Substantive, Verben und Adjektive auf Basis des relevanten Wortbildungsmusters	ZIRK = Zirkumfigierung KOMP = Kompositum KOMP1 = KOMP mit Ableitung im Erstglied KOMP2 = KOMP inkl. DIM synth KOMP = synthetisches KOMP KONV INF = Infinitivkonversion KONV mor = Konversion morph. IMPLIZIT = implizite Derivation Stamm(vokal)änderung SUFF = Suffigierung FSUFF = Fremdsuffigierung PP = Perfektpartizip PPräs = Präsenspartizip PREF = Präfigierung Präfix PTL = Präfigierung Partikel SIMP = Simplex, keine Derivation DIM = Diminutivierung	Ge-heg-e, be-ruh-ig-en Alt+papier Führ-er+schein Ohr+ring-erl Laut+sprech-er das Essen Sitz, Lauf Zug, Flug, Verband wunder-bar, Dieb-in Musik-ant, Poliz-ist ge-bog-en, er-staun-t schmerz-end, brenn-end ent-decken, er-zählen weg#bringen, zu#decken Obst, Ameise Has-i, Hund-erl, Mam-ilein
Prefixtype	Art des Präfixes (Verben)	PREF = untrennbares Präfix PTL = Partikel, trennbares Präfix DPTL = Doppelpartikel (trennbar) ZIRK = Zirkumfix	be-kommen, er-freuen rein#kommen, raus#gehen herum#hüpfen ent-schuld-ig-en
Affix_rel_txt	Angabe des relevanten Affixes oder der relevanten Derivation	Die zuletzt angewendete Derivation wird als die relevante angegeben	sagen → ver-sagen eben → un-eben → Un- eben-heit
Deriv2_txt	Angabe aller weiteren Wortbildungsmuster	UML = Umlaut / Ablaut stem mod. = Stammänderungen	Gefäng-nis (nis = rel. Affix, ä UML = 2. Derivation)
Base_txt	Angabe der Ableitungsbasis für die relevante Derivation	N = Substantiv V = Verb ADJ = Adjektiv PTLV = Partikelverb (trennbares Präfix) PREFV = Prefixverb (untrennbares Präfix) SUFFV = suffigiertes Verb ZIRKV = zirkumfigiertes Verb NUM = Numerale Fremdwort (z. B. lat.) E = Eigenname	

Zusatzkodierungen für Derivationen und Komposita			
Spalte	Beschreibung	Kategorie	Beispiel
Complexity	Angabe des Komplexitätsgrades des Lexems	0 = keine Ableitung, Simplizia 1 = eine Derivation (UML, PTL, PREF, Affigierung, Konversion, Komposita) 2 = zwei Ableitungen (siehe 1 + mehrere Affigierungen, synthetische Komposita) 3 = drei Ableitungen mehrere Affigierungen, UML, Konversion, Komposition, etc. 4 = vier oder mehr Ableitungen	Herd, fertig, verlier-fäll-, Wieg-e, Schubs, Heizung, telefon-ier-, be-komm-, leg-, Hand+schuh Vor-hang, Ab-sperr-ung, auff#leg-, gefähr-lich Auto+transport-er Ge-fäng-nis, Brief+träg-er, Ent-schuld-ig-ung, be-sänft-ig-en An-häng-er+kuppl-ung Un-durch-schau-bar-keit
Neologism	Eintrag Lemma ein Neologismus	0 = kein Neologismus 1 = Neologismus	Weck-er *Milch-er (jmd. der gerne Milch trinkt)
Opacity	Grad der Transparenz/Opazität	0 = transparent 1 = opak	mixen → Mix-er stellen → be-stellen
Error_chi_code	Angabe eines Fehlercodes Diese Codes können entweder alleine oder in Kombination auftreten zur Fehleranalyse der kindlichen Äußerungen	e*m = nur Endung wird verwendet ins*m = <i>insertion</i> , Einfügung eines Morphems, Lauts intf*m = Interfixfehler (KOMP) Weglassung, Ersetzung, Einfügung mod*m = Stammmodifikationsfehler om*m = <i>omission</i> , Weglassung p*m = Präfix/Partikelfehler phon*m = Phonetischer Fehler (Aussprache) s*m = Suffixfehler sub*m = Substitution - Ersetzung eines Morphems, Lauts STEM*m = Nur der Stamm alleine wird verwendet u*m = Umlautfehler	*vös statt nervös *dekor-er-iert statt dekoriert *Blumegart(e)n vs. Blumengarten *Lüger statt Lügner *ande statt andere *lor(e)n statt verloren *dazudagebn statt dazugegeben *Zuck statt Zuckerl *geheiraten statt geheiratet *nehm statt nehmen *Zuge statt Züge
Error_chi_txt	Detaillierte Fehlerbeschreibung	Texteingabe des Fehlers	Suffix t statt en

A2.3 Codierungsschema Elterninterviews

Tabelle 44. Auszug aus dem Codierungsschema zur Auswertung der Elterninterviews mit dem Fokus auf relevante Variablen für den Spracherwerb

Beschreibung	Label	Bezeichnung	Werte
KG-INPUT/Tag	TAX_KG_tg	Zeit, die CHI durchschnittlich an einem Wochentag im Kindergarten verbringt zum Interviewzeitpunkt X (1–3)	h m
HBP-INPUT/Tag	TAX_QT_tg	<i>quality time</i> , wertvolle Zeit, die mit der Hauptbezugsperson (HBP) gemeinsam verbracht wird, in Stunden und Minuten an einem durchschnittlichen Wochentag zum Interviewzeitpunkt X (1–3)	h m
NHBP-INPUT/Tag	TAX_ZH_tg	Zeit, die CHI durchschnittlich an einem Wochentag zu Hause bzw. in privater Obhut (auch Babysitter) verbringt, jedoch NICHT mit der HBP! zum Interviewzeitpunkt X (1–3)	h m

Beschreibung	Label	Bezeichnung	Werte
KG-INPUT/Wo	TAX_KG_wo	Zeit, die CHI durchschnittlich wöchentlich im KG verbringt zum Interviewzeitpunkt $X(1-3)$	h m
HBP-INPUT/Wo	TAX_QT_wo	<i>quality time</i> , wertvolle Zeit, die mit der HBP gemeinsam verbracht wird, in Stunden und Minuten in einer durchschnittlichen Woche, zum Interviewzeitpunkt $X(1-3)$	h m
NHBP-INPUT/Wo	TAX_ZH_wo	Zeit, die CHI durchschnittlich wöchentlich zu Hause bzw. in privater Obhut (auch Babysitter) verbringt, jedoch NICHT mit der HBP! Zum Interviewzeitpunkt $X(1-3)$	h m
L1 HBP	L1_HBP	Erstsprache der Hauptbezugsperson zum Aufnahmezeitpunkt 1 erhoben	1 Österr. Deutsch 2 Türkisch
L2 HBP	L2_HBP	Zweitsprache der Hauptbezugsperson	0 keine Zweitsprache 1 Österr. Deutsch 2 Türkisch 3 Spanisch 4 Englisch 5 Italienisch 6 Französisch 7 Kroatisch 8 Slowakisch 10 Dialekt 11 Kurmanji 12 Zaza 13 Romanian
L1 CHI	INTX_L1_CHI	Erstsprache des Fokuskindes zum Interviewzeitpunkt $X(1-3)$	1 Österr. Deutsch 2 Türkisch
L2 CHI	INTX_L2_CHI	Zweitsprache des Fokuskindes zum Interviewzeitpunkt $X(1-3)$	1 Österr. Deutsch 2 Türkisch 3 Spanisch 4 Englisch 5 Italienisch 6 Französisch 7 Kroatisch 8 Slowakisch 10 Dialekt 11 Kurmanji 12 Zaza 13 Romanian
Sprachlicher INPUT	INX_Deu_wo	Inputzeit mit Deutsch (L1 oder L2) pro Woche (hochgerechnet) im KG und Zuhause (mit oder ohne HBP) zum Interviewzeitpunkt $X(1-3)$	h m
	INX_Trk_wo	Inputzeit mit Türkisch (L1) pro Woche (hochgerechnet), Zuhause (mit oder ohne HBP) und teilweise im KG zum Interviewzeitpunkt $X(1-3)$	h m
	INX_L1_wo	Inputzeit mit der Erstsprache (entweder Türkisch oder Deutsch) hochgerechnet zum Interviewzeitpunkt $X(1-3)$	h m
	INX_L2_wo	Inputzeit L2 pro Woche (hochgerechnet), zu Hause zum Interviewzeitpunkt $X(1-3)$	h m
	L1BPX_Anz_Ges (auch für L2 & L3 erfasst)	Anzahl der sprachlichen Bezugspersonen für L1 gesamt zum Aufnahmezeitpunkt $X(1, 2)$	Z
	L1BPX_tg (auch für L2 & L3 erfasst)	Anzahl der Bezugspersonen, die täglich mit dem Fokuskind in L1 sprechen, zum Aufnahmezeitpunkt $X(1, 2)$	Z

Der frühe Erwerb der deutschen Derivationsmorphologie

Beschreibung	Label	Bezeichnung	Werte
	L1BPX_wo (auch für L2 & L3 erfasst)	Anzahl der Bezugspersonen, die mehrmals pro Woche mit dem Fokuskind in L1 sprechen, zum Aufnahmezeitpunkt $X(1, 2)$	Z
	L1BPX_ein_wo (auch für L2 & L3 erfasst)	Anzahl der Bezugspersonen, die einmal pro Woche mit dem Fokuskind in L1 sprechen, zum Aufnahmezeitpunkt $X(1, 2)$	Z
	L1BPX_mon (auch für L2 & L3 erfasst)	Anzahl der Bezugspersonen, die einmal pro Monat mit dem Fokuskind in L1 sprechen, zum Aufnahmezeitpunkt $X(1, 2)$	Z
	L1BPX_selten (auch für L2 & L3 erfasst)	Anzahl der Bezugspersonen, die weniger als einmal pro Monat mit dem Fokuskind in L1 sprechen, zum Aufnahmezeitpunkt $X(1, 2)$	Z
Sprachverhalten produktiv	L1_Onset_J	Sprachbeginn der Erstsprache in Jahren	0 vor dem ersten Jahr 1 mit 1;0 (1 Jahr, 0 Monate) 2 mit 1;6 (1 Jahr, 6 Monate) 3 mit 2;0 (2 Jahre, 0 Monate) 4 mit 2;6 (2 Jahre, 6 Monate) 5 nach 2;7
	L2_Deu_Onset_J	Sprachbeginn der Zweitsprache (Deutsch) in Jahren (nur für L1 Türkisch Kinder)	0 mit einem Jahr 1 mit 1;6 (1 Jahr, 6 Monate) 2 mit 2;0 (2 Jahre, 0 Monate) 3 mit 2;6 (2 Jahre, 6 Monate) 4 mit 3;0 (3 Jahre) 5 nach 3;1 6 bis jetzt noch nicht
	2Wort_Onset_J	Erste Wortverbindungen mit Jahren	0 mit einem Jahr 1 mit 1;6 (1 Jahr, 6 Monate) 2 mit 2;0 (2 Jahre, 0 Monate) 3 mit 2;6 (2 Jahre, 6 Monate) 4 mit 3;0 (3 Jahre) 5 nach 3;1 6 bis jetzt noch nicht
Sprachproduktion	L1_prodX	Erstsprache, die das Fokuskind spricht zum Aufnahmezeitpunkt X , bei 1 erhoben und bei 2 kontrolliert	1 Österr. Deutsch 2 Türkisch
Sprachperzeption	L1_rezX	Erstsprache, die das Fokuskind versteht zum Aufnahmezeitpunkt X , bei 1 erhoben und bei 2 kontrolliert	1 Österr. Deutsch 2 Türkisch
produktiv	L2_prodX	Zweitsprache, die das Fokuskind spricht zum Aufnahmezeitpunkt X , bei 1 erhoben und bei 2 kontrolliert	0 wird nicht gesprochen/keine Zweitsprache
rezeptiv	L2_rezX	Zweitsprache, die das Fokuskind versteht, zum Aufnahmezeitpunkt X , bei 1 erhoben und bei 2 kontrolliert	1 Österr. Deutsch 2 Türkisch 3 Spanisch 4 Englisch 5 Italienisch 6 Französisch 7 Kroatisch 8 Slowakisch 10 Dialekt
Sprachniveau L1 produktiv	L1_prod2_SN	Sprachniveau des Fokuskindes der Sprachproduktion der Erstsprache zum Aufnahmezeitpunkt 2	0 Muttersprache 1 führt ausführliche Gespräche 2 führt einfache Gespräche 3 spricht einige Wörter
L1 rezeptiv	L1_rez2_SN	Sprachniveau des Fokuskindes der Sprachrezeption der Erstsprache zum Aufnahmezeitpunkt 2	0 Muttersprache 1 versteht ausführliche Gespräche 2 versteht einfache Gespräche 3 versteht einige Wörter
L2 produktiv	L2_prod2_SN	Sprachniveau des Fokuskindes der Sprachproduktion der Zweitsprache	0 Muttersprache 1 führt ausführliche Gespräche

Beschreibung	Label	Bezeichnung	Werte
		zum Aufnahmezeitpunkt 2	2 führt einfache Gespräche 3 spricht einige Wörter
L2 rezeptiv	L2_rez2_SN	Sprachniveau des Fokuskindes der Sprachrezeption der Zweitsprache zum Aufnahmezeitpunkt 2	0 Muttersprache 1 versteht ausführliche Gespräche 2 versteht einfache Gespräche 3 versteht einige Wörter
Interaktion Konversation	AKTX_Konv	Interaktion zwischen HBP und CHI, Konversation , spontane Unterhaltung beim Essen, Baden etc. über Erlebnisse oder Erinnerungen zum Aufnahmezeitpunkt <i>X</i> (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
	AKTX_Konv_wo	Zeitangabe der Konversation , wieviel spricht HBP mit CHI und GK (Geschwister) unter der Woche zum Aufnahmezeitpunkt <i>X</i> (2, 3)	h m
	AKTX_Konv_wo_allein	Zeitangabe der Konversation , wieviel spricht die HBP alleine mit dem Fokuskind unter der Woche zum Aufnahmezeitpunkt <i>X</i> (2, 3)	h m
	AKTX_Konv_we	Zeitangabe der Konversation , wieviel spricht die HBP mit CHI und GK (Geschwister) am Wochenende	h m
Spielen	AKTX_Spi	Interaktion zwischen HBP und CHI, Spielen, Basteln, Zeichnen, Puzzeln mit Puppe/Zug spielen, Brettspiele Häufigkeit dieser Interaktion zum Aufnahmezeitpunkt <i>X</i> (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Vorlesen	AKTX_Vor	Interaktion zwischen HBP und CHI, Vorlesen , gemeinsames Bilderbuch anschauen oder Büchervorlesen Häufigkeit dieser Interaktion zum Aufnahmezeitpunkt <i>X</i> (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Erzählen	AKTX_Erz	Interaktion zwischen HBP und CHI, Erzählen Geschichten erfinden oder nacherzählen Häufigkeit dieser Interaktion zum Aufnahmezeitpunkt <i>X</i> (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Singen, Reimen	AKTX_SR	Interaktion zwischen HBP und CHI, Singen und Reimen Lieder singen, Reime, Aufzählungen, Satzspiele zum Aufnahmezeitpunkt <i>X</i> (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Geschwister	GK_Anz_ges	Anzahl der Geschwister gesamt	Z
	GK_m	Anzahl der männlichen GK	Z
	GK_w	Anzahl der weiblichen GK	Z
	GK_aelter	Anzahl der älteren GK	Z
	GK_juenger	Anzahl der jüngeren GK	Z
	CHI_R	Geburtsreihung des Fokuskindes im Vergleich zu den GK	1 Erstgeboren 2 Zweitgeboren 3 Drittgeboren 4 Viertgeboren
MEDIEN	BuchX_L1	Bücher werden in der Erstsprache vorgelesen, Aufnahmezeitpunkt <i>X</i> (1, 2)	0 trifft nicht zu 1 trifft zu
Bücher (Printmedien)	BuchX_L2	Bücher werden in der Zweitsprache vorgelesen zum Aufnahmezeitpunkt <i>X</i> (1, 2)	0 trifft nicht zu 1 trifft zu

Der frühe Erwerb der deutschen Derivationsmorphologie

Beschreibung	Label	Bezeichnung	Werte
	BLR2_L1	Bücher, Lieder, Gedichte und Reime werden in der Erstsprache präsentiert zum Aufnahmezeitpunkt 2	0 trifft nicht zu 1 trifft zu
	BLR2_L2	Bücher, Lieder, Gedichte und Reime werden in der Zweitsprache präsentiert zum Aufnahmezeitpunkt 2	0 trifft nicht zu 1 trifft zu
Bücher gesamt	MEDX_B_wo	Zeitangabe, wieviel Zeit CHI mit Büchern und vorgelesen bekommen pro Woche zubringt zum Aufnahmezeitpunkt X (2, 3)	h m
YouTube (Neue Medien)	MED1_You	YouTube darf angesehen werden zum Aufnahmezeitpunkt 1	0 trifft nicht zu 1 trifft zu
	MEDX_YIS	Das Fokuskind kommt mit YouTube, Internet oder Smartphone in Berührung zum Aufnahmezeitpunkt X (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Fernsehen (Massenmedien)	TV1_L1	TV-Serien/Filme werden in der Erstsprache angeschaut zum Aufnahmezeitpunkt 1	0 trifft nicht zu 1 trifft zu
	TV1_L2	TV-Serien/Filme werden in der Zweitsprache angeschaut zum Aufnahmezeitpunkt 1	0 trifft nicht zu 1 trifft zu
	MEDX_F	Das Fokuskind kommt mit Fernsehen (Filme, TV, DVD) in Berührung zum Aufnahmezeitpunkt X (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
audiovisuell gesamt	MEDX_YISF_wo	Zeitangabe, wieviel Zeit CHI mit YouTube, Internet, Smartphone oder Fernsehen pro Woche zubringt zum Aufnahmezeitpunkt X (2, 3)	h m
auditiv (Digitale Medien)	MEDX_HB	Das Fokuskind hört Hörbücher zum Aufnahmezeitpunkt X (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
Radio (Massenmedien)	MEDX_R	Das Fokuskind hört Radio (Radio, CD, USB) zum Aufnahmezeitpunkt X (2, 3)	1 täglich 2 mehrmals pro Woche 3 einmal pro Woche 4 mehrmals pro Monat 5 nie
auditiv gesamt	MEDX_HBR_wo	Zeitangabe, wieviel Zeit CHI mit Hörbüchern und Radiohören pro Woche zubringt, Aufnahmezeitpunkt X (2, 3)	h m
Erweiterung des Lexikons	IN3_Neu_HBP	CHI lernt neue Wörter von der Hauptbezugsperson	0 trifft nicht zu 1 trifft zu
	IN3_Neu_Part	CHI lernt neue Wörter vom Partner	0 trifft nicht zu 1 trifft zu
	IN3_Neu_GK	CHI lernt neue Wörter durch Geschwisterkind(ern)	0 trifft nicht zu 1 trifft zu
	IN3_Neu_FR	CHI lernt neue Wörter durch Freunde und Freundinnen	0 trifft nicht zu 1 trifft zu
	IN3_Neu_OO	CHI lernt neue Wörter von Oma und Opa (Großeltern)	0 trifft nicht zu 1 trifft zu
	IN3_Neu_KG	CHI lernt neue Wörter durch den Kindergarten (Pädagogin oder Freunde)	0 trifft nicht zu 1 trifft zu
	IN3_Neu_MED	CHI lernt neue Wörter durch Medien (Fernsehen, YouTube, Radio, Internet)	0 trifft nicht zu 1 trifft zu

Anhang 3 Korpora – Gesamtübersichten

A3.1 JAN-Korpus

Tabelle 45. JAN-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS)

Alter	Anz. Aufn.	Dauer (Min)	Anz. Lemmas/Tokens		Anz. Äußerungen		MLU		SD	
			CS	CDS	CS	CDS	CS	CDS	CS	CDS
1;3	1	45	39/189	381/2.455	183	632	1,027	4,158	0,194	4,158
1;4	1	30	42/141	261/1.565	139	412	1,165	5,706	0,408	4,300
1;5	1	30	41/127	246/1.764	119	406	1,134	6,645	0,408	4,465
1;6	1	30	42/179	260/1.733	163	434	1,380	6,053	0,648	4,319
1;7	1	30	56/165	249/1.503	157	377	1,318	5,944	0,705	3,945
1;8	8	240	246/1.356	952/13.059	1.172	2.994	1,753	6,818	0,830	4,602
1;9	8	240	306/1.451	1.007/13.901	1.271	3.323	1,626	6,388	0,835	4,322
1;10	6	180	304/1.328	875/9.331	998	2.228	1,979	6,360	1,046	4,438
1;11	8	240	378/2.314	998/11.759	1.306	2.744	2,713	6,475	1,595	4,486
2;0	8	240	548/3.440	956/11.596	1.763	2.813	3,022	6,453	1,936	4,642
2;1	4	120	346/1.584	645/5.203	750	1.288	3,275	6,121	2,058	4,505
2;2	4	120	383/2.121	676/6.025	804	1.397	4,002	6,530	2,732	5,047
2;3	4	120	295/1.346	769/6.605	540	1.383	3,898	7,430	2,837	4,973
2;4	4	120	349/2.077	784/6.373	622	1.319	5,159	7,386	3,408	5,081
2;5	4	120	371/1.987	883/7.578	646	1.379	4,571	8,445	3,407	5,721
2;6	4	120	328/2.088	740/6.315	672	1.328	4,626	7,313	3,179	5,440
2;7	4	120	278/2.013	792/7.526	769	1.507	3,683	7,447	3,088	4,833
2;8	4	120	381/2.406	831/6.718	770	1.302	4,788	7,470	3,521	5,231
2;9	4	120	306/2.130	798/6.710	746	1.353	4,161	7,293	3,058	5,183
2;10	4	120	259/1.540	791/7.899	527	1.588	4,378	7,475	3,445	4,563
2;11	4	120	282/1.792	702/6.342	597	1.394	4,444	6,778	3,130	4,501
3;0	4	120	331/2.631	772/8.372	888	1.871	4,306	6,761	3,568	4,665
3;1	1	30	115/403	403/1.903	144	388	4,076	7,322	3,298	4,980
3;2	1	30	81/292	280/1.055	115	217	3,539	6,880	3,025	4,359
3;3	1	30	113/394	386/1.687	136	354	4,404	7,008	3,359	4,117
3;4	1	30	147/725	321/2.286	201	426	4,970	7,812	3,432	4,132
3;5	1	30	154/601	476/2.373	160	405	5,800	9,133	4,430	5,155
3;6	1	30	156/607	385/1.676	179	322	5,307	7,739	3,420	4,318
3;7	1	30	165/642	336/1.461	202	340	4,426	6,094	3,809	4,131
3;8	1	30	226/864	323/1.415	231	277	5,000	7,199	3,648	5,608
3;9	1	30	116/370	262/925	138	222	3,551	5,869	3,573	4,382
3;10	1	30	159/535	310/1.679	164	409	4,750	6,054	3,972	4,544
3;11	1	30	187/670	321/1.780	172	349	5,424	7,367	4,171	5,350
4;0	1	30	177/599	358/1.525	133	309	6,195	7,298	4,031	5,889
4;1	1	30	163/546	329/1.573	144	367	5,792	6,553	4,477	5,014
4;2	1	30	214/798	415/1.988	164	350	7,207	8,460	5,459	7,192
4;3	1	30	223/851	324/1.352	191	271	6,356	7,683	4,461	6,614
4;4	1	30	219/811	323/1.618	174	270	7,092	8,641	4,766	5,921
4;5	1	30	146/385	429/2.398	127	437	4,205	8,197	3,555	6,999
4;6	1	30	173/745	306/1.501	221	344	4,566	6,369	3,520	4,844
4;7	1	30	179/669	265/1.058	163	254	5,785	6,134	3,772	4,870
4;8	1	30	170/442	489/2.329	88	359	6,955	9,760	4,421	7,521
4;9	1	30	162/553	233/1.092	127	258	6,087	5,899	4,012	4,672
4;10	1	30	153/485	297/1.139	114	185	5,974	8,086	4,697	6,405
4;11	1	30	116/309	268/1.033	76	264	5,882	5,405	4,743	5,117
Gesamt	114	3.435	2.337/47.701	4.956/187.178	19.166	40.849	4,261	6,987	3,069	5,012

A3.2 KAT-Korpus

Tabelle 46. KAT-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS)

Alter	Anz. Aufn.	Dauer (Min)	Anz. Lemmas/Tokens		Anz. Äußerungen		MLU		SD	
			CS	CDS	CS	CDS	CS	CDS	CS	CDS
1;6	2	36	7/15	52/153	14	35	1,000	6,857	0,000	4,310
1;7			k e i n e A u f n a h m e							
1;8	1	33	11/31	42/108	32	28	1,031	6,071	0,174	2,576
1;9	1	2	11/18	36/121	18	50	1,000	3,400	0,000	2,638
1;10	1	28	13/33	91/234	32	64	1,031	5,297	0,174	3,715
1;11	3	55,5	30/141	195/837	146	263	1,000	4,700	0,000	3,663
2;0	4	77	66/378	353/2.447	359	711	1,184	5,045	0,460	3,435
2;1	3	33	59/258	222/1.100	208	340	1,365	4,606	0,673	3,147
2;2	2	32	38/170	178/707	123	172	1,602	6,012	0,862	3,606
2;3	3	92	146/581	522/4.128	460	1.134	1,674	5,715	1,037	3,636
2;4	2	60	137/502	519/3.738	335	988	2,036	5,886	1,439	3,720
2;5	2	60	129/436	368/1.919	236	545	2,441	5,734	1,968	3,754
2;6	3	90	180/772	534/3.251	353	841	3,241	6,021	2,554	4,286
2;7			k e i n e A u f n a h m e							
2;8	2	60	235/991	493/2.616	353	754	3,776	5,302	3,091	4,062
2;9	2	60	166/525	404/1.983	211	523	3,621	5,727	2,645	4,216
2;10	1	38	203/818	307/1.036	282	231	4,025	6,528	3,292	4,464
2;11	1	29	121/349	232/782	126	136	3,611	7,625	2,711	4,905
3;0	1	30	145/348	172/500	129	102	4,078	7,294	2,844	5,268
Gesamt	34	815,5	750/6.366	1.633/25.660	3.417	6.917	2,219	5,754	1,407	3,847

A3.3 LEN-Korpus

Tabelle 47. LEN-Korpus - Monatliche Übersicht über Anzahl der Aufnahmen, Aufnahmedauer, Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) im Input (CDS) und Output (CS)

Alter	Anz. Aufn.	Dauer (Min)	Anz. Lemmas/Tokens		Anz. Äußerungen		MLU		SD	
			CS	CDS	CS	CDS	CS	CDS	CS	CS
1;7	2	60	54/232	384/1.848	197	622	1,249	4,728	0,624	3,962
1;8	3	90	69/324	479/3.064	282	977	1,379	4,756	0,856	3,810
1;9	3	90	83/454	544/3.519	413	1148	1,351	4,810	0,686	3,913
1;10	3	90	117/726	509/3.130	527	1011	1,579	4,852	0,980	3,982
1;11	3	90	114/582	312/1.615	438	601	1,642	4,123	1,107	3,584
2;0	2	60	102/467	298/1.476	344	509	1,759	4,281	1,041	3,411
2;1	2	60	113/472	386/1.984	344	632	1,791	4,587	1,263	3,779
2;2	2	60	102/421	291/1.200	265	398	2,132	4,696	1,446	3,326
2;3	2	60	152/616	381/1.964	330	559	2,736	5,501	1,971	4,125
2;4	2	60	146/567	376/2.039	317	653	2,549	5,009	1,719	3,718
2;5	2	60	154/703	328/1.516	368	465	2,679	5,105	1,907	3,631
2;6	2	60	172/682	373/1.978	312	604	3,266	5,425	2,179	3,878
2;7	2	60	198/853	420/2.236	271	610	4,705	5,739	2,749	4,436
2;8	2	60	207/1.179	430/2.533	388	736	4,647	5,425	2,975	4,229
2;9	2	60	183/1.091	420/2.546	356	731	4,750	5,539	3,147	3,985
2;10	2	60	233/1.063	351/1.670	330	547	4,509	4,837	3,081	3,872
2;11	2	60	185/776	329/1.445	286	455	3,773	4,776	2,660	3,673
3;0	2	60	205/879	422/2.297	296	619	4,358	5,683	3,426	4,372
Gesamt	40	1200	853/12.087	1.939/38.060	6.064	11.877	2,825	4,993	1,879	3,871

A3.4 INPUT-Korpus

Tabelle 48. INPUT-Korpus HSES - Übersicht über Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) je Kind und Erhebungszeitpunkt (DP 1–4) im Input (CDS) und Output (CS)

HSES	Anz. Lemmas/Tokens ^a		Anz. Äußerungen		MLU		SD		
	DP	CS	CS	CDS	CS	CDS	CS	CDS	
D02-SIJ	1	167/631	324/1.591	243	465	3,317	5,159	2,568	3,881
	2	149/714	365/1.979	210	536	4,405	5,817	2,922	4,172
	3	212/830	318/1.360	302	454	3,884	4,469	2,906	3,366
	4	302/1.245	452/1.789	417	539	4,403	5,082	3,689	3,861
	<i>Gesamt</i>	<i>572/3420</i>	<i>900/6.719</i>	<i>1.172</i>	<i>1.994</i>	<i>4,002</i>	<i>5,132</i>	<i>3,021</i>	<i>3,820</i>
D03-NIP	1	102/336	208/1.238	131	369	3,489	5,149	2,854	3,484
	2	114/415	244/1.305	167	447	3,269	4,640	2,876	3,292
	3	183/637	286/1.844	251	551	3,697	5,160	3,037	3,463
	4	245/1.042	237/1.223	466	528	3,073	3,515	2,538	2,847
	<i>Gesamt</i>	<i>431/2.430</i>	<i>570/5.610</i>	<i>1.015</i>	<i>1.895</i>	<i>3,382</i>	<i>4,616</i>	<i>2,826</i>	<i>3,272</i>
D04-SOS	1	196/827	337/1.376	171	369	6,509	5,610	4,134	3,833
	2	158/713	439/2.276	254	635	4,240	5,521	2,957	4,021
	3	241/1.068	422/1.952	349	556	4,636	5,462	3,510	4,016
	4	345/1.717	349/1.562	421	445	6,124	5,440	4,127	4,501
	<i>Gesamt</i>	<i>602/4.325</i>	<i>966/7.166</i>	<i>1.195</i>	<i>2.005</i>	<i>5,377</i>	<i>5,508</i>	<i>3,682</i>	<i>4,093</i>
D07-NIG	1	237/939	402/2.130	403	605	3,469	5,686	2,515	3,637
	2	188/747	422/2.509	282	621	4,177	6,290	3,021	4,345
	3	221/1.138	350/1.678	350	441	4,546	5,909	3,981	4,251
	4	214/670	440/1.990	279	534	3,430	5,757	3,104	4,581
	<i>Gesamt</i>	<i>585/3.494</i>	<i>1.018/8.307</i>	<i>1.314</i>	<i>2.201</i>	<i>3,906</i>	<i>5,911</i>	<i>3,155</i>	<i>4,204</i>
D09-KAB	1	142/467	305/1.205	223	449	2,812	4,225	2,899	3,372
	2	201/637	325/1.822	297	609	3,114	4,895	2,634	3,538
	3	196/602	302/1.444	215	495	4,181	4,588	3,501	3,532
	4	224/918	297/1.314	328	448	4,198	4,652	3,086	3,697
	<i>Gesamt</i>	<i>499/2.624</i>	<i>745/5.785</i>	<i>1.063</i>	<i>2.001</i>	<i>3,576</i>	<i>4,590</i>	<i>3,030</i>	<i>3,535</i>
D11-DAG	1	154/516	289/1.746	183	424	4,235	6,474	3,045	3,869
	2	182/653	306/1.573	167	381	4,563	6,743	3,804	4,709
	3	204/889	340/1.703	301	498	4,432	5,414	3,424	4,270
	4	230/1.041	295/1.753	302	430	5,225	6,251	3,663	3,778
	<i>Gesamt</i>	<i>478/3.099</i>	<i>722/6.775</i>	<i>953</i>	<i>1.733</i>	<i>4,614</i>	<i>6,221</i>	<i>3,484</i>	<i>4,157</i>
D13-JOP	1	226/1.205	302/1.341	399	407	4,163	4,929	3,087	3,534
	2	241/859	309/1.347	375	379	3,304	5,715	2,652	4,007
	3	173/711	264/1.017	270	304	3,752	5,332	3,047	4,216
	4	256/1.315	329/1.436	423	392	4,229	5,890	3,354	3,819
	<i>Gesamt</i>	<i>579/4.090</i>	<i>738/5.141</i>	<i>1.467</i>	<i>1.482</i>	<i>3,862</i>	<i>5,467</i>	<i>3,035</i>	<i>3,894</i>
D15-LUN	1	186/815	275/1.548	369	499	3,117	4,926	2,320	3,375
	2	215/934	265/1.878	398	560	3,427	5,336	2,458	3,572
	3	190/877	290/1.740	266	483	4,767	5,358	3,524	3,642
	4	193/621	376/2.150	238	669	3,576	5,143	2,968	3,755
	<i>Gesamt</i>	<i>529/3.247</i>	<i>707/7.316</i>	<i>1.271</i>	<i>2.211</i>	<i>3,722</i>	<i>5,191</i>	<i>2,818</i>	<i>3,586</i>
D16-LEP	1	138/423	327/2.251	169	805	3,657	4,330	2,718	3,591
	2	152/697	326/2.363	271	728	3,768	4,791	2,711	3,970
	3	156/798	302/1.926	347	656	3,432	4,418	2,865	3,707
	4	208/1.167	322/2.025	333	622	4,994	4,857	3,231	3,793
	<i>Gesamt</i>	<i>380/3.085</i>	<i>711/8.565</i>	<i>1.120</i>	<i>2.811</i>	<i>3,963</i>	<i>4,599</i>	<i>2,881</i>	<i>3,765</i>

Der frühe Erwerb der deutschen Derivationsmorphologie

HSES	Anz. Lemmas/Tokens ^a		Anz. Äußerungen		MLU		SD		
	DP	CS	CDS	CS	CDS	CS	CDS	CS	CDS
D20-LUD	1	210/1.029	371/2.406	375	697	3,715	5,271	2,748	3,579
	2	177/813	305/1.525	323	482	3,746	4,811	2,957	3,527
	3	208/1.314	280/2.135	474	627	4,139	5,139	3,104	3,693
	4	286/1.393	366/2.222	465	674	4,265	4,942	3,330	3,616
	<i>Gesamt</i>	<i>552/4.549</i>	<i>759/8.288</i>	<i>1.637</i>	<i>2.480</i>	<i>3,966</i>	<i>5,041</i>	<i>3,035</i>	<i>3,604</i>
D23-SOK	1	215/912	259/1.209	340	429	3,788	4,347	2,757	3,086
	2	222/1.157	264/924	385	314	4,527	4,586	3,438	3,115
	3	208/1.007	263/1.087	393	394	3,819	4,127	3,007	3,181
	4	253/821	334/1.149	341	364	3,622	5,102	3,083	3,597
	<i>Gesamt</i>	<i>540/3.897</i>	<i>685/4.369</i>	<i>1.459</i>	<i>1.501</i>	<i>3,939</i>	<i>4,541</i>	<i>3,071</i>	<i>3,245</i>
D27-JOZ	1	191/503	326/1.734	201	478	3,891	5,757	3,079	4,115
	2	224/760	334/1.701	284	506	4,215	5,154	3,265	3,582
	3	262/1.189	351/1.544	424	487	4,139	4,959	3,075	3,957
	4	235/1.151	294/1.420	398	465	4,638	4,961	3,336	3,808
	<i>Gesamt</i>	<i>618/3.603</i>	<i>785/6.399</i>	<i>1.307</i>	<i>1.936</i>	<i>4,221</i>	<i>5,208</i>	<i>3,189</i>	<i>3,866</i>
D28-MOH	1	248/1.164	278/1.399	421	384	4,157	6,107	3,597	3,785
	2	271/1.130	317/1.443	395	378	4,091	6,063	3,701	3,833
	3	329/1.592	356/1.871	513	433	4,669	6,822	3,807	4,295
	4	384/1.860	316/1.319	534	334	5,384	6,323	4,247	4,681
	<i>Gesamt</i>	<i>758/5.746</i>	<i>742/6.032</i>	<i>1.863</i>	<i>1.529</i>	<i>4,575</i>	<i>6,329</i>	<i>3,838</i>	<i>4,149</i>
D29-SAK	1	222/769	375/1.741	273	475	4,300	5,789	2,994	3,758
	2	225/797	436/2.267	256	592	4,508	6,079	3,470	3,986
	3	234/856	399/2.292	307	608	4,329	5,929	3,405	4,449
	4	227/748	453/2.305	236	546	4,619	6,337	3,248	4,371
	<i>Gesamt</i>	<i>602/3.170</i>	<i>1.031/8.605</i>	<i>1.072</i>	<i>2.221</i>	<i>4,439</i>	<i>6,034</i>	<i>3,279</i>	<i>4,141</i>
D31-AND	1	279/1.286	421/2.148	431	653	4,116	4,923	3,159	3,675
	2	226/835	375/1.877	258	575	4,876	5,136	3,452	3,918
	3	242/938	346/1.826	313	468	4,776	6,143	3,560	4,731
	4	199/772	344/1.573	195	413	5,959	5,833	4,426	4,244
	<i>Gesamt</i>	<i>590/3.831</i>	<i>880/7.424</i>	<i>1.197</i>	<i>2.109</i>	<i>4,932</i>	<i>5,509</i>	<i>3,649</i>	<i>4,142</i>
HSES Gesamt	3.239/54.610	4.208/102.501	19.105	30.109	4,165	5,326	3,200	3,831	

a. Lemma- und Tokenfrequenzen dieser Tabelle basieren auf dem Excel-Datenblatt und unterscheiden sich daher in der Gesamtsumme von Tabelle 3. Für die Kalkulation dieser Tabelle ergeben sich die Gesamtsummen auf Basis der errechneten Lemmafrequenzen mittels einer Pivot-Tabelle in Microsoft® Excel. Je nach Gruppierung (HSES, LSES, alle Datenpunkte zusammen) weichen die Summenzahlen insofern voneinander ab, als jedes Lemma pro Gruppierung nur einmal gezählt wird. Dies betrifft auch die Summenzahlen (Lemmas) in den Tabellen 49, 57 und 58, die sich nicht aus den einzelnen Datenpunkten addieren, sondern die Gesamtzahl an Lemmas im gesamten Zeitraum wiedergeben.

Tabelle 49. INPUT-Korpus LSES - Übersicht über Lemma- und Tokenfrequenzen, Anzahl der Äußerungen, mittlere Äußerungslängen (MLU) und Standardabweichung (SD) je Kind und Erhebungszeitpunkt (1–4) im Input (CDS) und Output (CS)

LSES	Anz. Lemmas/Tokens		Anz. Äußerungen		MLU		SD		
	DP	CS	CS	CDS	CS	CDS	CS	CDS	
D05-FLB	1	134/429	362/1.721	175	499	2,794	5,269	2,136	3,617
	2	205/674	395/1.731	234	502	4,244	5,187	2,807	3,787
	3	201/694	363/1.976	328	695	2,988	4,046	2,709	3,565
	4	245/887	461/2.290	329	699	3,778	4,900	3,105	3,804
	<i>Gesamt</i>	<i>513/2.684</i>	<i>919/7.718</i>	<i>1.066</i>	<i>2.395</i>	<i>3,451</i>	<i>4,851</i>	<i>2,689</i>	<i>3,693</i>
D06-LAB	1	162/529	363/2.117	253	656	2,759	5,383	2,004	3,896
	2	160/953	272/2.005	389	629	3,548	4,887	2,546	3,313
	3	199/1.011	276/1.445	296	420	5,074	5,407	3,458	3,405
	4	202/871	319/1.693	282	513	4,585	5,148	3,082	3,436
	<i>Gesamt</i>	<i>446/3.364</i>	<i>710/7.260</i>	<i>1.220</i>	<i>2.218</i>	<i>3,992</i>	<i>5,206</i>	<i>2,773</i>	<i>3,513</i>
D08-FAJ	1	215/930	201/666	389	190	2,992	5,711	2,258	3,576
	2	266/1.161	196/774	530	210	3,038	6,205	2,282	3,427
	3	209/1.027	136/286	320	88	4,306	5,545	3,092	3,477
	4	215/923	179/535	309	148	4,479	5,892	3,374	3,685
	<i>Gesamt</i>	<i>588/4.041</i>	<i>412/2.261</i>	<i>1.548</i>	<i>636</i>	<i>3,704</i>	<i>5,838</i>	<i>2,752</i>	<i>3,541</i>
D10-LER	1	170/697	286/1.596	338	510	3,047	4,871	2,076	3,838
	2	140/594	333/1.598	250	456	3,576	5,450	2,599	4,182
	3	150/562	279/1.561	261	564	3,103	3,929	2,450	3,293
	4	184/752	304/1.481	280	442	4,021	5,066	3,219	3,828
	<i>Gesamt</i>	<i>407/2.605</i>	<i>718/6.236</i>	<i>1.129</i>	<i>1.972</i>	<i>3,437</i>	<i>4,829</i>	<i>2,586</i>	<i>3,785</i>
D12-JAS	1	135/521	256/1.161	278	406	2,633	4,429	2,082	2,804
	2	111/357	180/664	203	250	2,473	4,132	2,255	2,763
	3	178/790	261/1.010	381	411	2,801	3,742	2,169	2,583
	4	170/823	346/2.003	293	521	3,993	5,906	3,034	3,806
	<i>Gesamt</i>	<i>373/2.491</i>	<i>649/4.838</i>	<i>1.155</i>	<i>1.588</i>	<i>2,975</i>	<i>4,552</i>	<i>2,385</i>	<i>2,989</i>
D14-KEW	1	147/561	171/779	269	212	2,539	5,882	1,666	3,104
	2	137/711	276/1.531	321	446	3,340	5,780	1,992	3,170
	3	91/495	209/1.176	262	436	2,550	4,264	2,007	3,154
	4	159/513	257/1.375	299	388	2,217	5,410	1,887	3,108
	<i>Gesamt</i>	<i>371/2.280</i>	<i>524/4.861</i>	<i>1.151</i>	<i>1.482</i>	<i>2,662</i>	<i>5,334</i>	<i>1,888</i>	<i>3,134</i>
D17-TOB	1	139/770	167/663	274	179	4,255	5,961	3,142	3,810
	2	232/902	196/731	353	201	3,728	5,697	2,814	3,822
	3	176/926	226/1.055	348	310	3,491	4,952	2,797	3,616
	4	225/1.185	190/662	487	258	3,497	4,039	2,936	3,060
	<i>Gesamt</i>	<i>475/3.783</i>	<i>454/3.111</i>	<i>1.462</i>	<i>948</i>	<i>3,743</i>	<i>5,162</i>	<i>2,922</i>	<i>3,577</i>
D18-ALM	1	247/1.373	319/1.609	510	512	3,906	4,795	3,337	3,821
	2	222/1.523	234/1.422	509	440	4,397	5,002	3,492	3,707
	3	253/1.357	266/1.692	454	500	4,324	5,188	3,571	3,577
	4	307/1.509	332/1.643	463	489	5,006	5,153	3,559	3,706
	<i>Gesamt</i>	<i>629/5.762</i>	<i>666/6.366</i>	<i>1.936</i>	<i>1.941</i>	<i>4,408</i>	<i>5,035</i>	<i>3,490</i>	<i>3,703</i>
D19-VAS	1	182/608	248/1.042	261	323	3,490	5,257	2,783	3,929
	2	142/491	261/1.340	178	411	3,938	5,156	2,817	3,379
	3	172/738	282/1.308	415	458	2,287	4,247	2,203	3,402
	4	187/876	204/901	395	344	2,982	3,922	2,834	2,898
	<i>Gesamt</i>	<i>444/2.713</i>	<i>586/4.591</i>	<i>1.249</i>	<i>1.536</i>	<i>3,174</i>	<i>4,646</i>	<i>2,659</i>	<i>3,402</i>
D21-PAS	1	124/681	234/1.141	280	323	3,307	5,505	2,195	3,439
	2	149/1.041	286/1.511	379	405	4,074	5,741	2,884	3,672
	3	174/964	235/1.203	453	380	2,945	4,837	2,396	3,532
	4	185/938	219/965	409	301	3,281	4,844	2,812	3,524
	<i>Gesamt</i>	<i>365/3.624</i>	<i>578/4.820</i>	<i>1.521</i>	<i>1.409</i>	<i>3,402</i>	<i>5,232</i>	<i>2,572</i>	<i>3,542</i>

LSES	Anz. Lemmas/Tokens		Anz. Äußerungen		MLU		SD		
	DP	CS	CDS	CS	CDS	CS	CDS	CS	CDS
D22-LAS	1	185/704	249/849	244	275	4,180	4,825	2,689	3,665
	2	173/676	242/1.166	256	406	3,695	4,360	2,726	3,143
	3	159/547	200/798	220	304	3,477	3,928	2,718	3,177
	4	203/815	244/1.527	302	587	3,967	3,801	3,399	2,920
	<i>Gesamt</i>	<i>455/2.742</i>	<i>543/4.340</i>	<i>1.022</i>	<i>1.572</i>	<i>3,830</i>	<i>4,229</i>	<i>2,883</i>	<i>3,226</i>
D24-RAG-	1	155/491	304/1.610	206	437	3,398	6,176	2,244	4,328
	2	113/369	302/1.538	137	429	3,453	5,795	2,476	4,145
	3	156/418	226/792	133	206	4,368	6,330	3,346	4,446
	4	147/473	278/1.325	152	316	4,711	6,864	3,494	4,219
	<i>Gesamt</i>	<i>368/1.751</i>	<i>666/5.265</i>	<i>628</i>	<i>1.388</i>	<i>3,983</i>	<i>6,291</i>	<i>2,890</i>	<i>4,285</i>
D25-PHG	1	122/364	254/1.278	172	354	3,029	6,124	2,117	3,848
	2	181/806	288/1.584	277	445	4,361	5,640	2,423	3,816
	3	188/542	237/732	222	218	3,505	5,587	2,637	4,490
	4	160/637	290/1.331	173	301	5,202	7,176	3,560	3,921
	<i>Gesamt</i>	<i>447/2.349</i>	<i>654/4.925</i>	<i>844</i>	<i>1.318</i>	<i>4,024</i>	<i>6,132</i>	<i>2,684</i>	<i>4,019</i>
D26-JUB	1	23/59	176/759	49	306	1,306	3,922	0,542	2,919
	2	116/495	240/1.005	281	373	2,231	4,231	1,553	3,017
	3	146/545	235/839	192	297	4,083	4,465	2,583	3,165
	4	131/440	232/775	134	248	5,172	4,774	2,650	3,246
	<i>Gesamt</i>	<i>284/1.539</i>	<i>549/3.378</i>	<i>656</i>	<i>1.224</i>	<i>3,198</i>	<i>4,348</i>	<i>1,832</i>	<i>3,087</i>
LSES Gesamt	2.442/41.728	2.854/69.970	16.587	21.627	3,570	5,120	2,643	3,535	

Anhang 4 Verteilung der Wortbildungsmuster

Im Folgenden werden die Lemma- und Tokenfrequenzen in CS und CDS in allen Korpora nach Wortklasse und Wortbildungsmuster aufgelistet.

A4.1 Nominale Wortbildungsmuster in CS und CDS

Tabelle 50. Frequenzen in Lemmas und Tokens und prozentuale Verteilung der Wortbildungsmuster der Substantive in CS und CDS in allen Korpora im gesamten Zeitraum

Korpus	JAN	KAT	LEN	INPUT	INPUT
	(1;3-4;11)	(1;6-3;0)	(1;7-3;0)	HSES	LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	469/4.908	183/697	189/1.330	582/3.926	415/3.221
% N SIMP/N Ges.	44,75/60,34	57,73/61,52	62,58/73,77	44,46/66,25	51,81/72,84
N KOMP	295/1.654	66/119	48/105	407/811	183/379
% N KOMP/N Ges.	28,15/20,33	20,82/10,50	15,89/5,82	31,09/13,69	22,85/8,57
synth. KOMP	80/305	11/19	1/1	72/139	36/53
% syn. KOMP/N Ges.	7,63/3,75	3,47/1,68	0,33/0,06	5,5/2,35	4,49/1,2
DIM	50/159	20/230	28/184	65/221	61/169
% DIM/N Ges.	4,77/1,95	6,31/20,30	9,27/10,21	4,97/3,73	7,62/3,82
ZIRK	1/2	0/0	0/0	2/21	0/0
% ZIRK/N Ges.	0,1/0,02	0,0/0,0	0,0/0,0	0,15/0,35	0,0/0,0
PREF	8/23	0/0	0/0	5/14	5/15
% PREF/N Ges.	0,76/0,28	0,0/0,0	0,0/0,0	0,38/0,24	0,62/0,34

Korpus	JAN	KAT	LEN	INPUT	INPUT
	(1;3-4;11)	(1;6-3;0)	(1;7-3;0)	HSES	LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SUFF	96/495	17/32	23/150	118/348	60/192
% SUFF/N Ges.	9,16/6,09	5,36/2,82	7,62/8,32	9,01/5,87	7,49/4,34
FSUFF	9/211	7/12	3/4	8/41	8/76
% FSUFF/N Ges.	0,86/2,59	2,21/1,06	0,99/0,22	0,61/0,69	1,00/1,72
KONV INF	2/10	0/0	2/4	7/31	2/2
% KONV INF/N Ges.	0,19/0,12	0,0/0,0	0,66/0,22	0,53/0,52	0,25/0,05
KONV mor	27/244	10/17	7/13	32/329	20/184
% KONV mor/N Ges.	2,58/3,00	3,15/1,50	2,32/0,72	2,44/5,55	2,50/4,16
IMPL (Stammmod.)	11/123	3/7	1/12	11/45	11/131
% IMPL/N Ges.	1,05/1,51	0,95/0,62	0,33/0,67	0,84/0,76	1,37/2,96
N Gesamt	1.048/8.134	317/1.133	302/1.803	1.309/5.926	801/4.422
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	828/11.705	361/1.866	356/2.569	792/6.661	534/4.371
% N SIMP/N Ges.	34,76/62,61	49,52/67,54	45,94/62,87	38,94/65,05	46,68/67,74
N KOMP	832/3.528	196/372	206/592	684/1.561	294/685
% N KOMP/N Ges.	34,93/18,87	26,89/13,46	26,58/14,49	33,63/15,24	25,70/10,62
synth. KOMP	241/677	30/56	38/58	150/270	76/126
% syn. KOMP/N Ges.	10,12/3,62	4,12/2,03	4,90/1,42	7,37/2,64	6,64/1,95
DIM	112/460	40/177	53/339	111/385	68/370
% DIM/N Ges.	4,70/2,46	5,49/6,41	6,84/8,30	5,46/3,76	5,94/5,73
ZIRK	5/20	1/2	0/0	4/44	1/1
% ZIRK/N Ges.	0,21/0,11	0,14/0,07	0,0/0,0	0,20/0,43	0,09/0,02
PREF	13/74	6/8	3/9	13/29	6/14
% PREF/N Ges.	0,55/0,40	0,82/0,29	0,39/0,22	0,64/0,28	0,52/0,22
SUFF	238/1154	48/155	79/398	190/636	111/436
% SUFF/N Ges.	9,99/6,17	6,58/5,61	10,19/9,74	9,34/6,21	9,70/6,76
FSUFF	24/411	10/38	12/42	20/88	12/64
% FSUFF/N Ges.	1,01/2,20	1,37/1,38	1,55/1,03	0,98/0,86	1,05/0,99
KONV INF	3/38	4/13	3/16	8/46	3/9
% KONV INF/N Ges.	0,13/0,20	0,55/0,47	0,39/0,39	0,39/0,45	0,26/0,14
KONV mor	54/404	23/49	16/46	40/422	26/272
% KONV mor/N Ges.	2,27/2,16	3,16/1,77	2,06/1,13	1,97/4,12	2,27/4,22
IMPL (Stammmod.)	32/225	10/27	9/17	22/98	13/105
% IMPL/N Ges.	1,34/1,20	1,37/0,98	1,16/0,42	1,08/0,96	1,14/1,63
N Gesamt	2.382/18.696	729/2.763	775/4.086	2.034/10.240	1.144/6.453

A4.2 Verbale Wortbildungsmuster in CS und CDS

Tabelle 51. Frequenzen in Lemmas und Tokens, und prozentuale Verteilung der Wortbildungsmuster der Verben in CS und CDS in allen Korpora im gesamten Zeitraum

Korpus	JAN	KAT	LEN	INPUT	INPUT
	(1;3-4;11)	(1;6-3;0)	(1;7-3;0)	HSES	LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	195/6.618	75/631	89/1.570	230/6.766	173/5.308
% V SIMP/V Ges.	33,45/79,42	43,10/74,41	40,27/78,38	29,26/79,57	29,73/79,82
PTL	325/1.341	77/161	115/375	449/1.303	335/1.044
% PTLV/V Ges.	55,75/16,09	44,25/18,99	52,04/18,72	57,12/15,32	57,56/15,7
PREF	25/177	5/16	5/16	57/246	36/160
% PREFV/V Ges.	4,29/2,12	2,87/1,89	2,26/0,80	7,25/2,89	6,19/2,41
ZIRK	1/1	0/0	0/0	2/3	1/1
% ZIRKV/V Ges.	0,17/0,01	0/0	0/0	0,25/0,04	0,17/0,02
SUFF	11/87	3/3	0/0	20/71	11/18
% SUFFV/V Ges.	1,89/1,04	1,72/0,35	0/0	2,54/0,83	1,89/0,27
KONV mor	15/55	8/25	10/40	17/76	17/50
% KONV/V Ges.	2,57/0,66	4,60/2,95	4,52/2,00	2,16/0,89	2,92/0,75
IMPL (Stammmod.)	11/54	6/12	2/2	11/38	9/69
% IMPLV/V Ges.	1,89/0,65	3,45/1,42	0,90/0,10	1,40/0,45	1,55/1,04
V Gesamt	583/8.333	174/848	221/2.003	786/8.503	582/6.650
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	325/26.275	134/3.977	178/5.201	291/14.866	226/10.434
% SIMP/V Ges.	22,84/80,06	28,63/75,77	28,39/72,48	23,70/77,49	23,62/74,67
PTL	897/5.059	258/940	362/1.592	760/3210	587/2793
% PTLV/V Ges.	63,04/15,41	55,13/17,91	57,74/22,19	61,89/16,73	61,34/19,99
PREF	98/732	26/114	37/149	97/545	75/309
% PREFV/V Ges.	6,89/2,23	5,56/2,17	5,90/2,08	7,90/2,84	7,84/2,21
ZIRK	3/5	1/1	4/7	2/6	3/3
% ZIRKV/V Ges.	0,21/0,02	0,21/0,02	0,64/0,10	0,16/0,03	0,31/0,02
SUFF	38/358	13/48	12/77	28/179	23/136
% SUFFV/V Ges.	2,67/1,09	2,78/0,91	1,91/1,07	2,28/0,93	2,40/0,97
KONV mor	36/164	19/78	20/95	30/233	28/170
% KONV/V Ges.	2,53/0,50	4,06/1,49	3,19/1,32	2,44/1,21	2,93/1,22
IMPL (Stammmod.)	26/228	17/91	14/55	20/146	15/129
% IMPLV/V Ges.	1,83/0,69	3,63/1,73	2,23/0,77	1,63/0,76	1,57/0,92
V Gesamt	1.423/32.821	468/5.249	627/7.176	1.228/19.185	957/13.974

A4.3 Adjektivische Wortbildungsmuster in CS und CDS

Tabelle 52. Frequenzen in Lemmas und Tokens, und prozentuale Verteilung der Wortbildungsmuster der Adjektive in CS und CDS in allen Korpora im gesamten Zeitraum

Korpus	JAN	KAT	LEN	INPUT	INPUT
	(1;3-4;11)	(1;6-3;0)	(1;7-3;0)	HSES	LSES
	CS	CS	CS	CS	CS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	97/1.160	27/99	43/182	112/1.307	92/902
% SIMP / ADJ Ges.	56,73/81,92	81,82/93,40	91,49/96,30	58,03/86,04	65,25/88,09
KOMP	14/16	1/1	0/0	13/22	5/8
% KOMP / ADJ Ges.	8,19/1,13	3,03/0,94	0,0/0,0	6,74/1,45	3,55/0,78
PP	11/18	0/0	0/0	6/7	4/6
% PP / ADJ Ges.	6,43/1,27	0,0/0,0	0,0/0,0	3,11/0,46	2,84/0,59
PPräs	1/1	0/0	0/0	3/4	0/0
% PPräs / ADJ Ges.	0,58/0,07	0,0/0,0	0,0/0,0	1,55/0,26	0,0/0,0
PREF	2/4	0/0	0/0	11/12	14/22
% PREF / ADJ Ges.	1,17/0,28	0,0/0,0	0,0/0,0	5,70/0,79	9,93/2,15
SUFF	42/197	5/6	4/7	44/159	19/78
% SUFF / ADJ Ges.	24,56/13,91	15,15/5,66	8,51/3,70	22,80/10,47	13,48/7,62
FSUFF	2/15	0/0	0/0	2/6	4/4
% FSUFF / ADJ Ges.	1,17/1,06	0,0/0,0	0,0/0,0	1,04/0,39	2,84/0,39
KONV	2/5	0/0	0/0	2/2	3/4
% KONV / ADJ Ges.	1,17/0,35	0,0/0,0	0,0/0,0	1,04/0,13	2,13/0,39
ADJ Gesamt	171/1.416	33/106	47/189	193/1.519	141/1.024
	CDS	CDS	CDS	CDS	CDS
	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
SIMPLEX	176/4.687	78/624	81/800	146/2.310	120/1.544
% SIMP / ADJ Ges.	42,21/80,70	18,71/10,74	19,42/13,77	35,01/39,77	28,78/26,58
KOMP	4/15	2/12	0/0	3/10	3/11
% KOMP / ADJ Ges.	6,00/0,71	1,20/0,10	0,24/0,02	5,76/0,62	2,64/0,34
PP	42/86	2/6	9/9	30/58	18/34
% PP / ADJ Ges.	10,07/1,48	0,48/0,10	2,16/0,15	7,19/1,00	4,32/0,59
PPräs	7/14	0/0	2/3	10/20	7/14
% PPräs / ADJ Ges.	1,68/0,24	0,0/0,0	0,48/0,05	2,40/0,34	1,68/0,24
PREF	16/35	1/1	7/15	17/20	10/17
% PREF / ADJ Ges.	3,84/0,60	0,24/0,02	1,68/0,26	4,08/0,34	2,4/0,29
SUFF	145/863	32/88	40/92	92/370	47/164
% SUFF / ADJ Ges.	34,77/14,86	7,67/1,52	9,59/1,58	22,06/6,37	11,27/2,82
FSUFF	2/67	1/1	2/4	3/6	4/6
% FSUFF / ADJ Ges.	0,48/1,15	0,24/0,02	0,48/0,07	0,72/0,10	0,96/0,10
KONV	4/15	2/12	0/0	3/10	3/11
% KONV / ADJ Ges.	0,96/0,26	0,48/0,21	0,0/0,0	0,72/0,17	0,72/0,19
ADJ Gesamt	417/5.808	121/738	142/924	325/2.830	220/1.810

Anhang 5 Derivationen in CS und CDS – Übersichten

A5.1 JAN-Korpus – N, V, ADJ DERIV je Monat

Tabelle 53. JAN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	JAN CS	JAN CDS	JAN CS	JAN CDS	JAN CS	JAN CDS
Korpus	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1;3	0/0	6/23	0/0	47/83	0/0	6/7
1;4	0/0	3/16	0/0	40/66	0/0	7/8
1;5	0/0	4/5	0/0	37/97	0/0	1/1
1;6	2/2	6/26	1/2	33/84	0/0	2/2
1;7	1/3	6/21	2/8	26/46	0/0	1/1
1;8	14/71	46/159	12/78	178/431	2/2	19/56
1;9	17/66	36/139	30/76	185/502	5/17	20/95
1;10	18/63	41/120	36/85	127/263	4/37	19/85
1;11	19/88	53/191	49/151	194/447	4/14	23/69
2;0	29/119	43/165	74/204	183/457	6/13	17/33
2;1	12/44	29/58	36/82	94/168	4/5	9/22
2;2	14/28	28/57	33/56	90/153	8/11	18/26
2;3	12/27	30/59	23/35	120/247	3/5	20/36
2;4	17/29	40/87	35/60	118/207	3/5	11/31
2;5	13/23	44/89	36/54	138/246	7/10	21/47
2;6	11/22	40/71	33/69	104/207	3/6	17/26
2;7	18/45	41/113	30/40	132/268	4/5	30/63
2;8	25/77	43/83	40/69	123/198	8/15	30/59
2;9	8/15	35/63	33/69	139/238	5/13	23/52
2;10	10/24	37/69	26/35	111/292	2/4	27/40
2;11	10/31	38/71	23/33	92/206	1/5	18/30
3;0	11/26	37/68	34/78	131/391	4/5	13/26
3;1	0/0	9/16	8/10	48/66	3/5	9/10
3;2	1/1	10/14	6/7	35/42	0/0	7/11
3;3	1/2	11/12	12/13	54/76	0/0	10/12
3;4	1/2	7/16	9/25	39/93	0/0	5/7
3;5	8/18	22/38	20/31	54/79	1/1	9/13
3;6	10/37	19/30	13/20	53/67	0/0	11/17
3;7	5/8	12/31	8/14	33/52	2/5	10/15
3;8	8/16	11/23	23/31	31/41	5/8	8/15
3;9	4/5	8/9	4/6	21/28	1/1	3/4
3;10	5/10	12/15	13/24	31/52	2/2	3/4
3;11	2/2	11/17	8/29	28/61	2/2	4/5
4;0	6/8	17/27	14/17	35/49	6/7	8/13
4;1	7/9	19/27	9/15	39/48	4/5	4/5
4;2	5/6	9/13	13/23	48/70	2/2	14/20
4;3	2/6	13/18	16/22	31/39	3/4	8/12
4;4	5/7	8/16	21/25	34/47	5/5	7/9

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
Korpus	JAN CS	JAN CDS	JAN CS	JAN CDS	JAN CS	JAN CDS
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
4;5	8/11	23/39	8/9	38/57	1/1	12/20
4;6	10/86	18/92	16/19	36/58	2/5	6/6
4;7	11/32	13/59	11/26	29/43	3/6	6/7
4;8	3/3	10/14	14/17	63/78	2/2	21/34
4;9	6/6	8/11	12/16	27/39	2/2	5/5
4;10	9/25	13/24	8/14	27/33	3/5	8/16
4;11	3/5	6/12	12/18	26/30	0/0	4/5
Gesamt	155/1.108	370/2.326	388/1.715	1.098/6.545	60/240	216/1.080

A5.2 KAT-Korpus – N, V, ADJ DERIV je Monat

Tabelle 54. KAT-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
Korpus	KAT CS	KAT CDS	KAT CS	KAT CDS	KAT CS	KAT CDS
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1;6	0/0	0/0	0/0	2/4	0/0	0/0
1;7			keine Aufnahme			
1;8	0/0	0/0	0/0	8/9	0/0	0/0
1;9	0/0	0/0	0/0	0/0	0/0	0/0
1;10	0/0	1/2	0/0	6/8	0/0	1/3
1;11	1/1	5/8	0/0	18/25	0/0	2/2
2;0	2/6	13/37	0/0	56/104	0/0	5/6
2;1	0/0	5/11	0/0	17/35	0/0	4/5
2;2	0/0	7/12	0/0	18/25	0/0	4/4
2;3	6/10	21/44	13/21	91/213	0/0	9/22
2;4	5/7	24/46	10/26	84/199	0/0	8/12
2;5	4/4	15/21	14/16	60/119	0/0	3/6
2;6	5/10	18/29	19/31	94/152	0/0	13/16
2;7			keine Aufnahme			
2;8	6/8	16/22	29/42	88/166	2/3	7/9
2;9	5/5	20/40	11/20	60/104	2/2	6/12
2;10	5/5	8/9	24/37	36/54	0/0	7/7
2;11	4/4	5/6	7/11	30/34	0/0	2/2
3;0	4/4	5/5	11/13	14/21	1/1	2/2
Gesamt	37/68	102/292	99/217	334/1.272	5/6	38/108

A5.3 LEN-Korpus – N, V, ADJ DERIV je Monat

Tabelle 55. LEN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) in CS und CDS

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
Korpus	LEN CS	LEN CDS	LEN CS	LEN CDS	LEN CS	LEN CDS
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1;7	0/0	17/36	0/0	64/101	0/0	3/4
1;8	1/1	18/31	2/2	73/153	0/0	16/19
1;9	4/9	27/57	3/4	91/195	0/0	8/12
1;10	3/29	27/69	2/2	97/178	0/0	17/20
1;11	4/12	14/23	2/5	46/73	0/0	1/1
2;0	1/3	13/22	3/3	26/47	0/0	3/3
2;1	3/3	15/24	1/1	53/79	0/0	6/6
2;2	6/26	7/21	8/12	43/71	0/0	2/2
2;3	4/11	11/17	15/40	61/99	0/0	6/6
2;4	6/13	12/29	13/22	61/116	0/0	2/2
2;5	6/16	10/17	20/47	50/77	0/0	2/3
2;6	3/6	6/11	23/40	61/110	0/0	3/3
2;7	5/8	10/15	25/38	65/126	1/1	4/4
2;8	2/2	17/35	26/45	69/140	1/1	6/8
2;9	6/11	16/45	20/41	65/137	1/1	7/11
2;10	5/5	13/19	28/40	50/108	0/0	3/6
2;11	6/13	12/18	21/36	42/62	0/0	3/3
3;0	9/15	17/39	30/55	58/103	2/4	8/10
Gesamt	36/183	122/528	132/433	449/1.975	4/7	60/123

A5.4 INPUT-Korpus – N, V, ADJ DERIV je Datenzeitpunkt

Tabelle 56. INPUT-Korpus - Gesamtübersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Erhebungszeitpunkt (1–4) in CS und CDS (HSES und LSES)

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	HSES CS	HSES CDS	HSES CS	HSES CDS	HSES CS	HSES CDS
Korpus INPUT	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
Datenpunkt						
1	66/148	119/331	178/388	362/1.029	12/36	38/107
2	69/174	118/350	182/381	405/1.143	21/31	52/105
3	66/266	113/397	204/430	385/1.035	25/50	62/137
4	81/241	123/285	271/538	442/1.112	34/72	68/134
HSES Gesamt	183/829	297/1.363	556/1.737	937/4.319	67/189	154/483
Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	LSES CS	LSES CDS	LSES CS	LSES CDS	LSES CS	LSES CDS
Korpus INPUT	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
Datenpunkt						
1	47/238	72/276	127/298	311/893	13/28	30/56
2	34/105	77/207	147/337	302/912	7/20	35/63
3	48/96	64/163	150/317	271/783	21/28	25/50
4	43/161	73/255	184/390	309/952	17/38	33/77
LSES Gesamt	106/600	172/901	409/1.342	731/3.540	44/114	89/246

Tabelle 57. INPUT-Korpus HSES - Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Kind und Erhebungszeitpunkt (1–4) in CS und CDS

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	Lemmas/Tokens		Lemmas/Tokens		Lemmas/Tokens	
	CS	CDS	CS	CDS	CS	CDS
D02-SIJ						
1	1/1	9/18	12/22	33/60	1/1	4/9
2	4/7	16/39	14/21	41/55	2/5	4/12
3	9/12	14/30	16/20	38/52	2/2	4/4
4	5/8	9/10	32/48	56/83	4/6	12/13
<i>D02-SIJ Gesamt</i>	<i>17/28</i>	<i>43/97</i>	<i>71/111</i>	<i>143/250</i>	<i>9/14</i>	<i>20/38</i>
D03-NIP						
1	9/10	14/30	3/3	28/43	1/1	2/3
2	5/14	17/32	9/14	42/60	3/3	2/3
3	7/11	7/11	10/11	37/55	2/4	3/4
4	7/16	8/18	12/16	23/33	5/7	5/8
<i>D03-NIP Gesamt</i>	<i>23/51</i>	<i>30/91</i>	<i>30/44</i>	<i>98/191</i>	<i>10/15</i>	<i>8/18</i>
D04-SOS						
1	3/6	6/13	12/18	44/52	4/6	6/6
2	7/9	18/31	14/18	64/101	2/2	9/10
3	12/24	17/21	23/26	58/86	1/1	11/12
4	10/20	7/11	45/74	38/65	3/5	2/3
<i>D04-SOS Gesamt</i>	<i>28/59</i>	<i>43/76</i>	<i>89/136</i>	<i>165/304</i>	<i>10/14</i>	<i>23/31</i>

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	Lemmas/Tokens		Lemmas/Tokens		Lemmas/Tokens	
D07-NIG	CS	CDS	CS	CDS	CS	CDS
1	6/13	15/20	18/22	54/76	2/4	15/25
2	6/8	13/28	17/23	61/96	2/2	8/14
3	5/10	11/19	25/33	57/82	2/2	10/23
4	7/13	20/43	18/20	53/75	3/4	11/16
<i>D07-NIG Gesamt</i>	<i>23/44</i>	<i>51/110</i>	<i>72/98</i>	<i>178/329</i>	<i>8/12</i>	<i>35/78</i>
D09-KAB	CS	CDS	CS	CDS	CS	CDS
1	4/7	9/10	13/16	34/59	0/0	4/4
2	5/8	7/14	12/13	40/73	1/1	5/6
3	1/1	8/8	17/17	34/47	4/6	7/12
4	8/15	9/12	22/31	34/51	7/15	5/5
<i>D09-KAB Gesamt</i>	<i>17/31</i>	<i>29/44</i>	<i>61/77</i>	<i>115/230</i>	<i>11/22</i>	<i>16/27</i>
D11-DAG	CS	CDS	CS	CDS	CS	CDS
1	11/22	19/52	22/33	43/73	0/0	2/2
2	9/14	14/29	17/26	46/59	2/2	3/4
3	7/9	11/25	11/13	47/69	1/2	3/3
4	10/29	12/27	25/31	54/97	2/2	5/5
<i>D11-DAG Gesamt</i>	<i>27/74</i>	<i>42/133</i>	<i>69/103</i>	<i>152/298</i>	<i>5/6</i>	<i>11/14</i>
D13-JOP	CS	CDS	CS	CDS	CS	CDS
1	9/28	14/25	17/22	26/36	2/7	2/2
2	7/14	12/20	26/40	43/65	1/1	2/2
3	2/4	8/12	14/23	40/52	2/2	3/3
4	7/22	6/6	31/48	54/66	2/2	3/4
<i>D13-JOP Gesamt</i>	<i>24/68</i>	<i>36/63</i>	<i>76/133</i>	<i>130/219</i>	<i>6/12</i>	<i>9/11</i>
D15-LUN	CS	CDS	CS	CDS	CS	CDS
1	2/2	3/4	18/29	38/61	0/0	7/24
2	5/13	11/25	20/38	52/125	0/0	6/6
3	13/49	15/74	23/33	49/82	1/1	4/6
4	4/5	12/19	11/12	60/94	5/6	9/13
<i>D15-LUN Gesamt</i>	<i>18/69</i>	<i>33/122</i>	<i>61/112</i>	<i>146/362</i>	<i>6/7</i>	<i>21/49</i>
D16-LEP	CS	CDS	CS	CDS	CS	CDS
1	3/6	8/20	13/16	55/92	1/2	3/5
2	2/3	12/21	8/16	48/98	1/3	8/14
3	6/10	10/11	8/12	42/71	2/3	8/19
4	11/23	11/17	16/21	64/103	3/3	7/14
<i>D16-LEP Gesamt</i>	<i>15/42</i>	<i>28/69</i>	<i>39/65</i>	<i>153/364</i>	<i>6/11</i>	<i>19/52</i>
D20-LUD	CS	CDS	CS	CDS	CS	CDS
1	8/11	19/35	14/32	56/132	1/1	5/7
2	1/1	5/9	17/32	52/69	0/0	8/8
3	12/47	16/93	24/59	39/116	1/1	6/7
4	6/14	18/29	17/35	44/101	3/3	7/10
<i>D20-LUD Gesamt</i>	<i>19/73</i>	<i>42/166</i>	<i>56/158</i>	<i>152/418</i>	<i>5/5</i>	<i>18/32</i>

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	Lemmas/Tokens		Lemmas/Tokens		Lemmas/Tokens	
D23-SOK	CS	CDS	CS	CDS	CS	CDS
1	5/12	9/17	20/35	31/68	2/2	1/1
2	9/29	8/12	31/50	42/54	3/5	4/5
3	9/28	7/11	13/33	28/40	4/5	3/3
4	3/3	5/6	24/27	46/59	2/2	5/5
<i>D23-SOK Gesamt</i>	<i>19/72</i>	<i>26/46</i>	<i>79/145</i>	<i>116/221</i>	<i>9/14</i>	<i>11/14</i>
D27-JOZ	CS	CDS	CS	CDS	CS	CDS
1	9/9	17/30	22/25	58/78	0/0	4/4
2	11/14	15/22	15/22	52/79	0/0	1/2
3	4/8	10/14	27/48	45/68	3/3	6/11
4	9/19	9/15	19/30	41/67	0/0	7/14
<i>D27-JOZ Gesamt</i>	<i>33/50</i>	<i>42/81</i>	<i>71/125</i>	<i>156/292</i>	<i>3/3</i>	<i>15/31</i>
D28-MOH	CS	CDS	CS	CDS	CS	CDS
1	5/6	6/12	40/58	43/71	3/11	4/7
2	9/14	8/9	27/32	54/84	1/1	4/6
3	8/16	11/17	39/50	40/73	5/8	6/9
4	11/25	12/18	53/91	31/56	5/8	7/7
<i>D28-MOH Gesamt</i>	<i>29/61</i>	<i>29/56</i>	<i>138/231</i>	<i>136/284</i>	<i>10/28</i>	<i>17/29</i>
D29-SAK	CS	CDS	CS	CDS	CS	CDS
1	6/6	18/29	24/36	43/65	0/0	4/4
2	12/14	21/41	20/24	54/89	1/1	3/3
3	7/23	12/34	13/17	57/85	0/0	8/9
4	5/5	19/28	20/28	64/99	4/4	7/10
<i>D29-SAK Gesamt</i>	<i>26/48</i>	<i>58/132</i>	<i>71/105</i>	<i>173/338</i>	<i>5/5</i>	<i>18/26</i>
D31-AND	CS	CDS	CS	CDS	CS	CDS
1	7/9	7/16	18/21	46/63	1/1	3/4
2	9/12	10/18	11/12	32/36	3/5	6/10
3	6/14	7/17	22/35	41/57	7/10	10/12
4	12/24	9/26	20/26	42/63	4/5	7/7
<i>D31-AND Gesamt</i>	<i>29/59</i>	<i>30/77</i>	<i>60/94</i>	<i>128/219</i>	<i>13/21</i>	<i>22/33</i>
HSES Gesamt	183/829	297/1.363	556/1.737	937/4.319	67/189	154/483

Tabelle 58. INPUT-Korpus LSES - Übersicht (Lemmas und Tokens) der derivierten Substantive (N DERIV), Verben (V DERIV) und Adjektive (ADJ DERIV) je Kind und Erhebungszeitpunkt (1–4) in CS und CDS

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	Lemmas/Tokens		Lemmas/Tokens		Lemmas/Tokens	
D05-FLB	CS	CDS	CS	CDS	CS	CDS
1	2/11	8/25	8/9	61/83	1/1	5/5
2	5/12	14/21	15/21	55/83	0/0	10/10
3	4/4	9/17	25/30	51/60	2/2	4/5
4	6/20	17/43	33/44	78/117	3/3	7/9
<i>D05-FLB Gesamt</i>	<i>14/47</i>	<i>33/106</i>	<i>73/104</i>	<i>198/343</i>	<i>6/6</i>	<i>23/29</i>
D06-LAB	CS	CDS	CS	CDS	CS	CDS
1	0/0	11/29	13/20	79/132	1/2	8/12
2	3/10	6/15	14/27	37/64	1/1	5/8
3	8/9	9/14	20/32	30/48	3/4	6/6
4	4/8	8/9	22/31	46/84	3/8	6/14
<i>D06-LAB Gesamt</i>	<i>13/27</i>	<i>26/67</i>	<i>61/110</i>	<i>147/328</i>	<i>7/15</i>	<i>22/40</i>
D08-FAJ	CS	CDS	CS	CDS	CS	CDS
1	5/57	2/9	19/28	35/42	1/2	4/5
2	7/11	4/5	35/55	31/51	0/0	2/2
3	7/13	3/3	29/45	15/21	1/1	0/0
4	5/7	2/2	25/39	26/41	3/5	3/6
<i>D08-FAJ Gesamt</i>	<i>20/88</i>	<i>10/19</i>	<i>95/167</i>	<i>86/155</i>	<i>5/8</i>	<i>9/13</i>
D10-LER	CS	CDS	CS	CDS	CS	CDS
1	11/25	12/32	15/17	42/68	2/4	4/5
2	1/1	11/27	17/29	56/107	0/0	1/1
3	7/13	9/19	5/6	47/95	1/1	2/3
4	7/12	15/31	19/22	48/75	0/0	2/3
<i>D10-LER Gesamt</i>	<i>22/51</i>	<i>40/109</i>	<i>46/74</i>	<i>158/345</i>	<i>3/5</i>	<i>8/12</i>
D12-JAS	CS	CDS	CS	CDS	CS	CDS
1	9/13	9/15	10/15	28/34	0/0	2/2
2	1/1	3/4	9/12	30/44	0/0	1/1
3	5/5	7/14	8/11	25/52	1/1	1/3
4	7/15	16/38	17/37	47/101	0/0	3/8
<i>D12-JAS Gesamt</i>	<i>17/34</i>	<i>28/71</i>	<i>41/75</i>	<i>105/231</i>	<i>1/1</i>	<i>6/14</i>
D14-KEW	CS	CDS	CS	CDS	CS	CDS
1	5/20	5/13	8/19	25/46	0/0	2/8
2	0/0	6/10	9/18	51/90	0/0	5/7
3	1/1	2/2	6/7	38/66	0/0	2/7
4	3/3	12/27	7/13	43/73	0/0	3/3
<i>D14-KEW Gesamt</i>	<i>9/24</i>	<i>19/52</i>	<i>26/57</i>	<i>112/275</i>	<i>0/0</i>	<i>11/25</i>
D17-TOB	CS	CDS	CS	CDS	CS	CDS
1	7/10	5/7	14/41	22/46	0/0	0/0
2	2/3	3/4	25/39	24/30	0/0	0/0
3	4/4	8/10	12/17	37/66	1/1	2/2
4	5/9	6/8	11/19	19/24	3/3	5/6
<i>D17-TOB Gesamt</i>	<i>15/26</i>	<i>20/29</i>	<i>58/116</i>	<i>80/166</i>	<i>4/4</i>	<i>7/8</i>

Wortklasse	N DERIV		V DERIV		ADJ DERIV	
	Lemmas/Tokens		Lemmas/Tokens		Lemmas/Tokens	
D18-ALM	CS	CDS	CS	CDS	CS	CDS
1	7/14	9/12	23/43	43/72	7/12	4/6
2	6/12	3/3	21/40	32/52	2/2	5/6
3	8/12	11/25	32/59	43/87	7/10	5/7
4	7/10	6/6	43/67	52/99	4/5	5/10
<i>D18-ALM Gesamt</i>	<i>24/48</i>	<i>26/46</i>	<i>102/209</i>	<i>126/310</i>	<i>15/29</i>	<i>17/29</i>
D19-VAS	CS	CDS	CS	CDS	CS	CDS
1	4/8	6/9	15/18	37/70	2/2	4/5
2	1/1	8/15	12/17	38/90	0/0	4/4
3	5/5	9/14	17/23	39/62	1/1	1/1
4	8/18	5/5	10/13	21/31	2/3	1/1
<i>D19-VAS Gesamt</i>	<i>15/32</i>	<i>24/43</i>	<i>49/71</i>	<i>106/253</i>	<i>5/6</i>	<i>9/11</i>
D21-PAS	CS	CDS	CS	CDS	CS	CDS
1	6/51	9/44	14/29	39/67	0/0	0/0
2	1/30	7/22	21/37	49/83	2/14	3/9
3	4/8	10/20	19/33	33/64	1/1	3/4
4	10/23	11/28	21/36	26/55	3/4	2/4
<i>D21-PAS Gesamt</i>	<i>16/112</i>	<i>30/114</i>	<i>65/135</i>	<i>117/269</i>	<i>6/19</i>	<i>7/17</i>
D22-LAS	CS	CDS	CS	CDS	CS	CDS
1	7/11	7/17	14/17	23/27	1/1	1/2
2	9/10	10/18	13/18	18/41	3/3	6/9
3	6/8	3/4	5/5	21/35	1/1	3/3
4	6/16	11/27	15/23	28/69	2/6	5/5
<i>D22-LAS Gesamt</i>	<i>23/45</i>	<i>26/66</i>	<i>45/63</i>	<i>67/172</i>	<i>7/11</i>	<i>11/19</i>
D24-RAG	CS	CDS	CS	CDS	CS	CDS
1	4/6	12/32	18/21	50/93	1/4	2/4
2	2/3	10/27	5/5	32/58	0/0	2/2
3	2/3	4/5	11/18	26/39	4/4	3/3
4	6/8	9/10	11/14	44/70	0/0	2/3
<i>D24-RAG Gesamt</i>	<i>14/20</i>	<i>32/74</i>	<i>38/58</i>	<i>128/260</i>	<i>4/8</i>	<i>7/12</i>
D25-PHG	CS	CDS	CS	CDS	CS	CDS
1	5/12	10/25	10/21	40/75	0/0	2/2
2	4/9	12/30	10/15	37/74	0/0	2/2
3	8/10	7/9	13/19	29/43	0/0	3/3
4	3/7	9/9	10/13	43/66	0/0	2/3
<i>D25-PHG Gesamt</i>	<i>19/38</i>	<i>36/73</i>	<i>41/68</i>	<i>125/258</i>	<i>0/0</i>	<i>7/10</i>
D26-JUB	CS	CDS	CS	CDS	CS	CDS
1	0/0	6/7	0/0	15/38	0/0	0/0
2	2/2	6/6	4/4	34/45	0/0	2/2
3	1/1	4/7	12/12	28/45	1/1	3/3
4	4/5	7/12	15/19	28/47	1/1	2/2
<i>D26-JUB Gesamt</i>	<i>7/8</i>	<i>20/32</i>	<i>30/35</i>	<i>83/175</i>	<i>2/2</i>	<i>7/7</i>
LSES Gesamt	106/600	172/901	409/1.342	731/3.540	44/114	89/246

Anhang 6 Komposita in CS und CDS – Übersichten

A6.1 JAN-Korpus

Tabelle 59. JAN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS

Korpus Alter	N KOMP		synth. KOMP		ADJ KOMP	
	JAN CS Lem./Tok.	JAN CDS Lem./Tok.	JAN CS Lem./Tok.	JAN CDS Lem./Tok.	JAN CS Lem./Tok.	JAN CDS Lem./Tok.
1;3	0/0	24/55	0/0	3/4	0/0	0/0
1;4	1/8	10/28	0/0	1/1	0/0	0/0
1;5	1/5	10/18	0/0	0/0	0/0	0/0
1;6	3/23	12/47	0/0	0/0	0/0	0/0
1;7	1/3	14/30	0/0	2/5	0/0	0/0
1;8	26/198	85/416	4/52	16/58	0/0	1/1
1;9	27/106	97/337	4/10	13/42	0/0	0/0
1;10	29/96	83/210	7/22	19/54	0/0	1/4
1;11	25/91	70/214	12/40	28/77	0/0	1/1
2;0	44/280	77/344	15/35	27/54	2/2	1/1
2;1	20/44	41/68	6/19	13/18	0/0	2/4
2;2	43/104	54/128	10/15	20/46	1/1	3/5
2;3	37/74	60/100	5/6	11/18	0/0	1/1
2;4	39/72	85/181	1/5	14/30	1/1	2/3
2;5	22/33	88/146	3/4	13/14	0/0	2/3
2;6	36/69	53/117	8/13	14/22	0/0	1/1
2;7	13/37	50/83	4/5	14/28	0/0	1/1
2;8	30/57	68/125	6/13	14/25	0/0	2/3
2;9	13/27	41/87	3/3	14/18	0/0	4/4
2;10	18/29	53/102	3/3	8/9	0/0	1/2
2;11	15/42	39/67	2/2	7/9	0/0	0/0
3;0	19/30	57/116	4/5	10/13	2/2	2/4
3;1	5/11	25/39	0/0	5/5	0/0	1/1
3;2	0/0	8/8	1/3	2/5	0/0	0/0
3;3	2/7	14/21	0/0	5/5	0/0	0/0
3;4	0/0	5/6	0/0	0/0	0/0	0/0
3;5	3/21	18/42	1/2	6/8	0/0	0/0
3;6	9/10	27/40	1/1	8/8	0/0	0/0
3;7	1/2	7/8	1/2	3/3	1/1	0/0
3;8	8/17	15/25	4/4	3/4	0/0	0/0
3;9	4/4	12/13	1/1	6/6	0/0	0/0
3;10	3/6	7/12	2/5	2/2	0/0	0/0
3;11	9/11	16/29	1/1	4/8	0/0	0/0
4;0	3/4	22/31	3/5	4/5	0/0	0/0
4;1	10/15	14/24	0/0	6/11	0/0	0/0
4;2	9/13	17/23	4/7	7/14	5/5	0/0
4;3	10/17	16/24	5/11	9/11	0/0	1/1
4;4	12/33	18/25	1/1	4/4	0/0	0/0

Korpus	N KOMP		synth. KOMP		ADJ KOMP	
	JAN CS	JAN CDS	JAN CS	JAN CDS	JAN CS	JAN CDS
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
4;5	9/12	32/55	2/2	6/12	0/0	0/0
4;6	0/0	4/9	0/0	1/1	0/0	0/0
4;7	9/17	6/8	2/2	3/3	0/0	0/0
4;8	5/6	23/40	2/3	6/7	0/0	1/1
4;9	3/8	5/6	1/1	3/4	3/4	0/0
4;10	7/10	9/15	2/2	5/6	0/0	0/0
4;11	2/2	5/6	0/0	0/0	0/0	0/0
Gesamt	295/1.654	832/3.528	80/305	241/677	14/16	25/41

A6.2 KAT-Korpus

Tabelle 60. KAT-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS

Korpus	N KOMP		synth. KOMP		ADJ KOMP	
	KAT CS	KAT CDS	KAT CS	KAT CDS	KAT CS	KAT CDS
Alter	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1;6	0/0	2/4	0/0	0/0	0/0	0/0
1;7	k e i n e A u f n a h m e					
1;8	0/0	1/1	0/0	0/0	0/0	0/0
1;9	0/0	0/0	0/0	0/0	0/0	0/0
1;10	0/0	1/2	0/0	0/0	0/0	0/0
1;11	0/0	6/11	0/0	0/0	0/0	0/0
2;0	0/0	14/24	0/0	1/1	0/0	0/0
2;1	1/1	9/20	0/0	0/0	0/0	0/0
2;2	2/11	2/9	0/0	0/0	0/0	0/0
2;3	6/11	34/57	1/2	6/11	1/1	1/1
2;4	7/13	32/51	0/0	4/6	0/0	0/0
2;5	5/12	18/26	1/2	7/8	0/0	1/1
2;6	9/10	28/44	0/0	2/6	0/0	1/1
2;7	k e i n e A u f n a h m e					
2;8	10/14	33/45	1/2	4/4	0/0	2/3
2;9	7/10	24/34	1/1	7/10	0/0	0/0
2;10	8/10	16/17	2/3	3/3	0/0	0/0
2;11	10/16	10/13	1/1	3/3	0/0	0/0
3;0	10/10	14/14	11/19	30/56	0/0	0/0
Gesamt	66/119	196/372	11/19	30/56	1/1	5/6

A6.3 LEN-Korpus

Tabelle 61. LEN-Korpus - Monatliche Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita in CS und CDS

Korpus Alter	N KOMP		synth. KOMP		ADJ KOMP	
	LEN CS Lem./Tok.	LEN CDS Lem./Tok.	LEN CS Lem./Tok.	LEN CDS Lem./Tok.	LEN CS Lem./Tok.	LEN CDS Lem./Tok.
1;7	0/0	20/34	0/0	3/6	0/0	0/0
1;8	1/1	26/52	0/0	7/10	0/0	0/0
1;9	1/2	44/77	0/0	9/10	0/0	1/1
1;10	1/1	19/34	0/0	1/1	0/0	0/0
1;11	0/0	20/36	0/0	1/1	0/0	0/0
2;0	1/1	12/18	0/0	2/2	0/0	0/0
2;1	1/6	11/18	0/0	0/0	0/0	0/0
2;2	1/1	16/23	0/0	3/4	0/0	0/0
2;3	3/18	15/28	0/0	3/4	0/0	0/0
2;4	4/4	19/24	0/0	4/4	0/0	0/0
2;5	0/0	7/10	0/0	1/2	0/0	0/0
2;6	5/7	14/33	0/0	0/0	0/0	0/0
2;7	2/2	20/26	0/0	3/3	0/0	0/0
2;8	4/4	12/29	0/0	2/2	0/0	0/0
2;9	8/20	26/52	0/0	1/1	0/0	0/0
2;10	13/25	17/31	0/0	3/3	0/0	0/0
2;11	2/5	15/16	1/1	1/2	0/0	0/0
3;0	4/7	27/51	0/0	3/3	0/0	0/0
Gesamt	48/105	206/592	1/1	38/58	0/0	1/1

A6.4 INPUT-Korpus

Tabelle 62. INPUT-Korpus - Übersicht (Lemmas und Tokens) der Nominalkomposita, synthetischen Komposita und Adjektivkomposita je Erhebungszeitpunkt (1–4) in CS und CDS (HSES und LSES)

Korpus INPUT	N KOMP		synth. KOMP		ADJ KOMP	
	HSES CS	HSES CDS	HSES CS	HSES CDS	HSES CS	HSES CDS
Datenpunkt	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1	108/192	189/384	19/35	36/70	2/2	5/6
2	117/191	199/416	18/23	43/69	1/1	8/12
3	126/204	216/356	18/32	45/59	7/9	10/10
4	130/224	225/409	30/49	47/68	4/10	5/8
HSES Gesamt	407/811	686/1.565	72/139	148/266	13/22	24/36

Korpus INPUT	N KOMP		synth. KOMP		ADJ KOMP	
	LSES CS	LSES CDS	LSES CS	LSES CDS	LSES CS	LSES CDS
Datenpunkt	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.	Lem./Tok.
1	55/91	94/165	11/16	20/31	1/1	1/1
2	51/97	95/213	7/10	25/43	0/0	4/4
3	43/65	87/137	11/14	13/18	3/5	6/8
4	64/126	90/171	8/13	26/33	2/2	4/7
LSES Gesamt	183/379	295/686	36/53	75/125	5/8	11/20

Anhang 7 Ableitungsbasen

A7.1 Basen derivierter Substantive (Tokens)

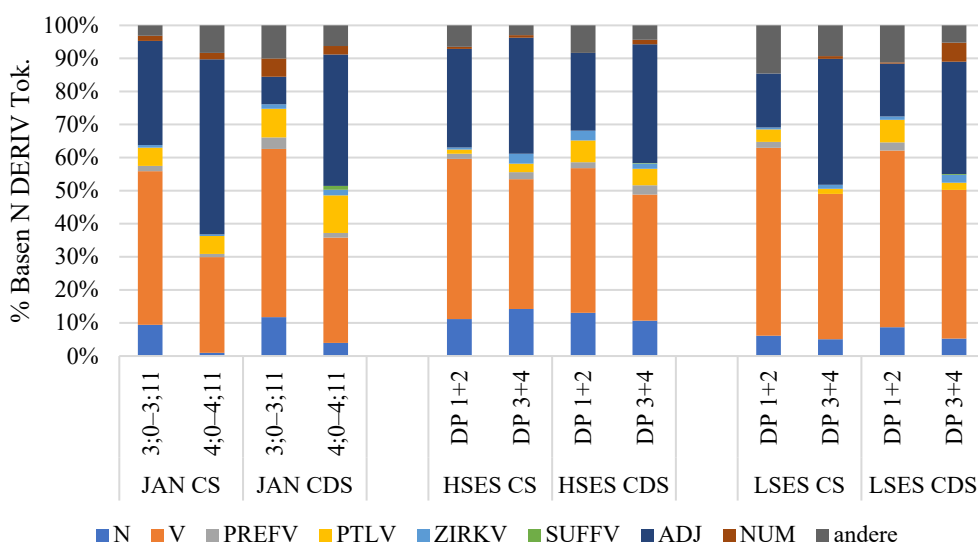


Abbildung 64. Ableitungsbasen derivierter Substantive (Tokens) im JAN-Korpus (3;0–4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

A7.2 Basen derivierter Verben (Tokens)

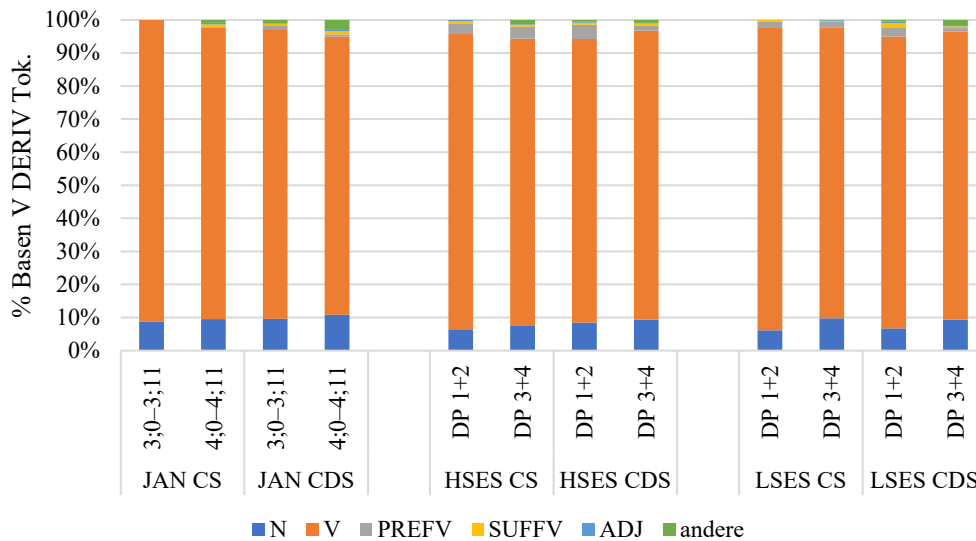


Abbildung 65. Ableitungsbasen derivierter Verben (Tokens) im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

A7.3 Basen derivierter Adjektive (Tokens)

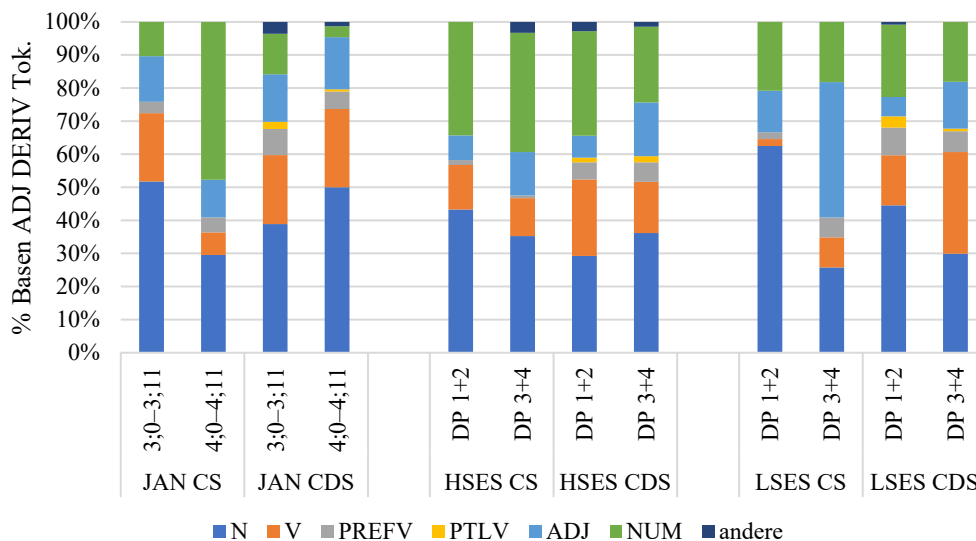


Abbildung 66. Ableitungsbasen derivierter Adjektive (Tokens) im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4) in CS und CDS

Anhang 8 Komplexität der Wortbildungsmuster

A8.1 Nominale Wortbildungsmuster in CS und CDS (Tokens)

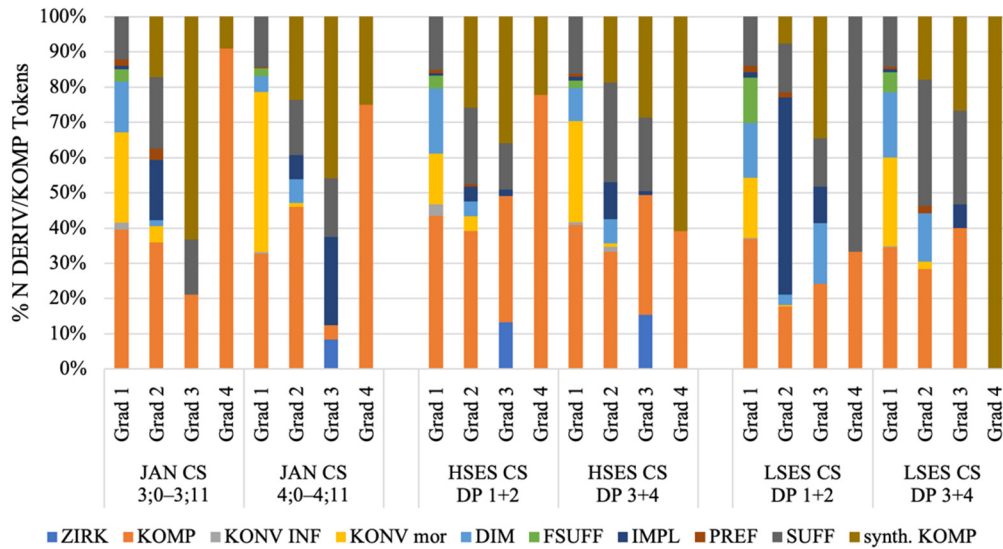


Abbildung 67. Prozentuale Verteilung nominaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

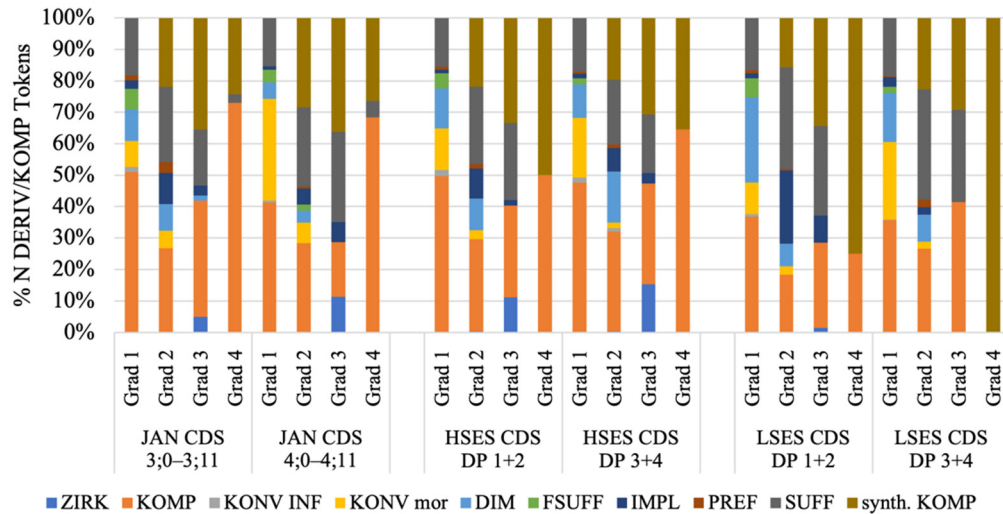


Abbildung 68. Prozentuale Verteilung nominaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

A8.2 Verbale Wortbildungsmuster in CS und CDS (Tokens)

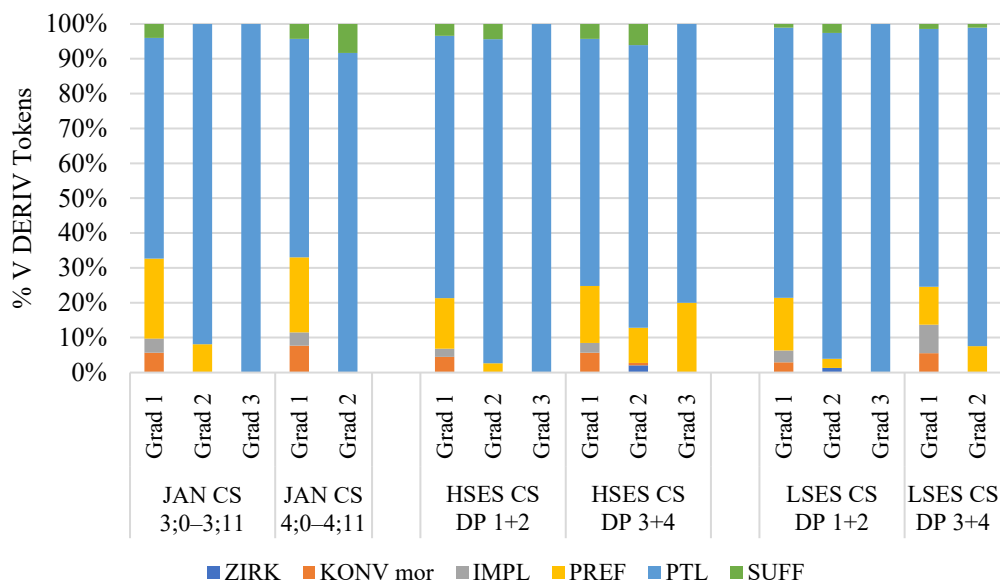


Abbildung 69. Prozentuale Verteilung verbaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

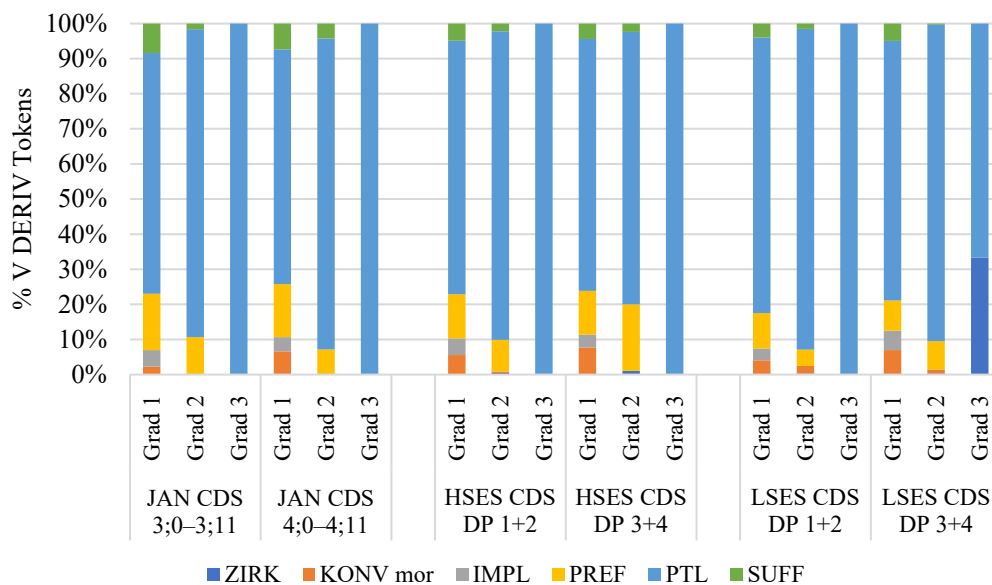


Abbildung 70. Prozentuale Verteilung verbaler Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

A8.3 Adjektivische Wortbildungsmuster in CS und CDS (Tokens)

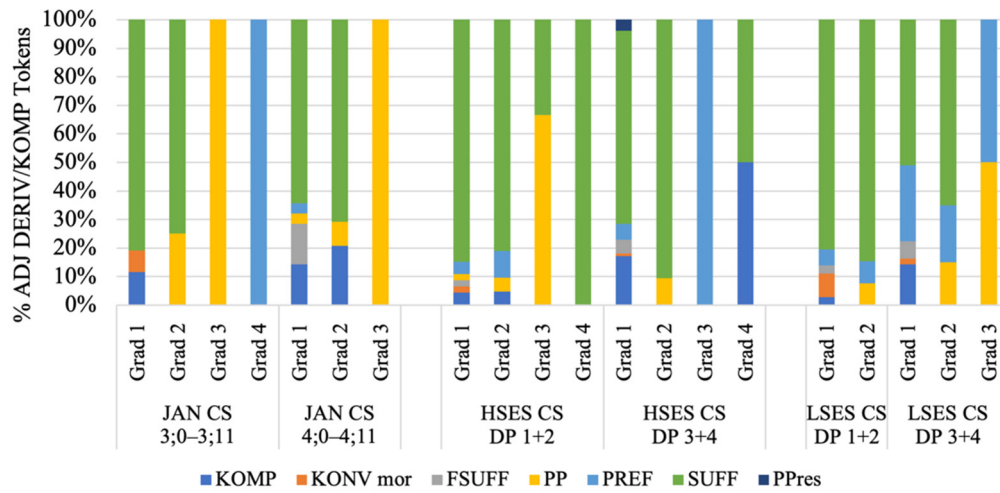


Abbildung 71. Prozentuale Verteilung adjektivischer Wortbildungsmuster (Tokens) je Komplexitätsgrad in CS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

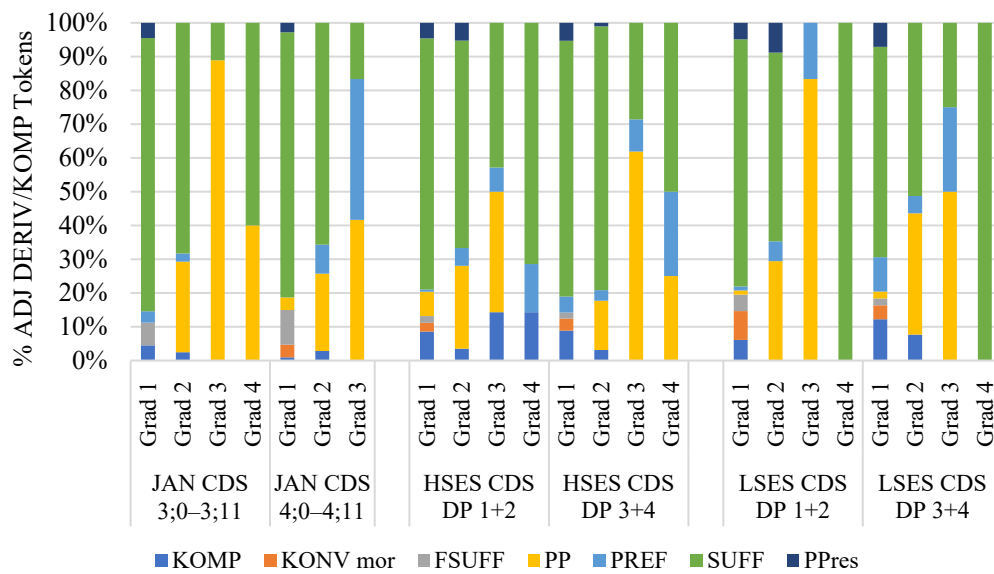


Abbildung 72. Prozentuale Verteilung adjektivischer Wortbildungsmuster (Tokens) je Komplexitätsgrad in CDS im JAN-Korpus (3;0-4;11) und INPUT-Korpus (DP 1+2, 3+4)

Anhang 9 Komplexität der Wortbildungsprodukte

Die Komplexitätsgrade abgeleiteter Substantive, Verben und Adjektive in CS und CDS wurden in Relation zur jeweiligen Grundgesamtheit aller Substantiv-, Verb- und Adjektivtokens berechnet, um die prozentuale Verteilung in den Korpora darzustellen. Die morphologischen Komplexitätsgrade wurden anhand der in Tabelle 5 aufgelisteten Grade bestimmt. Daher entspricht Grad 0 pseudo-derivierten Lexemen (z. B. **verlier-**, **gewinn-**, **fertig**), Grad 1 der Applikation eines Wortbildungsprozesses (z. B. ein Suffix, Konversion, Stammvokalveränderung, Komposition), Grad 2, 3 und 4 zwei, drei, vier oder mehr involvierte Wortbildungsprozesse.

A9.1 JAN-Korpus

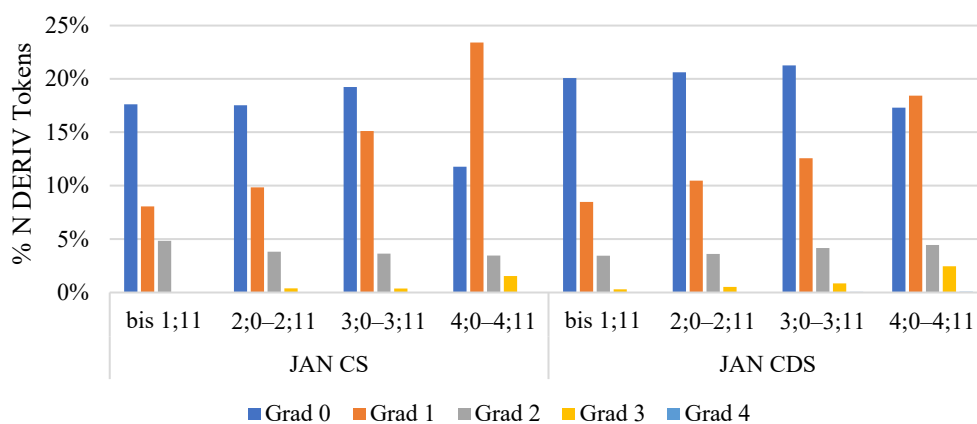


Abbildung 73. JAN-Korpus: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS

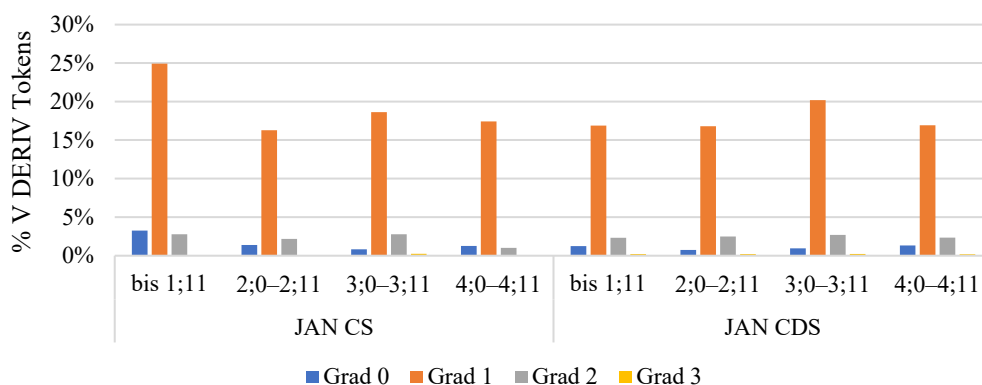


Abbildung 74. JAN-Korpus: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS

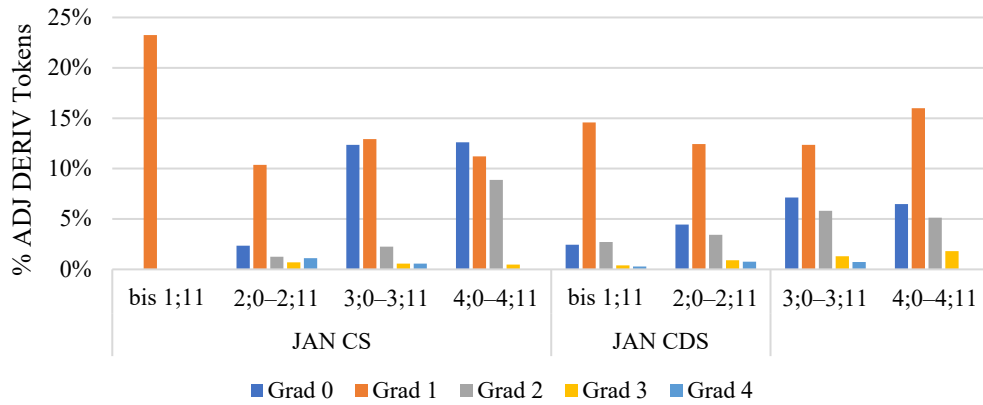


Abbildung 75. JAN-Korpus: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS

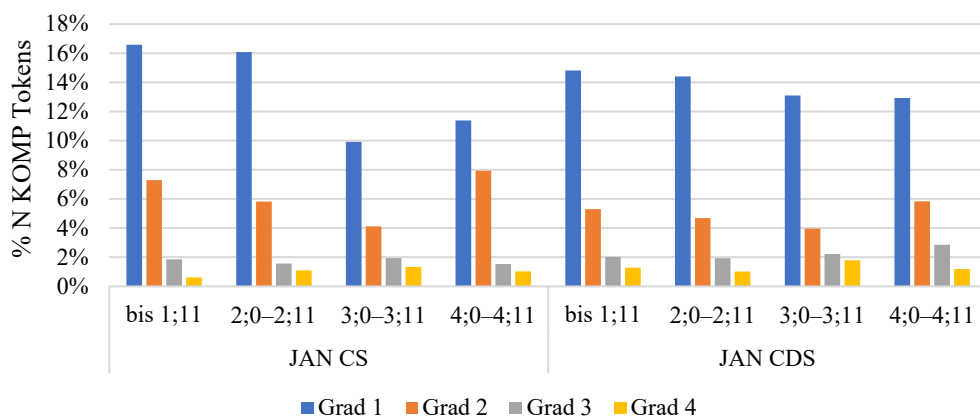


Abbildung 76. JAN-Korpus: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS

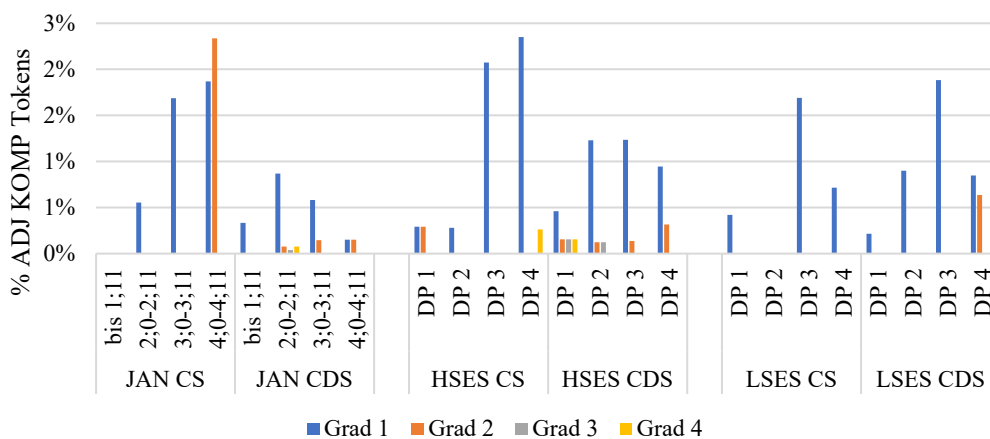


Abbildung 77. JAN- und INPUT-Korpus: Komplexitätsgrade von Adjektivkomposita (Tokens) in CS und CDS

A9.2 KAT-Korpus und LEN-Korpus

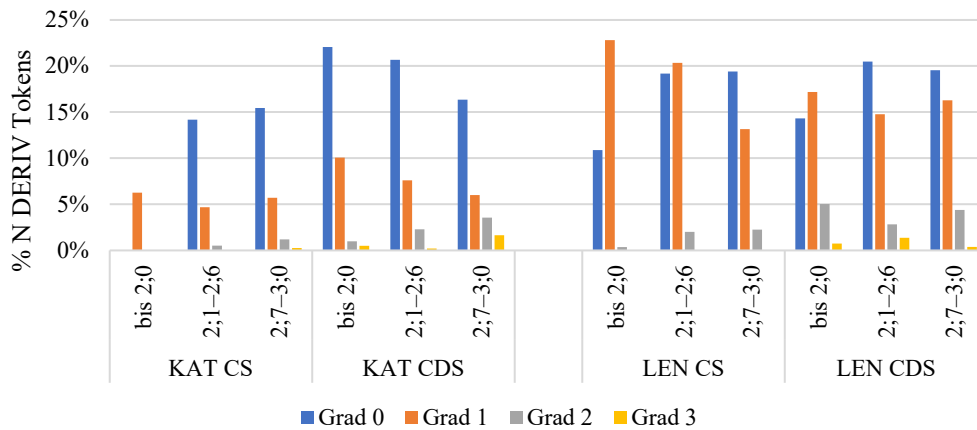


Abbildung 78. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS

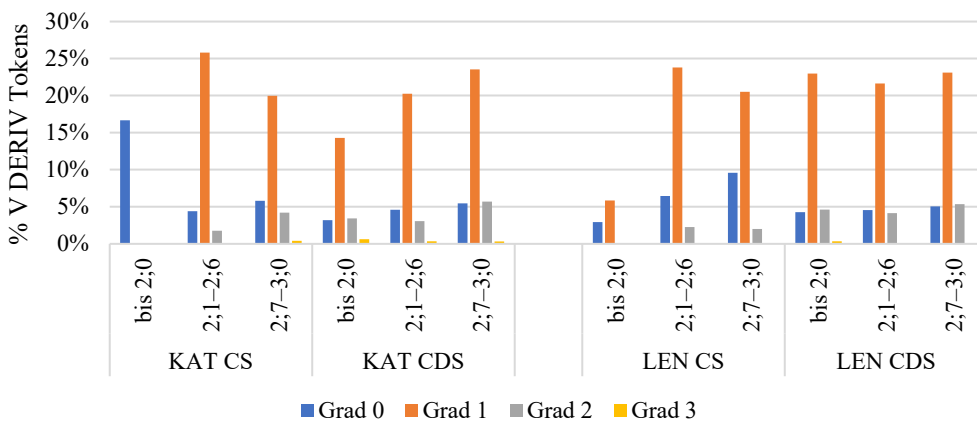


Abbildung 79. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS

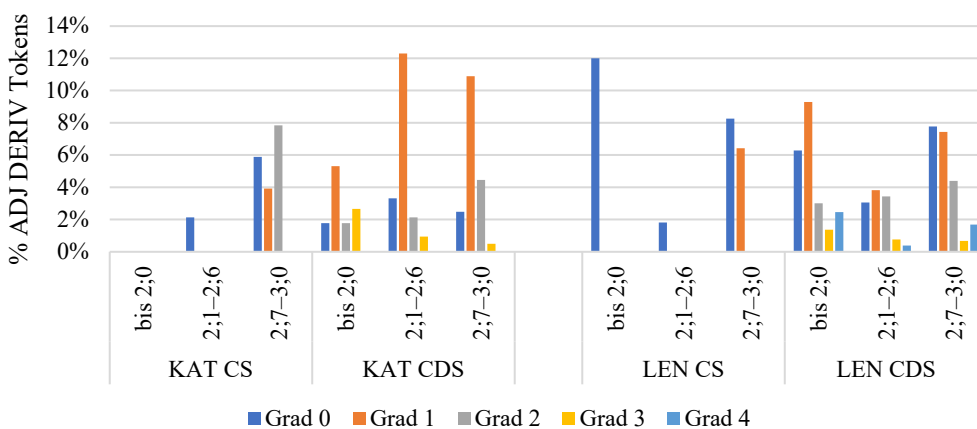


Abbildung 80. KAT- und LEN-Korpora: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS

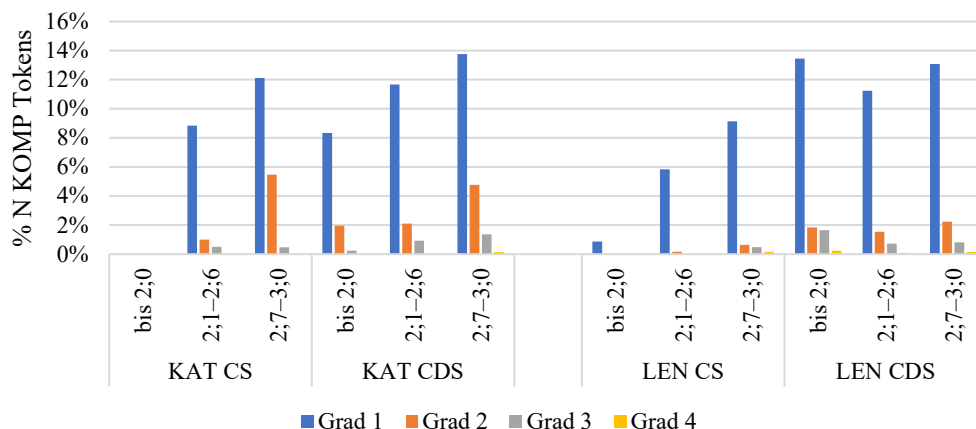


Abbildung 81. KAT- und LEN-Korpora: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS

A9.3 INPUT-Korpus

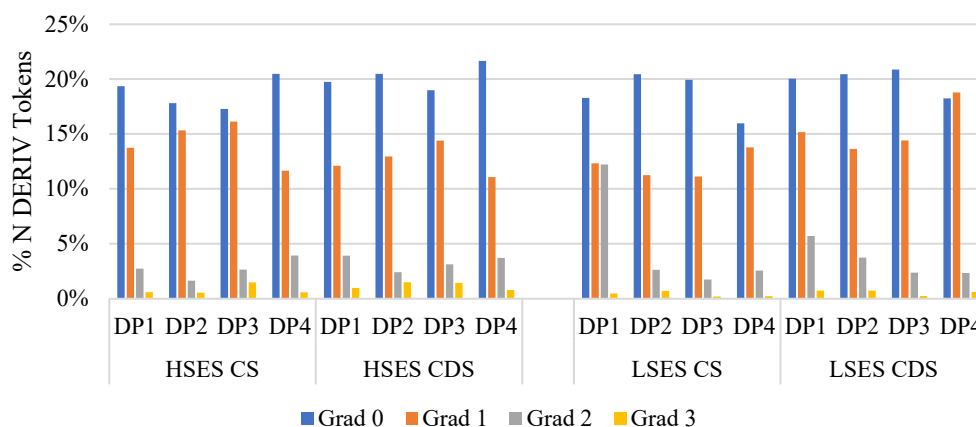


Abbildung 82. INPUT-Korpus: Komplexitätsgrade derivierter Substantive (Tokens) in CS und CDS

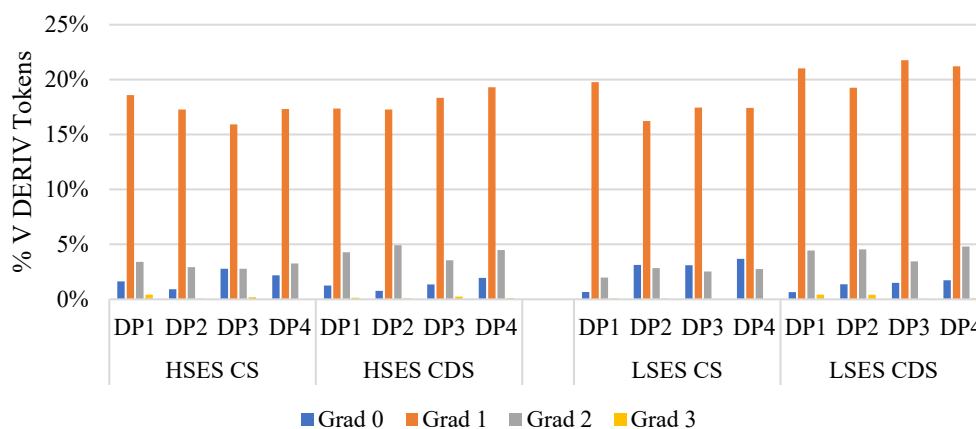


Abbildung 83. INPUT-Korpus: Komplexitätsgrade derivierter Verben (Tokens) in CS und CDS

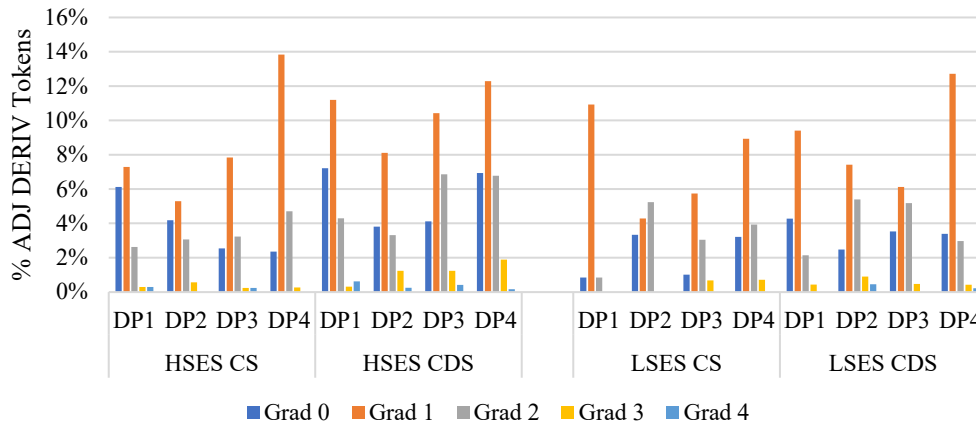


Abbildung 84. INPUT-Korpus: Komplexitätsgrade derivierter Adjektive (Tokens) in CS und CDS

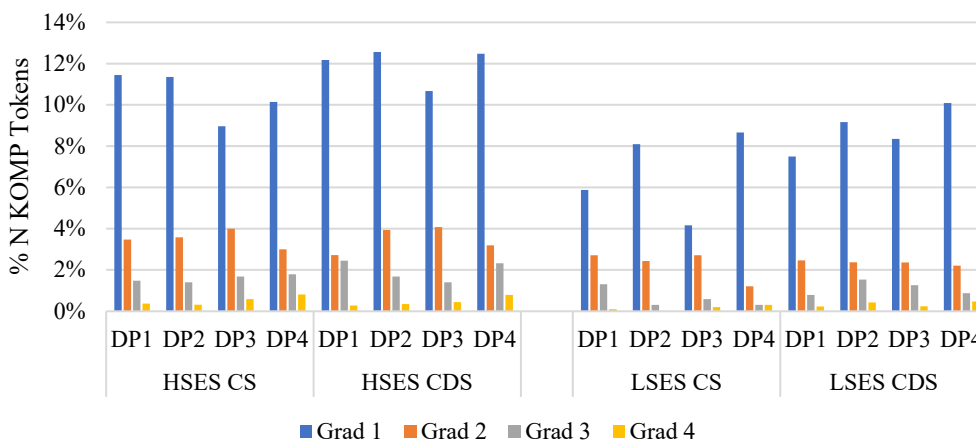


Abbildung 85. INPUT-Korpus: Komplexitätsgrade nominaler Komposita (Tokens) in CS und CDS

Anhang 10 Elizitationsexperiment

A10.1 Experimentalblatt

Auf dem Experimentalblatt wurde der Test zusätzlich zur Audio- und Videoaufnahme von der durchführenden Person schriftlich dokumentiert. Dazu wurden den Experimentatorinnen zwei Experimentalblätter zur Verfügung gestellt. Eines davon enthielt neben den Übungs- und Testsätzen ebenfalls eine ausführliche und genaue Anleitung zur Durchführung des Experiments wie in Abbildungen 86 und 87 ersichtlich. Dabei entspricht die Auflistung der Übungs- und Testsätze der Reihenfolge der Abfrage.

Derivationsexperiment	Datum
Experimentier:	Kind: Jahre

Experimentier, bitte das heutige Datum, den eigenen Namen, den Namen und das Alter des Kindes angeben

Lieber Experimentier, Bitte lies die Sätze deutlich vor, betone das zu derivierende Verb und gegebenenfalls kursiv geschriebene Wörter. Bitte achte darauf, die Sätze spannend zu formulieren! (Besonders möchte ich auf Satz 31 hinweisen, der ohne die Betonung von „darin“ möglicherweise die Bildung des Partizips eliziert)

→ Es hat sich beim Pilotexperiment gezeigt, dass die Kinder darauf achten, was man wie mitschreibt. Macht man zuviele Hakerl für richtige Antworten und schreibt dann einige Wörter auf, werden die Kinder unsicher. Daher besser jede Antwort kurz mitschreiben.

Einleitung: Jetzt würde ich mich wahnsinnig freuen, wenn Du bei dem Spiel hier auch so super mitmachst! Pass auf, ich erkläre dir wie das geht! Ich lese dir was vor und Du hilfst mir bitte den Satz fertig zu machen. Ich geb’ dir mal ein paar Beispiele:

A) Der Mann möchte gerne fahren. Er ist ein toller ... (Fahrer).
--> nach toller kurz warten, ob Kind etwas sagt, falls nicht vervollständigen!

B) Die Frau spielt sehr gerne. Sie ist ein gute ...? (Spielerin)
--> im Pilotexperiment haben die meisten Kinder hier schon von sich aus Spielerin gesagt, falls Kind nicht antwortet dann bitte selbst vervollständigen und nachfragen, ob es klar ist

C) Der Mann möchte den Bleistift spitzen. Dazu braucht er einen ...? (Spitzer)

D) Die tanzende Frau dreht sich. Sie macht eine schwungvolle ...? (Drehung)

wenn alles klar ist und das Kind die restlichen Sätze vervollständigt hat, man also das Gefühl hat die Aufgabe wurde verstanden, weiter mit:

..dann gehts jetzt los..:

6. Der Mann arbeitet viel. Er ist ein braver ...?	_____
17. Der Bub soll etwas rechnen. Dazu verwendet er einen ...?	_____
27. Am Montag gehen sie wandern. Sie unternehmen eine lustige ...?	_____
7. Die Frau geht gerne schwimmen. Sie ist eine tolle ...?	_____
21. Die Mutter möchte etwas mixen. Dazu braucht sie einen ...?	_____
1. Der Mann geht gerne laufen. Er ist ein guter ...?	_____
30. Das Flugzeug landet ganz sicher. Das war eine fabelhafte ...?	_____
19. Das Mädchen muss etwas radieren. Dazu braucht sie einen ...?	_____
10. Die Frau geht gerne tanzen. Sie ist ein gute ...?	_____
31. Bis Donnerstag muss der Bub noch üben. Dann hat er darin ganz viel ...?	_____
16. Der Vater muss in der Früh geweckt werden. Dazu hat er einen ...?	_____
28. Das Mädchen ist verkühlt. Sie leidet an einer ...?	_____

Abbildung 86. Experimentalblatt mit detaillierter Anleitung für den Experimentator zum Elizitationstest nominaler Ableitungen, Seite 1

Derivationsexperiment	Datum
2. Der Mann sammelt gerne Dinge. Er ist ein fleißiger ...?	_____
22. Die Frau möchte sich die Schuhe abstreifen. Dazu braucht sie einen ...?	_____
25. Die Feuerwehrmänner retten die Frau. Das war eine erfolgreiche ...?	_____
8. Die Frau verkauft jeden Tag etwas. Sie ist eine gute ...?	_____
3. Der Mann zaubert gerne. Er ist ein großer ...?	_____
29. Das Kind wird vom Opa begleitet. Das ist eine <i>sichere</i> ...?	_____
18. Der Bub möchte etwas tosten. Dazu braucht er einen ...?	_____
9. Die Frau lehrt in der Schule einiges. Sie ist eine gute ...?	_____
13. Der Mann soll ein Loch bohren. Dazu verwendet er einen ...?	_____
35. Das Mädchen wird befreit. Das war eine gelungene ...?	_____
14. Der Vater soll etwas öffnen. Dazu verwendet er einen ...?	_____
4. Der Mann geht gerne jagen. Er ist ein guter ...?	_____
24. Die Mutter entsaftet das Obst. Dazu nimmt sie einen ...?	_____
26. Der Bub hat sich verletzt. Er hat eine <i>schwere</i> ...?	_____
11. Die Frau malt schön. Sie ist eine tolle ...?	_____
32. Das Kind soll die Fehler verbessern. Emsig macht es die ...?	_____
20. Die Frau möchte etwas kleben. Dazu verwendet sie einen ...?	_____
33. Am Freitag wird das Mädchen geprüft. Sie lernt schon fleißig für die ...?	_____
5. Der Mann kann vieles backen. Er ist ein guter ...?	_____
23. Das Mädchen möchte etwas lochen. Dazu verwendet sie einen ...?	_____
34. Der Junge atmet sehr schnell. Er hat eine <i>ganz</i> schnelle ...?	_____
12. Die Frau kann wundervoll singen. Sie ist eine großartige ...?	_____
15. Der Mann soll die Wiese mähen. Dazu verwendet er einen ...?	_____
36. Am Dienstag wird der Arzt impfen. Die Kinder bekommen dann eine ...?	_____

VIELEN DANK, DASS DU SO TOLL MITGEMACHT HAST!!!!

Abbildung 87. Experimentalblatt zum Elizitationstest nominaler Ableitungen, Seite 2

A10.2 Codierungsschema für die Datenerfassung

Die Daten des Elizitationsexperiments zur Derivation wurden anhand des Codierungsschemas erfasst und ausgewertet. Neben den in Tabelle 63 gelisteten Einträge wurden überdies die Fallnummer, das Kürzel für die Experimentatorin, das Testdatum, sowie das anonymisierte Namenskürzel des Kindes, dessen Geburtsdatum, SES-Zugehörigkeit, Geschlecht, Erstsprache und Alter erfasst.

Tabelle 63. Codierungsschema für die Datenerfassung des Derivationsexperiments mit Beispielen

Label	Bezeichnung	Werte	Beispiel korrekt	inkorrekt
X_base_derivative_txt	gebildetes Wort (Antwort) X steht für die jeweilige Nummer des Testitems, base steht für die Ableitungsbasis	text	Entsafter	Wunde
X_base_deriv_genus	Genus des gebildeten Wortes	0 kein Genus (Verb) 1 maskulin 2 feminin 3 neutrum	1	2
X_base_deriv_category	Wortbildungskategorie des gebildeten Wortes	0 keine Kategorie 1 Nomen Agentis 2 Nomen Instrumenti 3 Nomen acti/actionis 4 Nomen loci	2	3
X_base_deriv_suffix	verwendetes Ableitungssuffix	0 kein Suffix 1 -er für NAG 2 -er für NIN oder NAC 3 -er-in 4 -ung 5 -el 6 -e 7 -heit 8 -in	2	6
X_base_deriv_umlaut	erhielt das Derivat einen Umlaut?	0 nein 1 ja	0	0
X_base_deriv_corr	das gebildete Wort ist das korrekt abgeleitete Zielwort	0 nein 1 ja	1	0
X_base_deriv_semicorr	das gebildete Wort enthält die richtige Ableitung, ist aber nicht das Zielwort	0 nein 1 ja	0	0
X_base_deriv_suff	wurde das korrekte Ableitungssuffix verwendet?	0 nein 1 ja	1	0
X_base_deriv_uml	wurde richtig umgelaute? wurde der Umlaut korrekt eingesetzt?	0 nein 1 ja	1	1
X_base_deriv_base	wurde die korrekte Ableitungsbasis verwendet?	0 nein 1 ja	1	0
X_base_deriv_error	Fehlercodierung für inkorrekte und semikorrekte Ableitungen	siehe Tabelle 28	0	10
X_base_autocorr	hat sich das Kind spontan selbst korrigiert?	0 nein 1 ja	0	0

A10.3 Frequenzen semikorrekter Derivate

Tabelle 64. Individuelle und Gesamthäufigkeiten der semikorrekten Ableitungen im Elizitationsexperiment in den Wortbildungskategorien nach SES-Zugehörigkeit

	Kind	NAG		NIN	NAC	Gesamt
		-er	-er-in	-er	-ung	
LSES	D06-LAB			3		3
	D12-JAS			1	1	2
	D18-ALM			2		2
	D19-VAS			3		3
	D21-PAS	1	1	3	3	8
	<i>LSES Ges.</i>	<i>1/15</i>	<i>1/15</i>	<i>12/60</i>	<i>4/60</i>	<i>18/180</i>
HSES	D02-SIJ			3	1	4
	D03-NIP			1	1	2
	D04-SOS			1		1
	D07-NIG			3		3
	D09-KAB	1				1
	D13-JOP	1	1	1	1	4
	D20-LUD			2		2
	D23-SOK			2	1	3
	D28-MOH			1		1
	D29-SAK			2		2
	D31-AND					0
	<i>HSES Ges.</i>	<i>2/33</i>	<i>1/33</i>	<i>16/132</i>	<i>4/132</i>	<i>23/396</i>
Gesamt	3/48	2/48	28/192	8/192	41/576	

Anhang 11 Paneldaten (Kindergarten)

Neben den Aufnahmen, die bei den Familien zuhause stattfanden, wurden ebenfalls Sprachaufnahmen im Kindergarten durchgeführt (siehe Tabelle 65). Diese Daten können Aufschluss darüber geben, wie sich ein Kind in Interaktion mit der primären Kindergartenpädagogin (KG1) im Vergleich zu seiner Interaktion mit der Hauptbezugsperson verhält. Dies kann vor allem im Hinblick auf den sprachlichen Input eine Rolle spielen in Bezug auf die Sprachentwicklung von Kindern aus bildungsfernen Familien (LSES).

Von dem Kind (D31-AND) sind leider aufgrund von Differenzen mit der Kindergartenleitung keine Kindergartenaufnahmen vorhanden. Bei dem Kind (D16-LEP) waren die Kindergartenpädagoginnen gleichermaßen auf den Aufnahmen vorhanden, weshalb keine eindeutige Zuordnung in primäre und sekundäre Kindergartenpädagogin möglich war. Deshalb wurden in diesem einen Fall die Daten von KG1 und KG2 zusammengezählt.

Tabelle 65. Paneldaten im Kindergarten CS und CDS (INPUT-Korpus)

Korpus INPUT		Erhebungszeitpunkt				Anz. Äußerungen		Anz. Worttokens	
		1	2	3	4	CS	CDS	CS	CDS
Kindergarten						(KG1)		(KG1)	
D02-SIJ	(m/H)	3;2	3;5	4;6	4;11	804	2.013	2.364	7.610
D03-NIP	(m/H)	3;1	3;4	4;5	4;9	339	1.498	1.085	6.497
D04-SOS	(w/H)	3;2	3;6	4;6	4;9	205	1.724	676	6.557
D07-NIG	(w/H)	2;11	3;2	4;2	4;6	751	2.162	2.144	8.500
D09-KAB	(w/H)	3;1	3;5	4;4	4;8	400	1.984	1.514	8.032
D11-DAG	(m/H)	3;1	3;5	4;6	4;9	290	2.107	629	8.227
D13-JOP	(m/H)	2;11	3;3	4;4	4;8	856	1.992	2.648	7.329
D15-LUN	(m/H)	3;0	3;3	4;3	4;8	203	1.787	483	6.626
D16-LEP	(w/H)	3;2	3;4	4;4	4;7	424	1.328	1.102	5.235
D20-LUD	(w/H)	3;2	3;6	4;6	4;9	683	2.345	1.905	8.277
D23-SOK	(w/H)	3;0	3;3	4;3	4;5	796	2.195	2.215	7.003
D27-JOZ	(m/H)	3;0	3;2	4;2	4;6	839	2.441	2.412	9.958
D28-MOH	(m/H)	3;0	3;4	4;4	4;8	1.045	1.894	3.400	7.111
D29-SAK	(w/H)	3;1	3;6	4;5	4;9	576	2.354	1.527	9.831
Summe HSES						8.211	27.824	24.104	106.793
D05-FLB	(m/L)	2;11	3;2	4;3	4;6	482	2.189	1.732	7.728
D06-LAB	(w/L)	3;1	3;5	4;5	4;8	797	2.499	1.932	8.130
D08-FAJ	(m/L)	3;0	3;3	4;3	4;6	685	2.024	1.735	7.739
D10-LER	(w/L)	3;1	3;4	4;4	4;7	588	2.683	1.520	9.617
D12-JAS	(w/L)	3;1	3;4	4;6	4;9	105	1.604	244	6.564
D14-KEW	(m/L)	3;1	3;4	4;4	4;7	280	2.130	577	7.935
D17-TOB	(m/L)	3;1	3;3	4;4	4;8	589	2.063	1.691	8.508
D18-ALM	(w/L)	3;0	3;4	4;4	4;7	812	2.230	2.580	8.644
D19-VAS	(w/L)	3;1	3;5	4;5	4;11	381	2.190	1.006	7.680
D21-PAS	(m/L)	3;3	3;6	4;6	4;9	1109	2.645	2.893	9.674
D22-LAS	(w/L)	2;11	3;4	4;3	4;6	541	1.757	1.548	6.944
D24-RAG	(m/L)	3;0	3;4	4;4	4;6	332	1.618	1.219	6.346
D25-PHG	(m/L)	3;0	3;4	4;4	4;6	488	2.161	1.498	7.556
D26-JUB	(m/L)	3;3	3;6	4;5	4;7	612	2.108	1.107	7.404
Summe LSES						7.801	29.901	21.282	110.469
Gesamt (L+HSES)						16.012	57.725	45.386	217.262

Anhang 12 Zusammenfassung

Die vorliegende Dissertation ist eine systematische korpuslinguistische Untersuchung der Kindersprache in Interaktion mit dem elterlichen Input in Bezug auf Derivationen im Vergleich zu Komposita. Insbesondere liegt ein Fokus auf Ableitungsmustern, die in der frühen Kindersprache eine tragende Rolle spielen. Überdies wird der Zusammenhang verschiedener Einflussfaktoren auf die morphologische Entwicklung im Bereich der Wortbildung bei Kindern von 3 bis 5 Jahren mit unterschiedlichem sozioökonomischen Hintergrund untersucht. Es werden einerseits longitudinale Spontansprachdaten von drei Kindern, sowie andererseits Paneldaten von 29 Kindern und ihren Hauptbezugspersonen umfassend analysiert. Alle Kinder wachsen in Wien auf und erwerben österreichisches Standarddeutsch als ihre Erstsprache. Die Untersuchung fasst die Ergebnisse aus anderen Studien, die die gleichen Kinder zu einem früheren Zeitpunkt abbilden, zusammen, stellt sie jedoch ausführlicher dar und erweitert sie um die Daten bis 5 Jahre. Das Mini-Paradigma-Kriterium zur Feststellung potenzieller Produktivität in der Kindersprache wird ebenfalls auf weitere Derivationsmuster angewandt und erweitert.

Zusätzlich wurde im Rahmen der Dissertation ein Experiment zur Elizitation nominaler Ableitungen entwickelt und mit 16 Kindern der Paneldaten im Alter von durchschnittlich 8 Jahren durchgeführt, um die weitere morphologische Entwicklung dieser Kinder im Bereich deverbaler nominaler Derivationen festzustellen.

Die wesentlichsten derivationellen Muster sind im Alter von durchschnittlich 3 Jahren in der Kindersprache vorhanden, wobei sich der Wortschatz der Kinder und Hauptbezugspersonen mit höherem sozioökonomischen Status aufgrund höherer Lemmafrequenzen vielfältiger zeigt.

Die Ergebnisse der Spontansprache und des Tests wurden den Faktoren Inputmenge, Büchervorlesen, Hörbücher und Radio, sowie weiterem Medienkonsum (YouTube, Internet, Smartphone, Fernsehen) gegenübergestellt, mit dem Resultat, dass ein günstiger Effekt zwischen der Menge des Büchervorlesens und der Entwicklung im Bereich der Derivationsmorphologie

beobachtet werden kann, während übermäßiger Medienkonsum nachteilige Auswirkungen mit sich bringt.

Anhang 13 Abstract

This PhD thesis is a systematic corpus linguistic study of child speech in interaction with child-directed speech with respect to derivations in comparison to compounding. In particular with a focus on derivation patterns, which play an essential role in early child speech. Furthermore, the relation of different factors influencing the morphological development in word formation in children with different socioeconomic backgrounds is investigated. The longitudinal spontaneous speech corpora of three children as well as panel data of 29 children acquiring Standard Austrian German and their main caretakers is analyzed up to almost 5 years of age. This investigation summarizes previous results with the same children at an earlier stage and extends them up to age 5. The mini-paradigm criterion for attesting potential productivity in child speech is also applied and extended to other derivational patterns.

In addition, an elicitation task for nominal derivations was developed and performed within this study. It was conducted with 16 children of the panel data at an average age of 8 years old to trace their further morphological development in the area of deverbal nominal derivations.

Around age 3 the most essential derivational patterns are present in child speech. Still, the vocabulary of the children and their main caretakers with higher socioeconomic status is more diverse given higher lemma frequencies.

The results of the spontaneous speech analyses and test data were contrasted with the variables of input quantity, read aloud, audio books and radio, as well as media consumption (YouTube, internet use, smartphone, television), showing that a positive effect can be observed between the amount of reading aloud and the morphological development in derivational morphology in child speech, while excessive media consumption shows adverse effects.