

Angela Anderka

Elterliches Sprachangebot und vorschulischer Spracherwerb

Eine empirische Analyse zu Zusammenhängen
und sozialen Disparitäten



Waxmann 2018
Münster • New York

Anhang

Anhang A: Stichprobe

Anhang B: Ergänzende Statistik

Anhang C: Transkripte

Anhang D: Kodierungen

Anhang A: Stichprobe

A 1. Vergleich der Stichprobe mit der Gesamtstichprobe

In den folgenden Tabellen sind die Kenndaten der Stichprobe und die Ergebnisse der *t*-Tests des Vergleichs der Stichprobe mit der Gesamtstichprobe zu den Messzeitpunkten drei, vier, fünf und sechs dargestellt. Dabei werden nur für die Vergleiche die Ergebnisse des *t*-Tests berichtet, für die sich statistisch bedeutende Differenzen auf dem 5-Prozent-Niveau ergeben.

Tabelle 54: Übersicht über demographische Kennwerte und Kompetenzmerkmale der Stichprobe und Ergebnisse des Vergleichs mit der Gesamtstichprobe zu Messzeitpunkt 3

3. MZP	<i>M</i>	<i>SD</i>	<i>Min-Max</i>	Differenzen zu BiKS-3–10 Gesamtstichprobe (<i>N</i> = 547) $\alpha = 0,05$
Alter in Monaten; <i>N</i> = 63	53.2	3.85	47-61	-2.58; $t(576) = 4.41$; s.; $d = 0.60$
Anteil Mädchen	49.2%			ns
HISEI; <i>N</i> = 56	54.48	16.65	25-88	ns
Wortschatz (PPVT zum 3. MZP); <i>N</i> = 64	53.44	18.55	45-139	ns
Grammatik (TROG zum 3. MZP); <i>N</i> = 64	29.75	6.67	16-43	ns

Anmerkungen: MZP = Messzeitpunkt; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; deutsche Forschungsversion: Rossbach, Tietze & Weinert, 2005); TROG-D = Test zur Überprüfung des Grammatikverständnisses (Fox, 2006); s = signifikant; ns = nicht signifikant, d = Effektstärke

Tabelle 55: Übersicht über demographische Kennwerte und Kompetenzmerkmale der Stichprobe und Ergebnisse des Vergleichs mit der Gesamtstichprobe zu Messzeitpunkt 4

4. MZP	<i>M</i>	<i>SD</i>	<i>Min-Max</i>	Differenzen zu BiKS-3–10 Gesamtstichprobe (<i>N</i> = 547) . α = 0,05
Alter in Monaten; <i>N</i> = 62	59.53	3.98	53-68	ns
Grammatik (TROG zum 4. MZP); <i>N</i> = 65	33.74	5.88	18-24	ns

Anmerkungen: MZP = Messzeitpunkt; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; deutsche Forschungsversion: Rossbach, Tietze & Weinert, 2005); TROG-D = Test zur Überprüfung des Grammatikverständnisses (Fox, 2006); s = signifikant; ns = nicht signifikant, d = Effektstärke

Tabelle 56: Übersicht über demographische Kennwerte und Kompetenzmerkmale der Stichprobe und Ergebnisse des Vergleichs mit der Gesamtstichprobe zu Messzeitpunkt 5

5. MZP	<i>M</i>	<i>SD</i>	<i>Min-Max</i>	Differenzen zu BiKS-3–10 Gesamtstichprobe (<i>N</i> = 547) . α = 0,05
Alter in Monaten; <i>N</i> = 60	65.50	3.82	59-73	-1.96; $t(509) = 3.42$; s; $d = 0.47$
HISEI; <i>N</i> = 51	53.27	17.00	25-88	ns
Wortschatz (PPVT zum 5. MZP); <i>N</i> = 61	78.90	18.55	45-139	ns
Grammatik (TROG zum 5. MZP); <i>N</i> = 61	36.43	4.74	23-44	ns

Anmerkungen: MZP = Messzeitpunkt; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; deutsche Forschungsversion: Rossbach, Tietze & Weinert, 2005); TROG-D = Test zur Überprüfung des Grammatikverständnisses (Fox, 2006); s = signifikant; ns = nicht signifikant, d = Effektstärke

Tabelle 57: Übersicht über demographische Kennwerte und Kompetenzmerkmale der Stichprobe und Ergebnisse des Vergleichs mit der Gesamtstichprobe zu Messzeitpunkt 6

6. MZP	<i>M</i>	<i>SD</i>	<i>Min-Max</i>	Differenzen zu BiKS-3–10 Gesamtstichprobe (<i>N</i> = 547) . α = 0,05
Alter in Monaten; <i>N</i> = 58	70.83	3.65	65-78	-2.11; $t(487) = 3.58$; <i>s.</i> ; $d = 0.50$
Grammatik (TROG zum 6. MZP); <i>N</i> = 58	36.90	4.96	25-45	ns

Anmerkungen: MZP = Messzeitpunkt; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; deutsche Forschungsversion: Rossbach, Tietze & Weinert, 2005); TROG-D = Test zur Überprüfung des Grammatikverständnisses (Fox, 2006); s = signifikant; ns = nicht signifikant, d = Effektstärke

Zu den vier Messzeitpunkten ergeben sich nur für das Alter signifikante Differenzen zwischen der Stichprobe der vorliegenden Arbeit und der BiKS-3–10 Gesamtstichprobe. Zum dritten Messzeitpunkt sind die Kinder in der Teilstichprobe durchschnittlich 2.6 Monate jünger sind als die Kinder der Gesamtstichprobe ($t(576) = 4.47$; $p < .001$; $d = .60$). Zum darauf folgenden vierten Messzeitpunkt sind die Differenzen deutlich geringer (0.23 Monate jünger als die Gesamtstichprobe) und nicht signifikant. Zum fünften und sechsten Messzeitpunkt ist die Substichprobe der vorliegenden Arbeit wiederum gut zwei Monate jünger als die Gesamtstichprobe (MZP 5: $t(509) = 3.42$, $p < .001$, $d = .47$; MZP 6: $t(487) = 3.58$, $p < .001$, $d = .50$). Die Unterschiede liegen mit den Effektstärken $d = .48$ bis $d = .60$ im mittleren Bereich.

Tabelle 58: Übersicht über Kompetenzmerkmale der Kinder mit einem Elternteil mit Migrationshintergrund und der Kinder ohne Migrationshintergrund zu den drei Messzeitpunkten

	TROG (<i>M/ SD</i>)		PPVT (<i>M/ SD</i>)	
	Kinder mit einem Elternteil mit Migrationshintergrund	Kinder ohne Migrationshintergrund	Kinder mit einem Elternteil mit Migrationshintergrund	Kinder ohne Migrationshintergrund
3. MZP	23.50/ 2.1 (<i>N</i> = 2)	29.95/ 6.7 (<i>N</i> = 62)	54.00/ -- (<i>N</i> = 1)	53.43/ 17.6 (<i>N</i> = 61)
4. MZP	35.00/ 1.4 (<i>N</i> = 2)	33.70/ 6.0 (<i>N</i> = 63)		
5. MZP	35.50/ 5.0 (<i>N</i> = 2)	36.96/ 4.8 (<i>N</i> = 59)	71.50/ 14.9 (<i>N</i> = 2)	19.16/ 18.7 (<i>N</i> = 57)

Anmerkungen: MZP = Messzeitpunkt; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; deutsche Forschungsversion: Rossbach, Tietze & Weinert, 2005); TROG-D = Test zur Überprüfung des Grammatikverständnisses (Fox, 2006)

Anhang B. Ergänzende Statistik

- B 1. Vergleich der Stabilitäten interindividueller Unterschiede
- B 2. Vergleich der Zusammenhänge mit dem sozialen Hintergrund zwischen den beiden Interaktionssituationen
- B 3. Vergleich der verschiedenen Indikatoren des sozialen Hintergrundes
- B 4. Vergleich der Eltern mit höherem und mit niedrigerem sozialen Status
- B 5. Vergleich der Kinder mit höherem und mit niedrigerem sozialen Hintergrund (HISEI) (sprachliche Kompetenztests)
- B 6. Regressionsmodelle

B 1. Vergleich der Stabilitäten interindividueller Unterschiede

Tabelle 59: Überprüfung der Stabilitäten interindividueller Unterschiede in der Interaktionssituation Bilderbuch auf statistisch bedeutsame Differenzen zwischen den verschiedenen zeitlichen Abständen

	3.-4. MZP		4.-5. MZP		3.-5. MZP	
	<i>N</i>	<i>r</i>	<i>N</i> und Ergebnis des Signifikanztests	<i>r</i>	<i>N</i> und Ergebnis des Signifikanztests	<i>r</i>
Dauer der Interaktion ¹	59 (58)	.71** (.64**)	53 (52) z = 0.251	.68** (.61**)	50 (49) z = 0.00	.69** (.64**)
Menge Elternsprache (Wörter)	60	.76**	53 z = 1.644*	.59**	51 z = 2.001*	.54**
Vollständige Sätze	60	.74**	53 z = 1.931*	.52**	51 z = 1.979*	.51**
Anteil der vollständigen Sätze	60	.52**	53 z = 1.434 ⁺	.29*	51 z = 1.638 ⁺	.25
MLU	59	.65**	50 z = 0.336	.61**	49 z = -0.993	.75**
Komplexe Syntax	59	.60**	50 z = 0.378	.55**	49 z = 1.817*	.32**
Anteil komplexe Syntax	59	.26 ⁺	50 z = 0.108	.24 ⁺	49 z = 1.237	-.02
Anzahl Nominalphrasen	60	.76**	53 z = 1.722*	.58**	51 z = 2.073*	.53**
Anzahl NP pro Satz	59	.50**	50 z = 0.872	.36**	49 z = -0.42	.56**
Lexikalische Vielfalt (Types)	60	.76**	53 z = 1.05	.66**	51 z = 1.301 ⁺	.63**
TTR	59	.44**	51 z = -1.044	.59**	50 z = -1.621 ⁺	.66**

Anmerkung: MZP= Messzeitpunkt; ⁺p < .10, *p < .05, **p < .01; 3.-4. MZP: bivariate Korrelationen zwischen Wert des 3. und 4. Messzeitpunktes; 4.-5. MZP: bivariate Korrelationen zwischen Wert des 4. und 5. Messzeitpunktes; 3.-5. MZP: bivariate Korrelationen zwischen Wert des 3. und 5. Messzeitpunktes

B 2. Vergleich der Zusammenhänge mit dem sozialen Hintergrund zwischen den beiden Interaktionssituationen

Tabelle 60: Ergebnisse der Signifikanztests der Zusammenhänge mit dem sozialen Hintergrund zwischen den beiden Interaktionssituationen

	3. MZP		4. MZP		5. MZP	
	<i>z-Wert</i>	<i>p</i>	<i>z-Wert</i>	<i>p</i>	<i>z-Wert</i>	<i>p</i>
Dauer ¹	0.45	.33	1.10	.14	0.81	.21
Menge Elternsprache (Wörter)	-0.40	.35	0.67	.25	0.37	.36
Vollständige Sätze	-0.17	.43	0.56	.29	0.84	.20
Anteil vollständige Sätze	0.55	.29	0.70	.24	-1.01	.16
MLU	-2.37	.009	0.27	.39	-1.05	.15
Komplexe Sätze	-0.35	.36	0.74	.23	-1.06	.15
Nominalphrasen	-0.17	.43	0.67	.25	0.37	.23
Lexikalische Vielfalt (Types)	-0.49	.31	1.02	.15	0.22	.41

Anmerkungen: ¹= Dauer der Interaktion ohne Extremwert; MLU= durchschnittliche Länge der vollständigen Sätze; MZP= Messzeitpunkt

B 3. Vergleich der verschiedenen Indikatoren des sozialen Hintergrundes

Tabelle 61: Bivariate Korrelationen ausgewählter Indikatoren der Elternsprache mit der schulischen Bildung der Hauptbetreuungsperson

	N	3. MZP	N	4. MZP	N	5. MZP
Dauer der Interaktion ¹	48	.26	48	.31*	48	.35*
Menge Elternsprache (Wörter)	50	.18	50	.22	50	.31*
Vollständige Sätze	50	.21	50	.24	50	.36**
Anteil der vollständigen Sätze	50	.06	50	.17	50	-.13
MLU	47	-.25	47	-.15	47	-.12
Komplexe Syntax	47	.13	47	.11	47	.13
Lexikalische Vielfalt (Types)	50	.26	50	.28*	50	.35*

Anmerkungen: ¹ = Dauer der Interaktion ohne Extremwert; MLU= durchschnittliche Länge der vollständigen Sätze; MZP= Messzeitpunkt; ⁺ $p < .10$, * $p < .05$, ** $p < .01$

Tabelle 62: Bivariate Korrelationen ausgewählter Indikatoren der Elternsprache mit der beruflichen Bildung der Hauptbetreuungsperson

	N	3. MZP	N	4. MZP	N	5- MZP
Dauer der Interaktion ¹	48	.02	48	-.05	48	.02
Menge Elternsprache (Wörter)	50	-.06	50	-.06	50	-.04
Vollständige Sätze	50	-.05	50	-.06	50	.01
Anteil der vollständigen Sätze	50	-.32*	50	-.33*	50	-.18
MLU	47	-.34*	47	-.10	47	-.10
Komplexe Syntax	47	.06	47	.05	47	-.01
Lexikalische Vielfalt (Types)	50	.01	50	-.06	50	.00

Anmerkungen: ¹ = Dauer der Interaktion ohne Extremwert; MLU= durchschnittliche Länge der vollständigen Sätze; MZP= Messzeitpunkt; ⁺ $p < .10$, * $p < .05$, ** $p < .01$

Tabelle 63: Ergebnisse der einfaktoriellen Varianzanalyse mit Messwiederholung für den Interaktionseffekt zwischen Messzeitpunkt und schulischer Bildung und für den Effekt der schulischen Bildung als unabhängiger Faktor

	<i>Interaktionseffekt</i>					<i>Effekt schulische Bildung</i>			
	<i>N</i>	<i>F-Wert</i>	<i>df</i>	<i>p</i>	η_p^2	<i>F-Wert</i>	<i>df</i>	<i>p</i>	η_p^2
Dauer der Interaktion ¹	48	0.27	4, 90	.90	.01	3.50 ¹	2, 45	.04	.14
Menge Elternsprache (Wörter)	50	0.98	4, 94	.42	.04	1.91 ²	2, 47	.16	.08
Vollständige Sätze	50	0.64	4, 94	.63	.03	2.44 ²	2, 47	.10	.09
Anteil der vollständigen Sätze	50	2.06	4, 94	.09	.08	0.87 ²	2, 47	.42	.04
MLU	47	0.93	4, 88	.45	.04	1.11 ²	2, 44	.34	.05
Komplexe Syntax	47	0.12	4, 88	.97	.01	0.76 ²	2, 44	.47	.03
Lexikalische Vielfalt (Types)	50	0.74	4, 94	.57	.03	3.00 ²	2, 47	.06	.11

Anmerkungen: Ergebnisse Post-Hoc-Vergleiche: ¹ = nur Hauptschulabschluss- Abitur signifikante Differenzen ($p < .05$); ² = keine signifikante Differenzen

Tabelle 64: Ergebnisse der einfaktoriellen Varianzanalyse mit Messwiederholung für den Interaktionseffekt zwischen Messzeitpunkt und beruflicher Bildung und für den Effekt der beruflichen Bildung als unabhängiger Faktor

	<i>Interaktionseffekt</i>					<i>Effekt berufliche Bildung</i>			
	<i>N</i>	<i>F-Wert</i>	<i>df</i>	<i>p</i>	η_p^2	<i>F-Wert</i>	<i>df</i>	<i>p</i>	η_p^2
Dauer der Interaktion ¹	48	0.77	2, 92	.68	.01	1.65	1, 46	.21	.04
Menge Elternsprache (Wörter)	50	0.72	2, 96	.49	.02	0.65	1, 48	.43	.01
Vollständige Sätze	50	0.62	2, 96	.54	.01	0.97	1, 48	.33	.02
Anteil der vollständigen Sätze	50	0.07	2, 96	.93	.00	0.69	1, 48	.41	.01
MLU	47	1.60	2, 90	.21	.03	2.81	1, 45	.10	.06
Komplexe Syntax	47	0.11	2, 90	.90	.00	0.00	1, 45	.98	.00
Lexikalische Vielfalt (Types)	50	1.25	2, 96	.29	.03	1.24	1, 48	.27	.03

*Anmerkungen: η_p^2 = Effektstärke; ¹ = Dauer der Interaktion ohne Extremwert;
MLU = durchschnittliche Länge der vollständigen Sätze*

B 4. Vergleich der Eltern mit höherem und mit niedrigerem sozialen Status

Tabelle 65: Mittelwerte und Testung der Differenz auf Signifikanz- 2 HISEI-Gruppen zu allen 3 Messzeitpunkten

	<i>MZP</i>	<i>HISEI</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t-Wert</i>
Dauer der Interaktion ¹	3	1	25	5.53	2.16	
		2	23	7.17	2.53	2.42*
	4	1	25	5.67	2.09	
		2	24	7.28	2.71	2.33*
	5	1	25	4.65	1.57	
		2	24	5.58	1.49	2.12*
Menge Elternsprache (Wörter)	3	1	26	544	290	
		2	24	733	321	2.18*
	4	1	26	494	268	
		2	24	722	362	2.55*
	5	1	26	422	194	
		2	24	542	182	2.26*
Vollständige Sätze	3	1	26	69	40	
		2	24	96	43	2.30*
	4	1	26	62	31	
		2	24	93	48	2.69*
	5	1	26	52	22	
		2	24	67	21	2.56*
MLU	3	1	26	6.15	0.59	
		2	24	6.03	0.69	0.65
	4	1	25	6.09	0.49	
		2	24	6.27	0.75	0.28
	5	1	24	6.16	0.63	
		2	24	6.22	0.84	0.61

<i>Fortsetzung</i> Tabelle 65	<i>MZP</i>	<i>HISEI</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t-Wert</i>
Komplexe Syntax	3	1	26	8.31	6.45	
		2	24	13.63	7.71	2.65*
	4	1	25	8.40	6.12	
		2	24	14.63	11.16	2.43*
	5	1	24	5.79	4.05	
		2	24	7.58	4.75	1.41
Nominalphrasen	3	1	26	168.15	94.22	
		2	24	229.12	100.83	2.73**
	4	1	26	154.38	83.13	
		2	24	226.42	113.99	2.62*
	5	1	26	130.62	58.54	
		2	24	166.46	57.34	2.27*
Lexikalische Vielfalt (Types)	3	1	26	158.58	48.93	
		2	24	195.96	55.91	2.52*
	4	1	26	176.46	67.12	
		2	24	237.08	76.12	2.99**
	5	1	26	147.35	56.08	
		2	24	178.71	42.98	2.21*

*Anmerkungen: ⁺p < .10, *p < .05, **p < .01; ^l=Dauer der Interaktion ohne Extremwert; MLU= durchschnittliche Länge der vollständigen Sätze; MZP= Messzeitpunkt*

Tabelle 66: Mittelwerte und Testung der Differenz auf Signifikanz- 2 HISEI-Gruppen für den über Messzeitpunkt 3 und 4 gemittelten Wert der erhobenen Indikatoren

	<i>HISEI</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t-Wert</i>
Dauer der Interaktion ¹	1	30	5.75	1.84	
	2	26	7.18	2.14	2.70**
Menge Elternsprache (Wörter)	1	30	495	196	
	2	26	705	289	3.13** ^{VN}
Vollständige Sätze	1	30	62	25	
	2	26	91	38	3.24** ^{VN}
MLU	1	30	6.04	0.48	
	2	26	6.13	0.66	0.57
Komplexe Syntax	1	30	7.33	4.17	
	2	26	13.60	7.65	3.72** ^{VN}
Nominalphrasen	1	30	152	61.30	
	2	26	220	89.77	3.28** ^{VN}
Lexikalische Vielfalt (Types)	1	30	165	50	
	2	26	215	56	3.54**

Anmerkungen: ⁺ $p < .10$, $*p < .05$, $**p < .01$; ¹=Dauer der Interaktion ohne Extremwert; MLU= durchschnittliche Länge der vollständigen Sätze, ^{VN}=T-Werte bei nicht gegebener Varianz-Homogenität (Levene-Test der Varianzgleichheit)

B 5. Vergleich der Kinder mit höherem und mit niedrigerem sozialen Hintergrund (HISEI) (sprachliche Kompetenztests)

Tabelle 67: Mittelwerte und Testung der Differenz auf Signifikanz- 2 HISEI-Gruppen

	<i>MZP</i>	<i>HISEI</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t-Wert</i>
TROG	3	1	30	28.27	6.83	
		2	26	31.88	5.97	2.09*
	4	1	30	31.93	6.06	
		2	26	35.89	5.06	2.63*
	5	1	30	35.07	5.13	
		2	26	37.77	4,09	2.16*
6	1	30	35.07	5.32		
	2	26	39.04	3.72	3.27** ^{VN}	
PPVT	3	1	25	52.80	17.39	
		2	22	57.68	13.77	1.06
	5	1	25	74.72	16.40	
		2	22	83.59	15.33	1.91 ⁺

Anmerkungen: ⁺ $p < .10$, * $p < .05$, ** $p < .01$; ^{VN} = *T-Werte bei nicht gegebener Varianz-Homogenität (Levene-Test der Varianzgleichheit)*

B 6. Regressionsmodelle

Tabelle 68: Ergebnisse des Regressionsmodells, in dem alle Variablen der Elternsprache enthalten sind -Varianzaufklärung der Veränderung grammatischer Kompetenzen der Kinder zu Messzeitpunkt 4

	R^2 alle Variablen: .44		Kollinearitätsstatistik	
	β	t	Toleranz	VIF
MW 3./4. MZP Menge Elternsprache	-0.78	-0.71	0.01	98.61
MW 3./4. MZP Vollständige Sätze	-0.37	-0.43	0.02	60.76
MW 3./4. MZP Lexikalische Vielfalt	-0.16	-0.39	0.08	12.85
MW 3./4. MZP Nominalphrasen	1.29	1.08	0.01	116.74
MW 3./4. MZP Komplexe Syntax	0.15	0.64	0.22	4.47
MW 3./4. MZP Dauer der Interaktion ¹	0.15	0.78	0.31	3.19

Anmerkungen: ⁺ $p < .10$, * $p < .05$, ** $p < .01$; ¹ = Dauer der Interaktion ohne Extremwert; MZP = Messzeitpunkt; VIF = Varianzinflations-Faktor; $N = 56$; Kontrollvariablen: Alter der Kinder zu Messzeitpunkt 3 und 4 und Werte grammatischer Kompetenzen zu Messzeitpunkt 3

Tabelle 69: Ergebnisse des Regressionsmodells, in dem der Faktor Menge enthalten ist -Varianzaufklärung der Veränderung grammatischer Kompetenzen der Kinder zu Messzeitpunkt 4

	R^2 alle Variablen: .42		Kollinearitätsstatistik	
	β	t	Toleranz	VIF
MW 3./4. MZP Faktor Menge	-0.07	-0.24	0.16	6.46
MW 3./4. MZP Komplexe Syntax	0.17	0.75	0.23	4.42
MW 3./4. MZP Dauer der Interaktion ¹	0.19	0.99	0.32	3.09

Anmerkungen: ⁺ $p < .10$, * $p < .05$, ** $p < .01$; Anmerkungen: ¹ = Dauer der Interaktion ohne Extremwert; MZP = Messzeitpunkt; VIF = Varianzinflations-Faktor; $N = 56$; Kontrollvariablen: Alter der Kinder zu Messzeitpunkt 3 und 4 und Werte grammatischer Kompetenzen zu Messzeitpunkt 3

Tabelle 70: Ergebnisse des Regressionsmodells für die verbliebenen drei Variablen der Elternsprache-Varianzaufklärung der Veränderung grammatischer Kompetenzen der Kinder zu Messzeitpunkt 4

	<i>R² alle Variablen: .42</i>		<i>Kollinearitätsstatistik</i>	
	β	<i>t</i>	<i>Toleranz</i>	<i>VIF</i>
MW 3./4. MZP Lexikalische Vielfalt	-0.10	-0.41	0.21	4.74
MW 3./4. MZP Komplexe Syntax	0.19	0.90	0.28	3.62
MW 3./4. MZP Dauer der Interaktion ¹	0.20	1.10	0.37	2.71

*Anmerkungen: ⁺p < .10, *p < .05, **p < .01; Anmerkungen: ¹ = Dauer der Interaktion ohne Extremwert; MZP = Messzeitpunkt; VIF = Varianzinflations-Faktor; N = 56; Kontrollvariablen: Alter der Kinder zu Messzeitpunkt 3 und 4 und Werte grammatischer Kompetenzen zu Messzeitpunkt 3*

Anhang C: Tanskripte

C 1. Beschreibung der transkribierten Dateien

C 2. Einarbeitung der studentischen Hilfskräfte

C 3. Segmentierung der gesprochenen Sprache

C 4. Transkriptionsvereinbarungen

C 5. Übereinstimmung der Tanskripte

C 1. Beschreibung der transkribierten Dateien

Von den Eltern wurden alle vorliegenden Videoaufnahmen von zwei Interaktionssituationen (Bilderbuch und Konstruktionsspiel) der drei Erhebungswellen transkribiert. Diese sind im Schnitt sieben Minuten lang. Insgesamt liegen 359 Transkripte vor (ca. 42 Stunden Videomaterial).

Bei den Eltern- Kind- Interaktionen stand das Aufnahmegerät in der Nähe des Eltern-Kind-Paares. Nicht immer wurde darauf geachtet, dass störende Geräusche vermieden wurden. So sind manche Aufnahmen schwerer verständlich, da z.B. Geschwisterkinder im Hintergrund reden oder Geräusche durch das geöffnete Fenster mit aufgenommen wurden.

C 2. Einarbeitung der studentischen Hilfskräfte

Die studentischen Hilfskräfte wurden anhand bereits vorliegender Transkripte geschult. Sie bekamen gemeinsam mit dem Transkript die Videodatei und arbeiteten sich so in das Anfertigen der Transkripte ein. Die Hilfskräfte wurden einzeln über die Transkriptionsvereinbarungen instruiert, da sie zu unterschiedlichen Zeitpunkten begonnen hatten. Sie bekamen nach jeweils einigen Transkripten Rückmeldung über die Qualität ihrer Transkripte. Die zu verändernden Aspekte wurden nach kurzer Zeit nochmals überprüft.

C 3. Segmentierung der gesprochenen Sprache

Die Hilfskräfte erhielten die Anweisung, die gehörte Sprache stets in kleinstmögliche, inhaltlich und grammatisch sinnvolle und korrekte Einheiten zu untergliedern.

Grammatische Formen, die zusammen gehören, werden jedoch nicht getrennt (z.B. ein Satzgefüge mit Haupt- und Nebensatz- „*Sie kaufen ein, weil sie einen Kuchen backen wollen*“- oder Satzteile, die durch ein gemeinsames Subjekt, das auf beide Verben bezogen ist, verbunden sind- „*Mama telefoniert und schmeißt den Schlüssel hinter das Kissen.*“)

In der gesprochenen Sprache gibt es neben den im schriftsprachlichen Sinne vollständigen Sätzen aber auch sprachliche Formen, die nur Interjektionen („*Hurra!*“, „*Oh je!*“, „*Mhm.*“) oder einzelne Phrasen (z.B. „*Unter dem*

Tisch. ‘’) enthalten. Diese werden ebenfalls als einzelne Analyseeinheiten (= kodierte Spracheinheiten, s. Kapitel „Methodik“) segmentiert.

Stehen Wörter scheinbar am Beginn eines Satzes, ohne die Funktion eines für den Inhalt oder die Struktur des Satz sinnvollen und wichtigen Satzgliedes, sondern eher mit der Absicht, die Aufmerksamkeit auf den folgenden Satz zu lenken, so werden diese Wörter vom darauf folgenden Satz als eigenes Segment abgetrennt (z.B. *genau, so, also, na, ja, nein*).

Hingegen werden am Satzende stehende Wörter, die den Inhalt des Satzes bekräftigen, im Segment belassen (z.B. *ne, ja, gell, oder*):

	0	1	2
X [v]	Genau.	Jetzt kommen sie, ne?	

Ebenso werden Ellipsen, die „[...]“ durch Beachtung des vorhergehenden Kontextes interpretierbare und sinnvolle separierbare Einheiten [...]“ (Gumpey & Berenz, 1993, 95) darstellen, als einzelne Segmente erfasst:

	0	1	2
X [v]	Eier sind keine mehr da.	(()) Ham sie vergessen wohl beim Einkaufen, ne?	

Wird ein Satz abgebrochen, und eine eventuelle Fortsetzung hat keinen Bezug zum abgebrochenen Satz, so wird dies ebenfalls als einzelnes Segment unterteilt:

	0	1	2
K [v]	Wenn sie dann//.	Später sucht sie den Schlüssel.	

Wird hingegen ein begonnener Satz repariert, d.h. einzelne Wörter oder Satzglieder werden ausgetauscht, und der Satzanfang hat einen grammatischen und inhaltlichen Bezug zur Fortsetzung des Satzes, so bleiben alle Wörter bzw. Satzglieder in einem Segment:

	0	1
X [v]	Als dann die/ der Maxi weiter puzzelt.	

Ebenso werden bei Sätzen, in denen einzelne Wörter oder Wortgruppen einmal oder mehrmals wiederholt werden (mit der Absicht der Intensitätssteigerung oder aufgrund eines sprecherseitigen Verbalisierungsproblems), alle Wörter in einem Segment belassen:

	0	1
X [v]	Seit vielen, vielen Jahren lebt der Igel schon bei Hempels im Garten.	

	0	1
X [v]	Dann, dann, dann müssen wir noch einkaufen.	

In der Variable „Satzvollständigkeit“ wird beim Kodieren die unterschiedliche Beschaffenheit der einzelnen Segmente notiert und entweder als Interjektion, satzwertige Strukturen, Abbrüche, Ellipsen oder vollständige Sätze kategorisiert.

C 4. Transkriptionsvereinbarungen

Im Folgenden werden die Transkriptionsvereinbarungen im Einzelnen beschrieben.

Tabelle 71: Transkriptionsvereinbarungen

	Erläuterung	Beispiele																				
Gliederung der gesprochenen Sprache	Erfolgt auf Satzebene. (→ ein Satz = ein Segment in der jeweiligen Speaker-Tier.) Orientierung an der Schriftsprache: In der mündlichen Sprache sind Grenzen häufig nicht klar → Segmentgrenzen werden bestimmt an Formen, wie sie in der Schriftsprache zu finden sind, z.B. Hauptsätze und Nebensätze. → Typische Formen der mündlichen Sprache wie Interjektionen, Lautmalereien oder „einleitende“ Satzglieder (wie „genau“, „ja“) werden abgetrennt, wenn sie kein erkennbar „sinnvolles“ Satzglied des Satzes sind und als eigenes Segment geschrieben.(s.u.)	<table border="1"> <tr> <td></td> <td>18 [44.3]</td> <td>19 [45.3]</td> <td>20 [47.3]</td> </tr> <tr> <td>PRE [v]</td> <td></td> <td></td> <td>I mean y</td> </tr> <tr> <td>ELK [v]</td> <td>Ehem.</td> <td></td> <td></td> </tr> <tr> <td>[m]</td> <td></td> <td>((laughter))</td> <td></td> </tr> <tr> <td>PPF n-1</td> <td></td> <td></td> <td></td> </tr> </table> <p><i>I</i>Was macht die Mama? <i>I</i>Und was macht der Maxi? <i>I</i></p> <p><i>I</i>Also. <i>I</i>Hm. <i>I</i></p>		18 [44.3]	19 [45.3]	20 [47.3]	PRE [v]			I mean y	ELK [v]	Ehem.			[m]		((laughter))		PPF n-1			
		18 [44.3]	19 [45.3]	20 [47.3]																		
	PRE [v]			I mean y																		
	ELK [v]	Ehem.																				
	[m]		((laughter))																			
	PPF n-1																					
2 Hauptsätze: ohne Zusammenhang und mit jeweils eigenem Subjekt und Prädikat → 2 Segmente	<i>Ich gehe einkaufen und du kannst aufräumen.</i> → <i>Ich gehe einkaufen. Und du kannst aufräumen.</i>																					
Haupt- und Nebensatz: stehen in Zusammenhang → keine Trennung	<i>Er hat gesagt, dass er sich darum kümmern will.</i>																					
Einzelne Lautmalerei: am Satzanfang/-ende → Trennung in 2 Segmente	<i>Mmh, wer möchte heute vorlesen?</i> → <i>Mhm. Wer möchte heute vorlesen?</i>																					
Einzelne Lautmalerei: im Satz ○ keine Trennung	<i>Wer weiß, ähh, was da oben hinkommt?</i>																					
Einwortäußerungen: am Satzanfang/-ende, die kein erkennbar „sinnvolles“ Satzglied im Satz darstellen → Trennung in 2 Segmente	<i>Genau, dann fangen wir mit Schreiben an</i> → <i>Genau. Dann fangen wir mit Schreiben an</i> ABER Vorsicht: <i>. Genau das habe ich von euch erwartet.</i> Hier gehört „genau“ als Wort in einem Satzglied in den Satz mit hinein → nicht trennen!																					
Satzzeichen	Entsprechend dem Satzmodus. Nach jedem Äußerungsendzeichen (Satzpunkt) muss noch ein Leerzeichen gesetzt werden	Frage → ? Aussage → . Deutliche Aufforderung → !																				

Fortsetzung	Erläuterung	Beispiele
Ankündigungen	Eine Äußerung erfüllt die Funktion auf etwas Folgendes hinzuweisen. → Trennung vom folgenden Segment	Doppelpunkt am Ende der hinweisenden Äußerung z.B. „Mama sagt: I Wir haben die Eier vergessen. I “
Pausen	Pausen werden gekennzeichnet, aber nicht in ihrer Länge unterschieden.	(()) (Leerzeichen muss dazwischen gesetzt werden) Am Anfang eines Segments, wenn vor der Äußerung eine Pause war: „(()) Maxi puzzelt immer noch.“ Im Satz, wenn zwischen den Wörtern Pause war: „Die Mutter überlegt,(()) wo sie den Schlüssel hingelegt hat.“
Reparatur	Es liegt ein vollständiger Satz vor, der auf der sprachlichen Oberfläche repariert wird., d.h. der ursprünglich begonnene Satz wird in ihrer reparierten Version weitergeführt und abgeschlossen.	Schrägstrich ohne Leerzeichen im Anschluss an das Wort: Wort/LZ neues Wort z.B. „Als dann die/ der Maxi weiter puzzelt.“
Abbruch	Abbruch liegt vor, wenn ein begonnener Satz in seiner Verbalisierung unabgeschlossen bleibt und die Kommunikation anschließend von demselben oder einem anderen Sprecher mit einem neuen Satz fortgeführt wird.	Der vollständige Abbruch eines Satzes wird mit : „/!“ gekennzeichnet. z.B. „Wenn sie dann//.“ I „Später sucht sie den Schlüssel.“ I
Reihung	Wiederholung ganzer Wörter bzw. Wortgruppen oder ganzer Sätze (Funktion: Gewichtung, Intensitätssteigerung oder Aufzählung)	Gliederung durch Komma z.B. „Seit vielen, vielen, vielen Jahren.“
Repetition	Wiederholung identischer Wörter oder Wortgruppen, die aber eher ein sprecherseitiges Verbalisierungsproblem zum Ausdruck bringen.	Gliederung durch Komma z.B. „Nein, dann, dann, dann müssen wir noch einkaufen.“
Hochsprachliche Umgangssprache	Sehr stark von herkömmlicher Orthographie abweichende (v.a. dialektbedingt) Schreibweisen korrigieren Leichtere Abweichungen wie z.B. gebn statt geben oder Kugln statt Kugeln schafft das Programm → diese müssen nicht verändert werden	z.B. alloa → allein Guck ma → Guck mal aa → auch Was is' n da los? → Was ist denn da los Was denkst' n? → Was denkst denn

Fortsetzung	Erläuterung	Beispiele
verschlif-fene Äußerungen	Wenn einzelne Wörter (z.B.: „es“, „du“) verschliffen werden, müssen diese trotzdem ausgeschrieben werden.	Innerhalb eines Wortes: z.B. „ <i>stimmt's</i> → <i>stimmt es.</i> “ <i>Und die Mama will was aus' m Kühlschranks holn. → Und die Mama will was aus dem Kühlschrank holn</i> <i>Jetzt darfst'e anfangen. → Jetzt darfst du anfangen.</i>
fehlende Wörter (Ellipsen)	Ganz fehlende Wörter werden <u>nicht</u> dazugeschrieben!	z.B. „Is doch schon längst drin.“ (gemeint wäre wahrscheinlich: Es ist doch schon längst drin.)
Vorgelesenes	Vorgelesenes wird mit Anführungszeichen markiert Vorgelesenes wird bei der Kodierung der Äußerungen <u>nicht</u> gewertet!	z.B. <i>Schau mal! Ein Buch. „Die kleine Ente Nelly.“</i> und in der nonverbalen Spur wird liest vor vermerkt, (solange, wie die Mutter liest.)
frei Gesungenes	Frei Gesungenes wird ohne Anführungsstriche transkribiert	In der nonverbalen Spur wird singt vermerkt, (solange, wie die Mutter singt.)
Unverständliches	Zurückspulen und noch mal anhören! Wenn das nicht hilft: Manchmal bringt es was, es <u>erstmal</u> frei zu lassen und weiter zu machen und ganz am Ende noch mal die Stelle anzuhören, wenn man sich an die Redeweise der Mutter / des Kindes gewöhnt hat. Medium wechseln! (Lautsprecher oder andere Kopfhörer ausprobieren!) Jemand anderen hören lassen! Wenn alles nichts hilft, als unverständlich markieren!	z.B. „ <i>Maxi macht ein (Puzzle?)</i> .“
Kontrollhören	Ganz am Schluss der Transkription, (wenn alles transkribiert ist und man sich an die Redeweisen der Sprecher gewöhnt hat): Das ganze Transkript noch einmal laufen lassen: Anhören und mitlesen und gegebenenfalls ausbessern!	

C 5. Übereinstimmung der Transkripte

Segmentierung

Die Berechnungen der Übereinstimmungen der Segmentierung der Transkripte erfolgte für die Transkripte der Lehrersprache und der Elternsprache gemeinsam. Daran beteiligt waren Vannesa Röhrich und Marei Kotzerke, wissenschaftliche Mitarbeiterinnen im Teilprojekt 3 der Längsschnittstudie BiKS-3–10.²⁸

In der Literatur finden sich für diese Art von Transkripten, wie sie für die vorliegende Arbeit angefertigt wurden, keine Hinweise für die Berechnung der Übereinstimmung zwischen zwei oder mehr Transkribenden. Berechnungen wie sie bei phonetischen Transkripten durchgeführt werden, lassen sich nicht übertragen, da es bei ihnen um die Berechnung des Abstandes zwischen Transkribenden hinsichtlich von Konsonanten und Vokalen auf einer Abstandsmatrix geht (Vieregge, 1985). Als eine weitere Prozedur wird bei Van Borsel (1989) die Möglichkeit dargestellt, die Reliabilität aus dem Quotient der durchschnittlichen Anzahl der identisch transkribierten Segmente geteilt durch die Gesamtanzahl der Segmente zu berechnen. Diese Möglichkeit lässt sich gut auf den Vergleich der in einem Segment transkribierten Wörter übertragen und wurde für die vorliegenden Transkripte durchgeführt.

Die Übereinstimmung der Transkribenden hinsichtlich der Unterteilung der gehörten Sprache in einzelne Segmente wurde anhand – zufällig ausgewählter – drei Dateien der Lehrersprache und zwei Dateien der Elternsprache durchgeführt, die von allen Hilfskräften (meist 12 Personen) bearbeitet wurden. Diese enthielten jeweils im Schnitt knapp 160 Segmente, insgesamt waren es 790 Segmente. Alle Hilfskräfte bekamen dieselben Instruktionen. Abweichungen zwischen den Hilfskräften sind also auf eine uneindeutige Regelung oder ein ungenaues Arbeiten der jeweiligen Hilfskraft zurückzuführen

Die Berechnung der Übereinstimmungen fand dabei auf zwei Wegen statt: Zum einen wurden alle Transkribenden miteinander verglichen, d.h. jeder Transkribend wurde einzeln mit allen anderen verglichen, indem die Anzahl der übereinstimmenden Segmentierungen ins Verhältnis gesetzt wird zu allen von dem je-

28 Leitung Frau Prof. Dr. Sabine Weinert.

weiligen Pärchen vollzogenen Segmentierungen. Zum zweiten wurde die Segmentierung der einzelnen Hilfskräfte mit einer sogenannten „Masterdatei“ verglichen. Diese entstanden in Zusammenarbeit mit Vanessa Röhrich und Marei Kotzerke für jede der fünf Dateien. Es wurde dann beurteilt, ob das Segment der Masterdatei von der entsprechenden Hilfskraft in gleicher Weise segmentiert wurde oder ob nicht. In Tabelle 72 sind die Ergebnisse für Übereinstimmung der Hilfskräfte hinsichtlich ihrer Segmentierung für beide Berechnungs-Prozeduren zusammengefasst.

Tabelle 72: Ergebnisse der Überprüfung der Segmentierung der Sprache durch studentische Hilfskräfte (Angaben der Übereinstimmung in Prozent)

	Bilderbuch	Konstruktionsspiel	HSU	Deutsch	Mathematik	insgesamt
Übereinstimmung auf Grundlage der Vergleiche aller Hilfskräfte miteinander						
Anzahl der Hilfskräfte	9	12	12	12	12	
Durchschnittliche Übereinstimmung aller Hilfskräfte	83.30	91.00	85.61	8.13	88.03	87.31
Geringste Übereinstimmung	73.50	82.71	72.65	72.46	80.52	
Höchste Übereinstimmung	93.79	95.54	100.00	91.91	100.00	
Übereinstimmung auf der Grundlage des Vergleichs der einzelnen Hilfskraft mit der Masterdatei						
Anzahl der Hilfskräfte	10	12	12	12	12	
Durchschnittliche Übereinstimmung aller Hilfskräfte	80.63	90.17	86.98	87.67	90.66	87.22
Geringste Übereinstimmung	72.04	84.67	77.10	76.00	84.96	
Höchste Übereinstimmung	89.78	95.33	93.39	94.67	94.24	

In beiden Berechnungsprozeduren haben die Transkribenden eine Übereinstimmung von gut 87 Prozent untereinander bzw. mit der Masterdatei.

Transkribierte Wörter

Transkribenden unterscheiden sich nicht nur in Entscheidungen zur Unterteilung der gesprochenen Sprache in analysierbare Einheiten (s.o.), sondern nehmen auch die Wörter an sich subjektiv unterschiedlich wahr. So werden beispielsweise die Transkribenden umso korrekter das Gehörte wiedergeben je vertrauter sie mit dem jeweiligen Dialekt der Sprecher sind. Auch die Qualität der Wiedergabe der Audio- oder Videoaufnahme beeinflusst die Verständlichkeit des Gesprochenen. Dazu kommen noch die grundsätzlich unterschiedliche Wahrnehmungsfähigkeit und Verarbeitungsgeschwindigkeit, die sich auf die Qualität des Transkripts auswirken können.

Zur Überprüfung der unterschiedlichen Wahrnehmung der gehörten Sprache und ihrer schriftlichen Darstellung wurden insgesamt acht zufällig ausgewählte Transkripte der Elternsprache von jeweils zwei Personen in unterschiedlichen Paarungen doppelt transkribiert- insgesamt waren acht Hilfskräfte daran beteiligt. Die acht Transkripte enthielten insgesamt 3450 Wörter. Beim jeweiligen Vergleich der beiden Transkripte wurde für jede der beiden Hilfskräfte notiert, wenn sie ein Wort mehr aufgeschrieben hat (Fall 1). Außerdem wurde festgehalten, wenn Wörter unterschiedlich gehört wurden (Fall 2).

In den Segmenten, in denen sich die beiden Transkripte jeweils unterscheiden, betrifft es überwiegend nur eines der Wörter, die zusätzlich oder anders notiert wurden.

Wie aus Tabelle 73 ersichtlich wird, kommen gut sechs Prozent Wörter jeweils nur in einem der beiden Transkripte im Paarvergleich vor. Die Anzahl der Wörter, die von einer Person jeweils mehr gehört werden, verteilt sich dabei gleichmäßig auf die beiden beteiligten Transkribenden. Tendenziell gibt es einzelne Hilfskräfte, die im Vergleich zu den anderen mehr Wörter aufgeschrieben haben.

Tabelle 73: Ergebnis der Überprüfung der Übereinstimmung der transkribierten Wörter

	Anzahl absolut	durchschnittlicher Anteil an allen Wörtern (3450) der 8 Transkripte	min- max (einzelne Paar- vergleiche)
Wörter, die von einer der beiden Hilfskräfte mehr aufgeschrieben wurden (alle 8 Paarvergleiche)	217	6.29%	2.50-13.79%
Wörter, die von den beiden Hilfskräften unterschiedlich aufgeschrieben wurden (alle 8 Paarvergleiche)	160	4.64%	1.52-7.55%
Summen	377	10.93%	

Circa 4.5 Prozent der Wörter werden von den Transkribenden jeweils unterschiedlich aufgeschrieben. Diese unterschiedliche Wahrnehmung der gesprochenen Sprache wirkt sich auf den nächsten Schritt, die Kodierung der Merkmale der Elternsprache aus. So schlägt sich die unterschiedliche Anzahl an wahrgenommenen Wörtern in der Anzahl der Wörter pro Segment wieder, d.h. die Länge einzelner Segmente wird sich unterscheiden, je nachdem welches Transkript zugrunde liegt (betrifft ca. 60% der gefundenen Abweichungen).

In gut 15 Prozent der beobachteten Abweichungen verändert sich durch die unterschiedliche Notation auch die Kodierung der Satzvollständigkeit. So wird vom ersten Transkribenden beispielsweise „*Mit so nem Puzzle.*“ und vom zweiten „*Was ist denn mit dem Puzzle?*“ notiert. Entsprechend anders wird bei der Kodierung die Satzvollständigkeit beurteilt.

Insgesamt ist die Übereinstimmung der transkribierten Wörter mit 89 Prozent zufriedenstellend und erfordert kein nochmaliges Überprüfen der Transkripte.

Anhang D: Kodierungen

D 1. Kodiervereinbarungen

D 2. Interrater-Übereinstimmungen

D 3. Ermittlung der Type-Token-Relation

D 1. Kodiervereinbarungen

Ebenso wie für das Anfertigen der Transkripte, wurden die Hilfskräfte von der Autorin instruiert und erhielten die Möglichkeit anhand bereits kodierter Transkripte, sich in das Kodieren einzuarbeiten. Die Kodierungen erfolgten in Excel bzw. SPSS und es wurde die Satznummer, das Segment und die einzelnen kodierten Merkmale der Elternsprache jeweils als eine Variable in der Datei angelegt. Die von den Hilfskräften in Excel kodierten Dateien wurden von der Autorin nach der Kodierung wiederum in SPSS-Dateien übertragen. In dem folgenden Kodiersystem sind die einzelnen Variablen beschrieben.

Tabelle 74: Kodiervereinbarungen

Variable	Inhalt	Kodierung
Satz_Nr	Nummer des Segments in kontrollgehörtem Transkript	Keine → fortlaufende Nummerierung
Satz	Segment aus Transkript wortwörtlich übernommen	Keine
MLU	Länge des Segments in Wörtern (auch „verschliffene“)	0 = Lautmalerei 1 = ein Wort 2 = zwei Wörter usw.
Satz_Vollst	Beurteilung der Vollständigkeit des Satzes	0 = Lautmalerei 1 = eingliedriger Satz 2 = Ellipse 3 = vollständiger Satz 4 = Abbruch
Verb	Anzahl der Verben in einem Satz	0 = kein Verb 1 = ein Verb 2 = zwei Verb usw.
Nebens	Anzahl der Nebensätze	0 = kein Nebensatz 1 = ein Nebensatz 2 = zwei Nebensätze usw.
NP	Anzahl der Nominalphrasen im Satz	0 = keine Nominalphrase 1 = eine Nominalphrase 2 = zwei Nominalphrasen usw.
Anmerkungen		Gr, WDH,

Anmerkungen: Gr = Grammatik im Satz nicht korrekt; WDH = Wiederholung des vorhergehenden Satzes

D 2. Interrater-Übereinstimmungen

Für die Berechnung der Übereinstimmungen der Kodierungen zwischen den Kodierern werden insgesamt elf Vergleiche vorgenommen. Dabei gibt es zwei verschiedene Berechnungen zur Übereinstimmung der Kodierungen:

1. Es werden alle kodierten Spracheinheiten zur Berechnung des Anteils der übereinstimmenden Kodierungen als Basis verwendet.
2. Es werden für die Berechnung des jeweiligen kodierten Merkmals nur die Segmente als Berechnungsgrundlage herangezogen, in denen dieses Merkmal auch enthalten ist. Dies bedeutet z.B., dass in einem Transkript mit insgesamt 149 kodierten Spracheinheiten, in dem insgesamt nur vier Nebensätze vorkommen, das Ergebnis der zweiten Berechnung sehr stark von dem der ersten Berechnung abweichen kann: z.B. hat der eine Kodierer nur drei der vier Nebensätze kodiert, der zweite Vergleichskodierer hingegen alle vier. Nach der ersten Art der Berechnung (alle kodierten Spracheinheiten als Referenzgrundlage) ergibt sich eine Übereinstimmung von 99.33 Prozent. Verwendet man hingegen nur die vier Nebensätze als Berechnungsgrundlage ergibt sich eine Übereinstimmung von 75 Prozent.

Für die komplexe Syntax ergibt sich eine Übereinstimmung von 98.72 Prozent (1. Berechnungsart) bzw. 76.01 Prozent (2. Berechnungsart). Nimmt man den Durchschnitt der Übereinstimmungen über alle kodierten Merkmale der Elternsprache sind es 97.00 Prozent (1. Berechnungsart) bzw. 82.56 Prozent (2. Berechnungsart).

Alle Kodierungen wurden nochmals von der Autorin überprüft und gegebenenfalls korrigiert.

D 3. Ermittlung der Type-Token-Relation

Für die Erhebung der Anzahl verschiedener Wörter (Types) und der Relation dieser zu der Gesamtzahl der Wörter insgesamt (Type-Token-Relation) wurde die frei verfügbare Software Antconc von Anthony (2012)²⁹ verwendet. Diese zählt zum einen die Anzahl der Wörter insgesamt als auch die Anzahl verschiedener Wörter. In das Programm können auch sogenannte Lemmalisten hochgeladen werden, die für eine Grundform des Wortes Einträge für Flexionen und Deklinationen der Verben und Nomen ermöglichen. Damit können diese als zu der Grundform zugehörig erkannt und als dieses Wort gezählt werden. Kommt z.B. das Wort: *gehen* in zwei weiteren Sätzen in der Form „*geht*“ oder „*gegangen*“ vor, so erkennt die Software dies als zu einem Wort zugehörig und zählt dies als ein Type und nicht als drei verschiedene. Die Lemmaliste für die vorliegende Arbeit wurde von der Autorin auf der Basis der Wörter in den Transkripten erstellt.

Um bei der Berechnung der Type-Token-Relation eine vergleichbare Basis für die unterschiedlich langen Transkripte zu schaffen, wurden die Transkriptionsdateien in gleichlange Dateien unterteilt (jeweils 86 Wörter lang, da dies die Anzahl der Wörter des kürzesten Transkripts ist) und die Type-Token-Relation jeweils einzeln berechnet. Daraus wurde dann für jede Datei die durchschnittliche Type-Token-Relation berechnet.

29 Abgerufen am 29.06.2012, <http://www.laurenceanthony.net/>