

L1-Phonemerwerb als dynamischer Prozess

**Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig–Maximilians–Universität
München**

**vorgelegt von
Arne Schneider
aus
München**

Referent: Frau PD Dr. Elsen

Korreferent: Herr Prof. Dr. Vennemann

Tag der mündlichen Prüfung: 11. Februar 2013

Meinen Eltern

Danksagung:

Ich habe in dieser Arbeit ein wenig die Perspektive auf einen Aspekt der menschlichen Entwicklung verändert und in diesem Prozess auch selbst eine Wandlung vollzogen!

Meine Dankbarkeit gilt allen, die mich dabei unterstützt haben.

Für ihre Bereitschaft mich zu betreuen und die guten Hinweise möchte ich Frau PD Dr. Elsen von ganzem Herzen danken! Und auch für Ihre interessanten Hauptseminare! Ebenso bin ich Herrn Prof. Dr. Vennemann und Herrn Prof. Dr. Hinz für ihre offene Kritik sehr dankbar!

Herrn Prof. Dr. Smith möchte ich für die freundliche Bereitstellung seines englischen Phonemerwerbskorpus danken, der aber entgegen meiner ursprünglichen Absicht wegen der Konzentration auf die beiden deutschen Korpora zunächst nicht berücksichtigt werden konnte.

Herrn Prof. Dr. Zaefferer und Frau PD. Dr. Pouplier möchte ich für Ihre interessanten Seminare und die anregenden Diskussionen danken.

Ich möchte Frau Dr. Annamarie Döderlein für ihre Gastfreundschaft in München und Grainau sowie für viele gute Ratschläge ganz herzlich danken und auch ihrer Schwester Frau Adele Döderlein meinen Dank aussprechen!

Besonders dankbar bin ich meinen Eltern Anne und Eberhard Schneider sowie meiner Schwester Rieke, die alle fest von diesem Projekt und meiner Arbeit überzeugt waren.

Kurzzusammenfassung:

Für den kindlichen Spracherwerb (L1-Phonemerwerb) werden Grundzüge eines Modells entwickelt, das auf die Annahme von angeborenen Regeln verzichtet und von einer Selbststeuerung des Prozesses ausgeht, der durch die kognitive Entwicklung, die Entdeckung kausaler Beziehungen, vorangetrieben und durch die Perzeption eigener und fremder Lautäußerungen über eine immanente Abstandsregelung innerhalb eines begrenzten artikulatorischen und phonologischen Raumes gesteuert wird.

Die Silbe wird in diesem Prozess als kleinste Lauteinheit angenommen, aus der sich onto- und phylogenetisch Phoneme über die Bildung von Schnittmengen als nützliche illusorische Einheiten herausbilden (herausgebildet haben).

Da der Spracherwerb des Kindes nicht als Lernprozess verstanden wird, sondern primär als Ausbildung eines eigenständigen Systems, das durch die Perzeption der Zielsprache beeinflusst, aber nicht ursächlich gesteuert wird, wird auf alle Vergleiche zwischen kindlichen Lautäußerungen und ihnen zuzuordnenden Lautäußerungen der Zielsprache verzichtet. Diese Zuordnung wird kritisch gesehen, da semantisch ebenfalls nur von Schnittmengen zwischen kindlichen Lautäußerungen und denen der Zielsprache auszugehen ist.

Die Ergebnisse der quantitativen Untersuchung von zwei Korpora werden als Indiz für die Annahme gewertet, dass der Entwicklungsprozess des Phonemsystems strukturell dem organischen Wachstum in der belebten Natur ähnelt, z.B. dem eines Baumes.

Inhalt

1.	Einleitung	16
2.	Vorbemerkungen zum Modell des evolutionären Phonemerwerbs und Darstellung der grundlegenden Differenzen zu anderen Theorien	18
2.1.	Diskurstheorien	19
2.1.1.	Tomasellos (2006) Evolutionstheorie des Menschen als Beispiel einer Diskurstheorie und das Erkennen von Kausalität durch den Menschen	23
2.1.2.	Piaget	27
2.2.	Phonetische Steuerung des Lauterwerbs und phonetische Begründung der Diskurstheorien - zum Verhältnis von Phonologie und Phonetik	30
3.	Das Phonemerwerbsmodell dieser Arbeit	34
4.	Eigenschaften und Implikationen des Phonemerwerbsmodells	42
4.1.	Der Kausale Motor	42
4.1.1.	Das Erkennen kausaler Zusammenhänge durch das Kind	42
4.1.2.	Der Kausale Motor, die Sprachentstehung und -entwicklung	43
4.2.	Das Schnittmengenkonzept	49
4.2.1.	Schnittmengen im Silbensystem eines einzelnen Sprechers	49
4.2.1.1.	Die funktionale Selbststeuerung der Silbenbildung und die damit verbundene Schnittmengenbildung als ein von Beginn des Phonemerwerbs an wirkender Prozess	50
4.2.1.2.	Substanzbezogenheit der Schnittmengenbildung	50
4.2.1.3.	Hinweise auf die Schnittmengenbildung ohne direkte Einbeziehung des Bewusstseins	51
4.2.1.3.1.	Die Nachbarschaftsdichte eines Wortes	52
4.2.1.3.2.	Die phonotaktische Wahrscheinlichkeit eines Wortes	53
4.2.1.4.	Bewusste Verfügbarkeit der Silbenschnittmengen: Von starker Silbenbindung der Schnittmengen hin zu phonematischen Einheiten.	53
4.2.2.	Schnittmengen zwischen den Silbensystemen unterschiedlicher Sprecher	56
4.2.2.1.	Bests (1995) Modell der Phonemunterscheidungsfähigkeit	57
4.3.	Phoneme als phonetisch-physikalische Invarianten gegenüber denselben als illusionäre Einheiten	59
4.3.1.	Das Verhältnis zwischen einem messbaren akustischen Sprachsignal oder den artikulatorischen Bewegungsabläufen und einer durch einen Sprachteilnehmer wahrgenommenen Phonemfolge	60

4.3.2.	Phoneme als gehörte Gesteninvarianten in der <i>motor theory</i> der Sprachperzeption mit Hilfe eines angeborenen Gestenübersetzungsmoduls	61
4.3.3.	Phoneme als Silbensystemgrößen in Form von Silbenschnittmengen	64
4.4.	Sprachuniversalien und das organische Silbenwachstum	65
4.4.1.	Universalität als angeboren postulierte Sprachstruktur	65
4.4.2.	Universalität als Optimierung von Sprachformen entlang ordinaler Qualitätsskalen	66
4.4.3.	Universalität als allgemeine während des Sprachgebrauchs operierende neuronale Verarbeitungsmechanismen	68
4.4.4.	Universalität als Gesetze des organisch wachsenden Silbensystems	69
4.5.	Ausmaße der Eigenständigkeit des kindlichen Phonemerwerbs und sein Verhältnis zur Zielsprache	70
4.5.1.	Vorschläge anderer Ansätze zum Umfang einer kindbezogenen systematischen Komponente und deren Zielsprachverhältnis	70
4.5.2.	Verhältnis des kindlichen Silbensystems zur Zielsprache im Ansatz des organischen Silbenwachstums	73
4.6.	Das Verhältnis der Perzeption und der Produktion im Phonemerwerb	73
4.6.1.	Die perzeptiven Fähigkeiten des Kindes seit seiner Geburt	73
4.6.2.	Die Produktion des Gehörten durch das Kind	75
4.7.	Problematische Aspekte der Zuordnung von kindlichen Lautformen zu Zielsprachwörtern	77
4.7.1.	Hinweise auf das eigene und individuelle Bedeutungssystem eines Kindes	78
4.7.2.	Bedeutungsbezogene Diskurstheorien zum Spracherwerb und problematische Aspekte der Gruppierung jeweils mehrerer kindlicher Lautformen nach Zielsprachwörtern	80
4.7.3.	Aus der Gruppierung von kindlichen Lautformen sich ergebende Phänomene als Gründe für die Annahme des Wortes als grundlegender Einheit im Phonemerwerb	82
4.8.	Die Rolle der Funktionalität in Lindbloms (1999, 2000) Emergenzphonologie	83
5.	Die Silbe als eine grundlegende Einheit	87
5.1.	Die Silbe in der Sprachverarbeitung	87
5.2.	Levelt et al.s (1999) Modell des Lexikonzugriffs während der Sprachproduktion	89
5.3.	Funktionen der Silbenkonstituenten während des Silbenwachstums	91
5.4.	Vorstellung und Merkmale verschiedener Silben(erwerbs)-theorien	95
5.4.1.	Silben(erwerbs)theorien	96
5.4.1.1.	Das Koordinationsmodell (im Rahmen der Artikulatorischen Phonologie)	96

5.4.1.2.	Stetsons (1951) Silbendefinition	98
5.4.1.3.	Die Frame-Content-Theorie (MacNeilage (1998, 2008))	101
5.4.1.4.	Die Sonoritätshierarchie	103
5.4.1.5.	Die Rolle der Silbe in Moskowitzs (1970) Phonemerwerbsansatz	108
5.4.1.6.	Die Silbe als Epiphänomen in Ohalas (2008) Kontrastansatz	109
5.4.2.	Notwendige Merkmale einer Silbendefinition	112
6.	Phonem(erwerbs)theorien und -aspekte	113
6.1.	Weitere Unterscheidungsmerkmale der vorgestellten Theorien	113
6.2.	Vorstellung einiger Phonem(erwerbs)ansätze	114
6.2.1.	Konnektionismus	114
6.2.2.	Jakobsons Theorie der distinktiven Merkmale und deren Anwendung auf den Phonemerwerb	117
6.2.3.	Bybees Gebrauchsbasierte Phonologie (2001)	120
6.2.4.	Vihmans und Crofts sog. "Radikale schablonenbasierte Phonologie"	124
6.2.5.	Pierrehumberts Probabilistische Phonemtheorie	127
6.2.6.	DeBoers Vokalevolutionssimulation	129
6.2.7.	Artikulatorische Phonologie	132
6.2.8.	Der Erwerb artikulatorischer Muster	135
7.	Untersuchungen von Konstituentenverteilungen in Korpora	141
7.1.	Anforderungen an die Beschaffenheit von Korpora zur Untersuchung des Phonemerwerbs	141
7.2.	Beschreibung der verwendeten Korpora und Aufbereitung der Daten	143
7.2.1.	Das Annalena-Korpus (Elsen 1991)	143
7.2.2.	Das ISA-Korpus (Lindner 2009)	144
7.3.	Verteilungsanalyse der Silbenkonstituenten in den Korpora	145
7.3.1.	Ausgangslage, Fragestellung und methodisches Vorgehen	145
7.3.2.	Darstellungsform der Ergebnisse und ihre Grundquantitätsarten	147
7.3.3.	Aufführung und Wiederholung einiger häufiger Begriffe	149
7.3.4.	Zur Gliederung des Vorgehens bei der Analyse	150
7.3.5.	Statische Verteilungen (nicht-relational)	150
7.3.5.1.	Silbenpositionsbezogene Gesamthäufigkeiten von Einzelphonemen	151
7.3.5.2.	Silbengesamthäufigkeitsverteilung der häufigsten Silbenkerne	153
7.3.5.3.	Häufigkeitsverteilungen der Silbengesamthäufigkeiten von Silbenbestandteilen einer bestimmten Art	155

7.3.5.4.	Verteilungen der Binärrelationsarten im Gesamtkorpus	158
7.3.5.5.	Gesamthäufigkeiten von Silben in den Lautformen	159
7.3.6.	Statische Verteilungen (relational)	161
7.3.6.1.	Verteilungen der häufigsten Köpfe und Kodas in der Gesamtschalenmenge	161
7.3.6.2.	Relation aller Silbenkerne in ihrem Verhältnis von neuen zu bereits verwendeten Schalen	163
7.3.7.	Statisch-dynamische Verteilungen	165
7.3.7.1.	Quantitative und zeitliche Relationen in der Silbenentwicklung	165
7.3.8.	Dynamische Verteilungen (nicht-relational)	169
7.3.8.1.	Silbenproduktivitätsänderungen im Zeitverlauf	169
7.3.8.2.	Wachstumsgeschwindigkeiten von Silbenbestandteilen	172
7.3.8.3.	Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in gleichgroßen Silbenmengen im Zeitverlauf	177
7.3.8.4.	Wiederverwendung von Silbenbestandteilen	182
7.3.8.5.	Anteil der jeweils von den 3, 10 und 20 häufigsten Kernen gebildeten Silbenmenge im Zeitverlauf	183
7.3.9.	Dynamische Verteilungen (relational)	186
7.3.9.1.	Zeitliches Kombinationsverhalten konkreter Köpfe mit Kodas	186
7.3.9.2.	Vergleich zeitbezogener Binärrelationsmengen	190
7.3.9.3.	Das Verhältnis aus einer Silbenproduktivitätszunahme und dem Schalenwachstum	195
7.3.9.4.	Schalengemeinsamkeiten von Silbenkernen	206
7.3.9.5.	Zunehmende "Wiederverwendung" von Konstituenten bei der Silbenbildung	214
7.3.9.6.	Zunehmende Streuung bei der Silbenbildung	215
7.3.9.7.	Silbenkernpaarbasierte Phasenbildung	216
7.3.9.8.	Die Bildung von Köpfen und Kodas unter dem Gesichtspunkt phonetischer Merkmale und in Bezug auf die Silbenkernentwicklung	225
7.3.9.9.	Verhältnis von bereits aufgetretenen komplexen Schalen zu nicht-komplexen Schalen oder bereits aufgetretenen C1-Komplexen zu einfachen C1-Köpfen	237
7.3.9.10.	Anteil der Silben in gleich großen Zukunftsteilzeiträumen, die noch vollständig aus Binärrelationen des Vergangenheitszeitraums gebildet werden	241
7.3.9.11.	Relation zwischen dem Silbenanfang und dem ersten Phonem eines Silbenkerns hinsichtlich ihrer phonetischen Lage (Artikulationsort, horizontale Zungenlage)	244
7.3.10.	Diskussion der Beobachtungen und Ergebnisse der Verteilungsuntersuchungen in den Korpora	246
8.	Ausblick	257

9.	Die Digitalisierung, Auswertung und Notation der Korpora	258
10.	Literatur	260

Verzeichnis der Abbildungen und Tabellen

Abb.	1	: Evolutionär kognitives Kommunikationsmodell	45
Abb.	2	: Sonoritätshierarchie des Deutschen	105
Abb.	3	: Silbenpositionsbezogene Silbengesamthäufigkeit von /t/	152
Abb.	4	: Silbenpositionsbezogene Silbengesamthäufigkeit von /ç/	153
Abb.	5	: Silbengesamthäufigkeiten der 20 häufigsten Nuklei (Annalena-Korpus)	154
Abb.	6	: Silbengesamthäufigkeiten der 20 häufigsten Nuklei (Isa-Korpus)	154
Abb.	7	: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Köpfe	156
Abb.	8	: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Nuklei	156
Abb.	9	: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Kodas	157
Abb.	10	: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Schalen	158
Abb.	11	: Häufigkeitsverteilung der Lautformengesamthäufigkeiten aller Silben (Annalena-Korpus)	160
Abb.	12	: Häufigkeitsverteilung der Lautformengesamthäufigkeiten aller Silben (Isa-Korpus)	161
Abb.	13	: Gesamthäufigkeit der 30 häufigsten Köpfe in der Schalen gesamtmenge	162
Abb.	14	: Gesamthäufigkeit der 30 häufigsten Kodas in der Schalen gesamtmenge	162
Abb.	15	: Streudiagramm des Schalenverhältnisses und der Silbengesamthäufigkeit aller Nuklei (Annalena-Korpus)	164
Abb.	16	: Streudiagramm des Schalenverhältnisses und der Silbengesamthäufigkeit aller Nuklei (Isa-Korpus)	164
Abb.	17	: Zeitliche Anordnung aller Schalen nach dem Zeitpunkt ihres ersten Auftretens und Darstellung ihrer Silbengesamthäufigkeiten	166
Abb.	18	: Streudiagramm des Zeitpunktes des ersten Auftretens und der Silbengesamthäufigkeit von Schalen	167
Abb.	19	: Streudiagramm der /a/-Silbengesamthäufigkeit eines Kopfes und des Zeitpunktes des ersten Auftretens beim Nukleus /a/	168
Abb.	20	: Streudiagramm der /a/-Silbengesamthäufigkeit einer Koda und des Zeitpunktes ihres ersten Auftretens beim Nukleus /a/	168
Abb.	21	: Kumulative Häufigkeit unterschiedlicher Silben (Annalena-Korpus)	170
Abb.	22	: Kumulative Häufigkeit unterschiedlicher Silben (Isa-Korpus)	170
Abb.	23	: Kumulative Häufigkeiten aller Nuklei	173
Abb.	24	: Kumulative Häufigkeiten aller Nuklei (mit Kennzeichnungen)	174

Abb.	25	:	Kumulative Häufigkeit des Laterals /l/	175
Abb.	26	:	Kumulative Häufigkeit des Kopfs /ð/	175
Abb.	27	:	Kumulative Häufigkeit des Kopfs /f/	176
Abb.	28	:	Kumulative Häufigkeit des Kopfs /θ/	176
Abb.	29	:	Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in je 50 Silben im Zeitverlauf (Annalena-Korpus)	177
Abb.	30	:	Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in je 50 CVC-Silben im Zeitverlauf (Annalena-Korpus)	180
Abb.	31	:	Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in je 10 Silben im Zeitverlauf (Isa-Korpus)	181
Abb.	32	:	Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in je 10 CVC-Silben im Zeitverlauf (Isa-Korpus)	182
Abb.	33	:	Anteil bereits aufgetretener Konstituenten und Binärrelationen an 20-Silbenzeiträumen im Zeitverlauf	183
Abb.	34	:	Anteil der 3 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben	184
Abb.	35	:	Anteil der 10 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben	185
Abb.	36	:	Anteil der 20 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben	185
Abb.	37	:	Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /p/	186
Abb.	38	:	Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /m/	189
Abb.	39	:	Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /b/	189
Abb.	40	:	Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 2760-2780)	191
Abb.	41	:	Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 3260-3280)	192
Abb.	42	:	Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 6060-6080)	193
Abb.	43	:	durchschnittliche Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Annalena-Korpus)	194
Abb.	44	:	durchschnittliche Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Isa-Korpus)	195
Abb.	45	:	Sitzungsbezogene Anzahl neuer Silben (Annalena-Korpus)	196
Abb.	46	:	Kumulative sitzungsbezogene Anzahl unterschiedlicher Schalen und Nuklei (Annalena-Korpus)	197
Abb.	47	:	Verhältnis aus der kumulativen Anzahl unterschiedlicher Schalen zu der unterschiedlicher Nuklei (Annalena-Korpus)	198
Abb.	48	:	Kumulative Anzahl unterschiedlicher Köpfe, Kodas und Nuklei (Annalena-Korpus)	199
Abb.	49	:	Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /a/ im Zeitverlauf (Annalena-Korpus)	201

Abb.	50	: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /ɪ/ im Zeitverlauf (Annalena-Korpus)	201
Abb.	51	: Sitzungsbezogene Anzahl neuer Silben (Isa-Korpus)	202
Abb.	52	: Kumulative sitzungsbezogene Anzahl unterschiedlicher Schalen und Nuklei (Isa-Korpus)	203
Abb.	53	: Verhältnis aus der kumulativen Anzahl unterschiedlicher Schalen zu der unterschiedlicher Nuklei (Isa-Korpus)	203
Abb.	54	: Kumulative Anzahl unterschiedlicher Köpfe, Kodas und Nuklei (Isa-Korpus)	204
Abb.	55	: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /a/ im Zeitverlauf (Isa-Korpus)	205
Abb.	56	: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /ɪ/ im Zeitverlauf (Isa-Korpus)	205
Abb.	57	: Schalengemeinsamkeiten von Silbenkernpaaren (Annalena-Korpus)	207
Abb.	58	: Schalengemeinsamkeiten von Silbenkernpaaren (Isa-Korpus)	208
Abb.	59	: Streudiagramm der Schalengemeinsamkeit und des Minimums der Silbengesamthäufigkeiten aller Silbenkernpaare (Annalena-Korpus)	208
Abb.	60	: Streudiagramm der Schalengemeinsamkeit und des Minimums der Silbengesamthäufigkeiten aller Silbenkernpaare (Isa-Korpus)	209
Abb.	61	: Streudiagramm des ersten Zeitpunkts des Erreichens und Beibehaltens von Nicht-Inklusionsverhältnissen der Schalenmengen von Silbenkernpaaren und der Größe jeweils von Gruppen der häufigsten Silbenkerne (Annalena-Korpus)	211
Abb.	62	: Ausmaß der Schalengemeinsamkeiten der Silbenkerne /a/ und /ɪ/	212
Abb.	63	: Zeitlicher Verlauf der maximalen paarweisen Schalengemeinsamkeiten der 20 häufigsten Silbenkerne	213
Abb.	64	: Streudiagramm des ersten Zeitpunkts des Erreichens und Beibehaltens von Nicht-Inklusionsverhältnissen der Schalenmengen von Silbenkernpaaren und der Größe jeweils von Gruppen der häufigsten Silbenkerne (Isa-Korpus)	213
Abb.	65	: Verhältnis aus der Anzahl der mehrmals zu der der einmal verwendeten Konstituenten im Zeitverlauf	215
Abb.	66	: Zeitliche Anordnung aller Schalen und Darstellung ihrer einmaligen oder mehrfachen Verwendung	216
Abb.	67	: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 1-720	217
Abb.	68	: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 721-1290	218
Abb.	69	: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 1291-1800	218
Abb.	70	: Häufigkeitsverteilung der Silbenkerngruppen auf dem 1-ten Rang	223
Abb.	71	: Häufigkeitsverteilung der Silbenkerngruppen auf dem 2-ten Rang	223
Abb.	72	: Häufigkeitsverteilung der Silbenkerngruppen auf dem 3-ten	224

		Rang	
Abb.	73	: Häufigkeitsverteilung der Silbenkerngruppen auf dem 4-ten Rang	224
Abb.	74	: Differenzierung der Köpfe nach Artikulationsarten im Zeitverlauf	227
Abb.	75	: Differenzierung der Kodas nach Artikulationsarten im Zeitverlauf	227
Abb.	76	: Differenzierung der Köpfe aller /a/-Silben nach Artikulationsarten im Zeitverlauf	228
Abb.	77	: Differenzierung der Kodas aller /a/-Silben nach Artikulationsarten im Zeitverlauf	229
Abb.	78	: Differenzierung der Köpfe aller /ɪ/-Silben nach Artikulationsarten im Zeitverlauf	229
Abb.	79	: Differenzierung der Kodas aller /ɪ/-Silben nach Artikulationsarten im Zeitverlauf	230
Abb.	80	: Differenzierung der Köpfe aller /ʊ/-Silben nach Artikulationsarten im Zeitverlauf	230
Abb.	81	: Differenzierung der Kodas aller /ʊ/-Silben nach Artikulationsarten im Zeitverlauf	231
Abb.	82	: Differenzierung der Köpfe nach (Nicht-)Stimmhaftigkeit im Zeitverlauf	231
Abb.	83	: Differenzierung der Kodas nach (Nicht-)Stimmhaftigkeit im Zeitverlauf	232
Abb.	84	: Differenzierung der Köpfe aller /a/-Silben nach (Nicht-)Stimmhaftigkeit im Zeitverlauf	232
Abb.	85	: Differenzierung der Kodas aller /a/-Silben nach (Nicht-)Stimmhaftigkeit im Zeitverlauf	233
Abb.	86	: Häufigkeitsverteilung der Artikulationsortmerkmale in allen Phonemen aller Köpfe	233
Abb.	87	: Häufigkeitsverteilung der Artikulationsortmerkmale in allen Phonemen aller Kodas	234
Abb.	88	: Differenzierung der Köpfe nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf	234
Abb.	89	: Differenzierung der Kodas nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf	235
Abb.	90	: Differenzierung der Köpfe aller /a/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf	235
Abb.	91	: Differenzierung der Köpfe aller /ɪ/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf	236
Abb.	92	: Differenzierung der Köpfe aller /ʊ/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf	236
Abb.	93	: Anteil (nicht-)komplexer Schalen an der Schalenbildung im Zeitverlauf	238
Abb.	94	: Vergleich der durchschnittlichen C ₁ -Komplexsilbenanteile mit den entsprechenden durchschnittlichen C ₁ -Silbenanteilen im Zeitverlauf (Annalena-Korpus)	239
Abb.	95	: Vergleich der durchschnittlichen C ₁ -Komplexsilbenanteile mit den entsprechenden durchschnittlichen C ₁ -Silbenanteilen im Zeitverlauf (Isa-Korpus)	239

Abb.	96	: Vergleich der /f/-Komplexsilbenanteile mit den entsprechenden /f/-Silbenanteilen im Zeitverlauf	240
Abb.	97	: Vergleich der /g/-Komplexsilbenanteile mit den entsprechenden /g/-Silbenanteilen im Zeitverlauf	240
Abb.	98	: Vergleich der /d/-Komplexsilbenanteile mit den entsprechenden /d/-Silbenanteilen im Zeitverlauf	241
Abb.	99	: Anteil der vollständig aus Binärrelationen der Vergangenheit (1.-800.-te Silbe) gebildeten Silben von Zukunftsteilzeiträumen	242
Abb.	100	: Anteil der vollständig aus Binärrelationen der Vergangenheit (1.-2500-te Silbe) gebildeten Silben von Zukunftsteilzeiträumen	243
Abb.	101	: Häufigkeitsverteilung von Artikulationsort und horizontaler Zungenlage aller C-V-Paare der Anfangssilbenmenge	244
Abb.	102	: Kumulative Häufigkeit des Körpers /ba/	245
Abb.	103	: Kumulative Häufigkeit des Körpers /ga/	245
Abb.	104	: Kumulative Häufigkeit des Körpers /fi/	246
Abb.	105	: Die Selbststeuerung des Silbenwachstums durch eine Abstandsregelung	246
Tabelle	1	: Zehn dem Zielsprachwort <i>pen</i> zugeordnete kindliche Lautformen (Ferguson und Farwell 1975 in Piske 2001: 61)	81
Tabelle	2	: C ₁ C ₂ -Kopf-Komplexe des Standarddeutschen	104
Tabelle	3	: C ₁ C ₂ -Koda-Komplexe des Standarddeutschen	104
Tabelle	4	: Anzahl unterschiedlicher Schalen, Körper und Reime in der Silbengesamtmenge	158
Tabelle	5	: Anzahl aller unterschiedlichen Silben und Lautformen und Summe aller Silbenhäufigkeiten im Gesamtkorpus	160
Tabelle	6	: Wachstumsratenänderungen der Anzahl neuer Silben	172
Tabelle	7	: Silbenkernpaarbasierte Phaseneinteilung der ersten beiden Ränge	219
Tabelle	8	: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 3 u. 4	221
Tabelle	9	: Häufigkeit der Phasenbildung (1. u. 2. Rang) eines Kernpaares	222

1. Einleitung

Mein Interesse an der Kindersprache wurde durch Jakobsons "Kindersprache"(1968) geweckt. Obwohl Jakobson und mit ihm nahezu die gesamte Linguistenwelt diese Schlussfolgerung abgelehnt hätte, weckten die von ihm beobachteten Gesetzmäßigkeiten im Phonemerwerb des Kindes und die in einem umgekehrten Verhältnis stehenden Sprachverluste bei Aphasikern in mir die Vorstellung, dass der Spracherwerb sich dynamisch vollzieht und im Zusammenhang der evolutionären Entwicklung des Menschen gesehen werden sollte.

Wenn die Evolution im Spracherwerb gelten sollte, sind alle Wege zu Vorstellungen von angeborenen Regeln verstellt und es ist erforderlich, alle Fragen des Spracherwerbs in einem evolutionären Zusammenhang zu erklären: Warum spricht der Mensch, warum wachsen Kinder in einen bestimmten Sprachzusammenhang hinein, warum und wie verändert sich die Sprache, welchen Einfluss haben die Sprachorgane auf die Sprachentwicklung?

Es sollte doch möglich sein, Ansätze zu einem schlüssigen Modell der Sprachentwicklung zu formulieren, wobei der Gegenstand der Untersuchung auf den Phonemerwerb zu beschränken war, da er sich für eine Quantifizierung anbietet. Die Quantifizierung eines evolutionären Prozesses erfordert jedoch Kategorien, die das Phonemmaterial ordnen. Unmittelbar aus den phonetisch-physiologischen Grundlagen des Sprach- und Phonemerwerbs lassen sich diese Kategorien nicht gewinnen, da die Artikulationsorgane nicht aus eigener Initiative tätig werden, sondern als Werkzeuge geistiger Prozesse.

Die Silbe bietet sich als grundlegende Einheit des Spracherwerbs an, da dies zum einen artikulatorisch begründet ist. Sie spielt aber auch wegen ihres gehäuftten Auftretens als Einzelsilbe oder Reduplikation im frühen Spracherwerb eine wichtige Rolle, da sie in dieser Phase der Sprachentwicklung offensichtlich mit einer Bedeutung für das Kind verbunden ist und kognitiv und kommunikativ genutzt wird. Welche konkrete Bedeutung den Silbenfolgen eines Kindes zugeordnet werden kann, lässt sich nur aus dem Kontext vermuten, aber nicht eindeutig klären. Diese Überlegung führte zu der Annahme, dass Bedeutungsschnittmengen für die Kommunikation zwischen Sprechern relevant sind und dass eine Zuordnung von kindlichen Lautformen zu angenommenen Bedeutungen von Lautformen der Zielsprache willkürlich ist und die Eigenständigkeit der kindlichen Sprachentwicklung übersieht. Deshalb wurde vollkommen auf die möglichen Bedeutun-

gen von kindlichen Lautäußerungen als Ordnungskategorie des Lautmaterials verzichtet, stattdessen wird die Existenz bedeutungsvoller sprachlicher Lautformen angenommen, die aber für die Analyse des Lautsystems keine Rolle spielen. Mit diesem Schritt waren weitere Fragen verbunden: In welchem Zusammenhang stehen die Lautformen des Kindes, wenn sie nicht direkt an eine Zielsprachform gebunden sind?

Im Sinne einer evolutionären Erklärung des Spracherwerbs musste sich die Antwort aus der Entwicklung der kindlichen Lautformen selbst ergeben, aus einem eigenen Systemzusammenhang, der zwar von der Zielsprache beeinflusst, aber nicht ursächlich gesteuert wird. Zum Nachweis des angenommenen Zusammenhangs wurde das Konzept der Schnittmengen auf die Silben des Kindes angewendet. Der Gedanke der Schnittmengen ermöglicht, die Dynamik eines Prozesses und seiner Veränderungen zu erfassen und sich von statischen Zuordnungen zu lösen. Die Silben stehen über Schnittmengen in einem systeminternen Zusammenhang, der zum einen als autonom, aber gleichzeitig auch als adaptiv an die Zielsprache zu sehen ist.

Aus dieser Stärkung der Seite des Kindes gegenüber vermuteten Einflüssen der Zielsprache, ergaben sich auch der Begriff des Silbensystems des Kindes und weitere Akzentuierungen der Kindposition wie der des kausalen Motors, etc., die in einen größeren theoretischen Zusammenhang gestellt wurden.

Schwierigkeiten ergaben sich auch bei der Bestimmung der Silben innerhalb der Korpora: Wie soll die phonologische Einheit der Silbe, die in manchen Theorien überhaupt keine Beachtung findet oder eher als Fremdkörper behandelt wird, der sich aus den invarianten Phonemen zusammensetzt und nur dann verwendet wird, wenn es die Beschreibung von phonologischen Schwierigkeiten erfordert (Bell, Hooper 1978: 4), bei der Untersuchung der Korpora berücksichtigt werden?

Zunächst habe ich den Schnittmengenbegriff auf die Einzelphoneme in der Silbe angewendet, ohne Bezug zur Silbenposition. Doch es zeigte sich, dass sowohl die Positionen der Phoneme als auch ihre Zusammenfassung zu Silbenkonstituenten in der Entwicklung von Bedeutung sind, so dass der Schnittmengenbegriff in seiner jetzigen Form auf Silbenkonstituenten angewendet wird.

Immer wieder war mit der Untersuchung der Schnittmengen auch die Frage nach der Silbendefinition aktuell: So wurden mehrere Varianten entwickelt, die Silbengrenzen durch Algorithmen automatisch zu bestimmen, z.B. durch eine dynamische Berücksichtigung von Merkmalen bereits gebildeter Silben wie ihrer Schalen oder der Häufigkeit von Einzelkonstituenten oder -phonemen. Die Ergebnisse der verschiedenen Einteilungen stell-

ten sich aber jeweils als nicht zufriedenstellend heraus, insbesondere hinsichtlich der zwischen zwei Silbenkernen befindlichen Konsonantenkomplexe.

Bei der Untersuchung der Korpora habe ich zunächst das Augenmerk auf abstrakte Muster gelegt, die im Korpus vermutet wurden. Doch mit der fortschreitenden Analyse rückten immer mehr die allgemeinen Wachstumsprinzipien in den Mittelpunkt der Aufmerksamkeit, die sich auf den Kontrast zwischen der Ökonomie der Silbenbildung und ihrer Differenzierung reduzieren lassen. Es zeigten sich Anzeichen von Regelmäßigkeiten, die einer zufälligen Silbenbildung durch das Kind widersprechen. Vielmehr deuten sie auf einen "organischen" Prozess der Selbststeuerung hin.

2. Vorbemerkungen zum Modell des evolutionären Phonemerwerbs und Darstellung der grundlegenden Differenzen zu anderen Theorien

Als ich am Anfang meiner Arbeit zur kindlichen Sprachentwicklung die verschiedenen wissenschaftlichen Positionen untersuchte, fragte ich mich, ob sich ein gemeinsamer Nenner formulieren lässt, der eine Einordnung und Beurteilung der Forschungsansätze ermöglicht, eine Art Lackmustest der Linguistik. Trotz vieler Unterschiede im Einzelnen ließen sich zwei entgegengesetzte Positionen feststellen unter der Voraussetzung, dass die Evolution in der Sprachentwicklung und dem Spracherwerb eine Rolle spielt. Unberücksichtigt bleiben in dieser Gegenüberstellung alle Theorien, die von angeborenen Regeln, etc. ausgehen. Die beiden zuletzt genannten Positionen können mit den beiden folgenden Fragen erfasst werden:

- 1.) Entsteht die Sprache aus dem Diskurs?
- 2.) Entwickelt sich Sprache im Einklang mit der Evolution?

Analog zu der Frage, ob ein Stoff sauer oder basisch ist, zwei entgegengesetzte Zustände, die sich zwar ausschließen, zwischen denen es aber Abstufungen gibt, soll die Frage formuliert werden: Hat Sprache sich im Einklang mit den Erkenntnissen der Evolution entwickelt, wobei diese Entwicklung fortgesetzt weiterläuft oder gibt es sprachspezifische Baupläne und Regelwerke, die in jedem Menschen von Beginn an die Sprachverwendung steuern? Die zuletzt genannte Ansicht ist zunehmend in Frage gestellt worden. Eine evolutionäre Sprachtheorie muss den Anspruch haben, die Entstehung des Sprachsystems aus Prinzipien der Evolution selbst erklären zu können.

Die Entwicklung der Menschheit wie der gesamten belebten Natur wird von ihrem Lebens- und Reproduktionswillen angetrieben, auf dessen Grundlage sich die Sprachfähigkeit des Menschen aus tierischen Signalsystemen heraus entwickelt hat. Um ihrem kogni-

tiven und kommunikativen Zweck dienen zu können, müssen sprachliche Signale eine vorrangige Eigenschaft aufweisen, die in ihrer Eindeutigkeit besteht. Dieser Zwang zur Eindeutigkeit ermöglicht die Entwicklung sprachlicher und insbesondere lautlicher Systeme als sich selbst steuernder Prozesse, die zu ihrer Erklärung auf keine externen Faktoren irgendwelcher Art angewiesen sind.

Den zunächst im Folgenden in dieser Arbeit diskutierten theoretischen Ansätzen ist gemeinsam, dass sie die Laut- und Sprachentwicklung nicht aus den immanenten Systemvoraussetzungen selbst erklären. Die Theorie einer ersten Gruppe, auf die in dieser Arbeit nicht weiter Bezug genommen wird, da sie wesentliche Momente der Sprachfähigkeit als angeboren ansieht, findet sich dennoch ungewollt auch in anderen Gruppen von Theorien, die eine evolutionäre Erklärung der Sprachentwicklung anstreben. Zu diesen gehören als zweite Gruppe alle Ansätze, die die Kommunikation oder den Diskurs als Ursache der Sprachveränderung verstehen. Die dritte Gruppe bilden Theorien einer phonetischen Basierung und Steuerung der Sprachveränderung. Übergänge zwischen diesen Ansätzen sind fließend, und sie stimmen in der Annahme von der Existenz subjektiver oder objektiver Optimierungsstrategien überein, die sich in der psychologischen Theorie Piagets zu einer deutlich teleologisch geprägten Weltsicht verdichten. Die drei Theoriegruppen sollen zunächst ansatzweise in ihrem inneren Zusammenhang dargestellt und diskutiert werden, bevor die Kritik in die Darstellung der hier vertretenen evolutionären Theorie überleitet.

2.1. Diskurstheorien

Findet die Sprachentwicklung im Einklang mit der Evolution statt, hat diese Auffassung eine Reihe von Konsequenzen: Vor allem kann das Wirken von Regeln nicht mehr vorausgesetzt werden, sondern diese müssen selbst in ihrer evolutionären Entstehung und Veränderung erklärt werden ebenso wie die Gesamtheit der sprachlichen Phänomene. Damit müsste die Entstehung des phonologischen Systems eines Kindes auf eine komplexe innere Kausalität zurückgeführt werden, die nicht das Ergebnis von Zufällen ist oder von angeborenen Faktoren. Bevor ich jedoch auf diesen Punkt näher eingehe, sollen einige theoretische Vorstellungen betrachtet werden, die den Anspruch erheben, den evolutionären Charakter der Sprache zu berücksichtigen. Gegenüber nativistischen Ansätzen sind sie der Ansicht, dass die Sprache erworben wird. Dem Kind werden z.B. Wörter der Zielsprache "gegeben", das Kind "übernimmt" sie, gibt sie jedoch in unvollkommener Form, d.h. fehlerbehaftet zurück. Allen in den folgenden Ausführungen vorgestellten

Theorien ist in unterschiedlichen Formen die Annahme gemeinsam, Spracherwerb, der Input für die sprachliche Entwicklung des Kindes, vollziehe sich in der Kommunikation (im Diskurs). Sie werden deshalb als *Diskurstheorien* bezeichnet. Diese grundsätzliche Annahme der Diskurstheorien und der in dieser Arbeit vertretene Standpunkt, den Spracherwerb als sich selbst steuerndes individuelles System unter perzeptiver Beeinflussung durch die Zielsprache zu sehen, schließen sich aus, zwischen ihnen gibt es keine Kompatibilität. Bei Tomasello (2006) haben die Erwachsenen konstant die Rolle des Sprechers inne und das Kind die des Hörers: Kinder müssen möglichst effizient und verlustfrei einen Großteil der bereits angehäuften Informationen der Menschheit "verinnerlichen" (ebd.). Dabei ist ihnen die angeborene intentionale Kognition besonders hilfreich: Sie ermöglicht den Kindern auch das Lernen sprachlicher Symbole und kulturellen Wissens (ebd.). Das Kind erscheint so als ein Speichermedium, das erfolgreich mit Informationen 'bespielt' werden kann, weshalb die Kinder von Tomasello auch folgerichtig als "Imitationsmaschinen"(ebd.) bezeichnet werden (ebd.). Die Kommunikation spielt auch bei deBoer (2001) die zentrale Rolle bei der Entstehung von Lautsystemen. Diese umschreibt er mit dem Begriff eines "Imitationsspiels"(ebd.), dessen Ziel es ist, eine Identität der Lautäußerungen von Sprecher und Hörer zu erreichen (ebd.). Das Kind wird in der Rolle des Imitierenden gesehen, indem deBoer sagt, dass "[...] [c]hanging pronunciation in order to match others more closely is also necessary for children learning a language"(ebd.: 52). In der Artikulatorischen Phonologie (Studdert-Kennedy (1992)) soll das Kind Wörter der Zielsprache in einer "acceptable pronunciation"(ebd.: 97) wiedergeben. Der Konnektionismus (Stemberger (1992)) rückt den Diskurs ebenfalls in den Vordergrund, weil angenommen wird, dass das Kind die Abweichungen seiner Lautäußerungen von der Zielsprache durch einen ständigen Abgleich korrigiere, auf der Grundlage einer Kenntnis der exakten Phonemsequenzen (ebd.). In Bybees (2004) Phonemansatz der Gebrauchs-basierten Phonologie ist die Kommunikation ebenfalls der primäre Bestandteil, weil der Gebrauch von Sprache im Mittelpunkt steht: Die Art, wie Sprache im Gebrauch verwendet wird, bedingt die Form derselben. Veränderungsmöglichkeiten der Sprache werden aus der Automatisierung von Bewegungen beim Sprechen und Hören abgeleitet (ebd.).

Welche Folgen hat die Annahme, dass das Kind nur eine reproduktive Leistung erbringt, die keine eigene produktive Leistung ist, oder deren Produktivität in der zunehmenden Adaption von Zielsprachstrukturen besteht?

Zunächst einmal wird der Gegenstand der Untersuchung festgelegt. Man könnte sagen, dass der Blick sich nicht auf die Kindersprache richtet, sondern auf vermutete inhärente Regelwerke, die erklären sollen, warum das Kind die Zielsprachwörter nicht exakt wiedergeben kann. Aus der Annahme der hauptsächlich reproduktiven Leistung des Kindes ergibt sich in der Folge eine weitere Annahme: Die Abweichungen des Kindes von den ihm "übergebenen" Wörtern werden als echte Fehler angesehen, weil es "eigentlich" die Wörter so äußern wolle, wie sie ihm übergeben wurden. Die kindliche Sprache ist also ein Fehlersystem, dessen Abweichungen von den zielsprachlichen Normen erklärt werden müssen¹. So sagt beispielsweise Stemberger (1992) als ein Vertreter des Konnektionismus:

The child [...] attempts to reproduce the perceived form faithfully. All changes are accidental; they are true errors, in which the child's system produces a result that it was not designed to produce. As the child's pronunciation changes over time, the improvements derive from increased accuracy in the access of phonological information. (Stemberger 1992: 166)

Warum weicht die Kindersprache von den Diskursvorgaben ab?

Wenn man davon ausgeht, dass die sprachlichen Äußerungen des Kindes Ergebnis einer unvollkommenen oder fehlenden Umsetzung von vorhandenen korrekten Regeln sind oder des inadäquaten Arbeitens eines neuronalen Netzwerkes (vgl. Smith 2010, Studdert-Kennedy 1992) oder einer nicht ausreichenden artikulatorischen Übung (vgl. Vihman 2007, Piske 2001), müssen die Abweichungen des Kindes von den "übergebenen" Wörtern entweder als Resultat einer fehlerhaft arbeitenden Systemkomponente dargestellt werden oder als Rückgriff auf bereits geübte artikulatorische Routinen, die aber nicht für alle "übergebenen" Wörter ausreichen.

Wie wird auf dieser Grundlage die Überwindung der Fehler erklärt?

Da sich das Fehlerkonzept des Spracherwerbs aus der Annahme herausbildet, dass der Spracherwerb eine Folge der Kommunikation ist, werden folgerichtig auch die Heilmittel für die Fehlerüberwindung in der Kommunikation gesehen und zwar insbesondere in pädagogischen Maßnahmen. Auf diese Weise wird der Spracherwerb zu einem pädagogischen Problem. In der Folge wird das Kind nicht mehr als Subjekt gesehen, sondern nur noch als Objekt sprachpädagogischer Einflüsse.

¹ vgl. auch die ausführliche Besprechung des Konnektionismus nach Stemberger (1992) in Abschnitt 6.2.1. oder die Besprechung des Konnektionismus in Elsen (1999: 21ff.), insbesondere:

Die kindlichen Abweichungen sind echte Fehler: die Kinder wollen so sprechen wie die Erwachsenen, können es aber (noch) nicht. Dies zeigen Fehlversuche bei der Aussprache verschiedener Wörter, Vermeidungen und Schwierigkeiten bei komplexen Sätzen. (Elsen 1999: 22, 207f.)

Die Fehler des Kindes werden dadurch überwunden, dass es durch einen ständigen Input während des Diskurses entweder über die Fehlerrückmeldung eines Lehrers oder die Musterextraktionsfähigkeiten eines neuronalen Netzwerkes die Zielsprachregularitäten in zunehmendem Maße adäquat repräsentiert.

Woher kommt die Input-Sprache?

Wenn wir annehmen, dass die Sprache des Kindes aus dem Diskurs entsteht, müssen wir fragen, woher diese Input-Sprache stammt. Die Antwort darauf könnte wie bei Tomasello (2006) lauten, dass sie das gespeicherte Ergebnis der Sprachtätigkeit früherer Generationen ist. Da wir aber wissen, dass Sprache sich ständig verändert und wir Mühe hätten, heute Menschen ohne Schwierigkeit zu verstehen, die vor achthundert oder tausend Jahren gelebt haben, ergibt sich wiederum die neue Frage, warum bei der Weitergabe von Sprache über Generationen tiefgreifende Veränderungen zu beobachten sind.

Ich habe diesen Zusammenhang herausgestellt und in den Vordergrund gerückt, weil es meiner Ansicht nach gute Gründe gibt, die Entstehung der Sprache allein aus dem Diskurs in Zweifel zu ziehen, was aber nicht so zu verstehen ist, dass das Kind seine Sprache vollständig eigenständig ausbildet.

Die verschiedenen Varianten der Diskurstheorien werfen weitere Fragen auf:

- 1.) Da nicht geklärt ist, woher die sprachlichen Elemente stammen, deren sich der Diskurs bedient, werden sie als vorhanden vorausgesetzt. Damit kommt man zu einem Perpetuum-Mobile-Modell, das die Erklärung von Veränderung innerhalb eines geschlossenen Systems anstrebt, ohne von außen zugeführte Energie.
- 2.) Die in der historischen Sprachbeschreibung augenfällige Sprachveränderung wird durch die Annahme von Prinzipien begründet, die subjektive oder objektive Optimierungsstrategien voraussetzen, z.B. durch die Bewertung von Sprachqualitäten.
- 3.) Eine Interpretation der empirischen Sprachbeobachtungen erfolgt ausschließlich durch den Bezug auf durch die Kommunikation motivierte Begriffe, deren Definition aber wegen der Missachtung von individuellen Systemaspekten zu kurz greift. So wird z. B. der Begriff der "Ökonomie" in Bezug auf die Bedürfnisse des Hörers und des Sprechers definiert, indem postuliert wird, der Sprecher artikuliere gerade so 'detailliert' bzw. 'aufwendig' wie es das Verständnis des Hörers erfordere (vgl. Lindblom 1990). Die Eingrenzung des Ökonomie-Begriffs wird aber nicht in Bezug auf das Sprachsystem eines einzelnen Sprachteilnehmers vorgenommen.

2.1.1 Tomasellos (2006) Evolutionstheorie des Menschen als Beispiel einer Diskurstheorie und das Erkennen von Kausalität durch den Menschen

Tomasello (2006) versucht in Auseinandersetzung mit den kognitiven Fähigkeiten nicht-menschlicher Primaten, eine Erklärung für die evolutionäre Entwicklung der kognitiven Fähigkeiten des Menschen zu finden. Es gibt große kognitive Übereinstimmungen, weshalb Tomasello (2006: 253) auch davon spricht, dass die "[m]enschliche Kognition [...] eigentlich eine Form der Primatenkognition" (Tomasello 2006: 253) sei. Auch hinsichtlich der Gene hat der Mensch mit seinem nächsten evolutionären Verwandten, dem Schimpansen, 99% gemeinsam (Tomasello 2006: 14). Trotz dieser umfassenden kognitiven und genetischen Gemeinsamkeiten habe der Mensch einen Umfang an Kultur vorzuweisen, insbesondere eine Sprache und technische Innovationen, wie sie bei Schimpansen oder anderen nicht-menschlichen Primaten nicht anzutreffen sei (Tomasello 2006).

Wie sind die kulturellen Unterschiede zwischen dem Menschen und den nichtmenschlichen Primaten evolutionär zu erklären?

Die Beantwortung der Frage wird nach Tomasello (2006) durch den vergleichsweise kurzen Zeitraum seit der evolutionären Trennung der menschlichen und der nicht-menschlichen Primaten vor 6 Millionen Jahren noch schwieriger, zumal wesentliche Hinweise auf Unterschiede sich vor allem auf die letzten 250.000 Jahre bezögen (Tomasello 2006: 15). Tomasello (2006: 28ff., 40ff.) weist auf die Intelligenz von nicht-menschlichen Primaten hin, die z.B. zu technischen Innovationen wie einer neuen Art des Termitenangelns führen würde (Tomasello 2006: 43), die aber nicht über Generationen erhalten blieben, weil es an einer effizienten und (nahezu) verlustfreien Form der Weitergabe von Informationen mangle (Tomasello 2006: 56). Nicht-menschliche Primaten erkennen zwar das Antezedenz-Konsequenz-Verhältnis von Ereignissen (Tomasello 2006: 36), dieses beziehe sich aber nur auf die äußere Erscheinungsform der Ereignisse (Tomasello 2006: 31ff.). Tomasello (2006: 31) spricht ihnen aber die kognitive Fähigkeit ab, über Antezedenz-Konsequenz-Verhältnisse in Beziehung stehende Ereignisse mittels des abstrakten Konstrukts einer kausalen Kraft zu verbinden, die beinhalte, dass mehrere äußere Ereignisse über dieselbe vermittelnde Kraft zu einem äußerlich beobachtbaren Ereignis führen könnten (Tomasello 2006: 38). Diese kognitive Fähigkeit, äußere Ereignis-

se zueinander kausal in Beziehung zu setzen, sieht Tomasello nur² beim Menschen gegeben (Tomasello 2006: 38). Evolutionär-historisch habe der Mensch zunächst Absichten anderer Artgenossen als Verursacher ihrer Handlungen erkannt, d.h. über die sog. "intentionale Kognition" (Tomasello 2006: 36) insbesondere die Handlungen anderer als zielgerichtet aufgefasst (Tomasello 2006: 38). Die intentionale Kognition habe der Mensch dann vom sozialen Bereich auf die Erkenntnis anderer nicht durch geistige Zustände verursachter Ereignisse auch in der unbelebten Natur übertragen und damit die "kausale Kognition" (Tomasello 2006: 36) entwickelt (Tomasello 2006: 38). Tomasello (2006: 18) nimmt an, dass die intentionale Kognition beim heutigen Menschen in Form der "biologische[n] Vererbung" (Tomasello 2006: 27) weitergegeben wird und damit angeboren ist (ebd.) und evolutionär "eine artspezifische kognitive Anpassung" (Tomasello 2006: 253) sei. Trotz ihrer Bezeichnung als evolutionäre Adaption, bleibt zu erklären, warum die intentionale Kognition als neuronal angeborenes "Modul" (Tomasello 2006: 76) unvermittelt in die Gene übernommen worden sein soll. Die fehlende logische Motivierung³ dieses Moduls der intentionalen Kognition spricht Tomasello (2006) auch an:

Die evolutionären Bedingungen, unter denen sich diese einzigartige sozio-kognitive Fähigkeit entwickelte, sind gegenwärtig unbekannt, aber eine mögliche Hypothese besagt, daß sie erst mit dem modernen Homo sapiens auftrat und daß sie wirklich das kognitive Hauptmerkmal ist, das moderne von prämodernen Menschen unterscheidet. (Tomasello 2006: 74)⁴.

² Die kognitive Fähigkeit, kausale Zusammenhänge zu erfassen, scheint im Tierreich weiter verbreitet zu sein und widerspräche damit der Annahme Tomasellos (2006), die kausale Kognition als eine artspezifische Fähigkeit des Menschen zu betrachten. So beobachteten Taylor et al. (2012) Neukaledonische Krähen, die mit einem Hilfsmittel Futter aus einer Röhre holten und sich dabei von einer kleinen Öffnung in einer Wand abwenden mussten, durch die sie zuvor die Bewegung eines durch diese Öffnung reichenden Stockes gesehen hatten (ebd.). Die Krähen verhielten sich unterschiedlich, je nachdem ob sie vor der Bewegung des Stockes einen Menschen hinter der Wand hatten verschwinden sehen, der anschließend wieder erschien und den Käfig verließ oder ob dieses Ereignis nicht stattfand: Im ersten Fall war die Anzahl der Überprüfungen des Bereichs hinter der Wand ("Versteck") größer bzw. die Anzahl der Abbrüche des Versuchs der Nahrungsaufnahme geringer. Taylor et al. (2012) schließen aus diesem Verhalten, dass Krähen die Bewegung des Stockes der hinter der Wand verschwundenen Person zuordneten (ebd.) und damit kausale Zusammenhänge verstehen können (ebd.). In der alternativen Situation hingegen, in der keine Person der Bewegung des Stockes zuzuordnen war, verhielten sich die Tiere vorsichtiger. Taylor et al. (2012) machen auch darauf aufmerksam, dass ihre Untersuchung die erste sei, die Hinweise auf eine kausale Kognition bei Tieren gebe (ebd.: 1) und dass weitere Studien anderer Arten für die Umwelt der Tiere angemessene "ecological situations" (ebd.: 1) schaffen müssten. Taylor et al. (2012) schließen zudem reines Assoziationslernen für das Verhalten der Krähen aus, weil dieses nicht erklären könne, warum die Krähen sich in beiden Situationen unterschiedlich verhielten, obwohl der Stimulus sich nicht änderte (ebd.).

³ Die Annahme eines angeborenen, artspezifischen Moduls der intentionalen Kognition wird durch die Einbeziehung seiner Rolle im Spracherwerb ebenso wenig begründet (s.u.).

⁴ Im Verständnis Tomasellos (2006) reduziert sich die evolutionäre Entwicklung des Menschen auf das unmotivierte Entstehen der intentionalen Kognition:

"Die vollständige Reihe angenommener Evolutionereignisse sieht folgendermaßen aus: Menschen entwickelten eine neue Form sozialer Kognition, die bestimmte neue Formen kulturellen Lernens er-

In seiner Theorie nimmt diese intentionale Kognition jedoch den zentralen Platz ein, denn sie soll sowohl historisch als auch ontogenetisch die kulturellen Unterschiede zwischen Menschen und nichtmenschlichen Primaten und insbesondere die Entstehung der Sprache erklären. Denn durch diese (im Verlauf der Evolution unvermittelt genetisch inkorporierte) intentionale Kognition sollen neue Formen des sozialen Lernens wie Imitation, Unterricht und Zusammenarbeit (Tomasello 2006: 16f.) ermöglicht worden sein. Sie sollen mit dem Beginn der genetischen Fixierung dieser kognitiven Fähigkeit die effiziente, weil verlustfreie Weitergabe bereits erlangter Informationen an die nächsten Generationen ermöglicht haben (Tomasello 2006: 16, 54ff.). Auf diese Weise erklärt Tomasello (2006: 16, 54ff.) die sukzessive wachsende, immer größere kulturelle Wissensanhäufung der Menschheit, weshalb er auch von "*kumulativer* [Hervorhebung Verfasser] kultureller Evolution" (ebd.) spricht und diesen Vorgang mit dem Betätigen eines Wagenhebers vergleicht (2006: 16, 54): Analog zu diesem technischen Mechanismus verhinder(te)n die neuen spezifischen Lernformen des Menschen einen Rückfall auf eine frühere Wissens- bzw. Informationsebene (Tomasello 2006: 16, 54ff.), weshalb Tomasello (2006) diesen Vorgang auch als "Wagenhebereffekt" bezeichnet (Tomasello 2006: 16, 54ff.). Die informationswahrende Funktion des neuronalen Moduls der intentionalen Kognition ist auch in Tomasellos (2006: 71) Konzeption des Spracherwerbs relevant (Tomasello 2006: 71ff.). Kinder müssen möglichst effizient und verlustfrei einen Großteil der bereits angehäuften Informationen der Menschheit "verinnerlichen" (Tomasello 2006: 163ff.). Dabei ist ihnen die angeborene intentionale Kognition besonders hilfreich, weil sie die entsprechenden notwendigen Lernprozesse wie die Imitation ermögliche: In "Szenen gemeinsamer Aufmerksamkeit" (Tomasello 2006: 127ff.) können Kinder dank der angeborenen intentionalen Kognition ihr Verständnis von sich selbst als intentionale Wesen analog auf den Erwachsenen übertragen und deshalb seine Handlungen als zielgerichtet wahrnehmen. Diese intentionale Kognition ermöglicht den Kindern nach Tomasello (2006) auch das Lernen sprachlicher Symbole. Sprachliche Symbole entsprächen dabei unterschiedlichen Perspektiven auf Ereignisse oder Objekte, die die Kognition bzw. die kognitiven Repräsentationen der Kinder durch diese unterschiedlichen Blickrichtungen veränderten (Tomasello 2006: 126). Es geht in Tomasellos (2006) Konzeption des Spracherwerbs also

möglichte, wodurch bestimmte neue Prozesse der Soziogenese und eine kumulative kulturelle Evolution möglich wurden." (Tomasello 2006: 18)

vorwiegend um eine Übernahme und "Verinnerlichung" (Tomasello 2006: 163ff.) von Wissen, weshalb Tomasello (2006: 72) auch davon spricht, "daß Kinder in der Zeit zwischen einem und drei Jahren insofern »Imitationsmaschinen« sind, als ihre natürliche Reaktion auf viele Situationen darin besteht, daß sie das tun, was die anderen in ihrer Umgebung tun. Tatsächlich ist das, was sie jeweils für sich selbst in den meisten Situationen zustande bringen, sehr begrenzt." (Tomasello 2006: 72). Zwar gesteht Tomasello (2006: 73) den Kindern auch eine stärkere Eigenkomponente "im Alter von vier bis fünf Jahren" (ebd.) zu, aber dies ändert nichts an seiner (Tomasello 2006) Konzeption des frühen Spracherwerbs als effizienter und möglichst verlustfreier Wissensübertragung von den Erwachsenen auf das Kind, welche evolutionär durch die genetische Fixierung der intentionalen Kognition zum kulturellen Vorsprung des Menschen vor allen anderen nicht-menschlichen Primaten geführt haben soll (s.o.).

In Tomasellos (2006) Überlegungen bleibt ungeklärt, wie die intentionale Kognition sich evolutionär entwickeln konnte, sie wird stattdessen vorausgesetzt. Eine intentionale Kognition im Sinne Tomasellos (2006) ist sicher Bestandteil des menschlichen Erkenntnisprozesses, sie kann aber nicht als Vorbedingung jeglicher Erkenntnis und Sprachentwicklung angesehen werden. Sprache und kulturelles Wissen bedingen sich zwar gegenseitig, sie sind aber nicht identisch: Die Weitergabe von Wissen setzt sprachliche Kategorien und Lernen voraus. Sprache und Denken bedingen sich gegenseitig und entwickeln sich auf der Grundlage eines praktischen Erkenntnisinteresses. Da die intentionale Kognition von anderen sprachlich basierten kognitiven Prozessen getrennt wird, verbleibt sie im Bereich der Empfindungen, wodurch der Übergang zur Sprachentwicklung und der mit ihr einhergehenden Entwicklung der Kognition offen bleiben. Die Fokussierung auf die "Weitergabe von Wissen" vernachlässigt den Aspekt der Entwicklung neuen Wissens als Bestandteil dieses praktischen Erkenntnisinteresses, das auch in einer kognitiven Theorie angesprochen werden sollte. Da Kinder als Objekte eines sprachlichen Lernprozesses angesehen werden, bleibt die Eigenständigkeit ihrer Leistung unberücksichtigt. Demgegenüber zeigt aber gerade die Neuschaffung von Sprachsystemen durch Kinder in Form der Kreolsprachen (Spitzer 2000: 200), dass der Erwerb einer Sprache mit der Entwicklung der Kognition einhergeht. Deshalb ist die Konzeption des Spracherwerbs primär als Wissensübernahme durch Sprachimitation irreführend, weil sie den Eindruck erweckt, Kinder würden lediglich Sprachhülsen übernehmen, die ihre geistigen Repräsentationen bzw. Fähigkeiten weiter veränderten.

2.1.2. Piaget

In ähnlicher Weise wie Tomasellos “Wagenhebereffekt” zu einer kumulativen kulturellen Evolution führen soll, sieht Piaget (in Montada: 1970) die menschliche Entwicklung als Äquilibrationsprozess (ebd.: 22) “als ein Fortschreiten von Zuständen geringeren Gleichgewichts, geringerer Stabilität zu Zuständen größeren Gleichgewichts, größerer Stabilität” (Montada 1970: 22). Dabei spielen die Prozesse der Assimilation und Akkommodation eine wesentliche Rolle (ebd.). "Nur das kann gesehen, begriffen und manipuliert werden, wofür Assimilationsschemata bestehen. Wer nicht zählen kann, kann keine Anzahl erfassen; wer nicht über den Begriff der Horizontalität verfügt, 'sieht' nicht, daß der Wasserspiegel horizontal bleibt, wenn die Lage eines Gefäßes verändert wird; wer nicht über Schemata der Klassenverschachtelung verfügt, erkennt nicht, daß die Oberklasse immer mehr Elemente umfaßt als die Unterklasse. Es ist nicht so, daß sich die "objektive" Umwelt einem passiven Geist als Information aufdrängt, im Gegenteil, ein aktiver Geist assimiliert sich die Umwelt, und er assimiliert das, was für ihn aufgrund seines Erkenntnis- und Verhaltensinstrumentariums assimilierbar ist" (ebd.: 22). Die Assimilationschemata übernehmen in dieser Theorie die analoge Funktion der “intentionalen Kognition” Tomasellos.

Auch Piagets Annahme unterscheidet sich vollständig von dem in dieser Arbeit vertretenen Ansatz, da im letzteren keinerlei Regeln oder Schemata zur Voraussetzung einer Entwicklung gemacht werden. Die Entwicklung des Silben- und Phonemsystems ist ein unbewusst verlaufender Prozess, dessen adaptive Orientierung auf die Zielsprache zwar perzeptiv entsprechend des Systemzustandes begrenzt ist, jedoch nicht auf die Existenz oder das Fehlen von “Regeln” zurückzuführen ist, sondern auf Verteilungsverschiebungen und -veränderungen innerhalb des Systems. Aus diesem Grunde ist auch die von Piaget betonte “Akkommodation” (in Montada 1970: 22) als Gegenentwicklung gegen eine unzureichende Assimilation oder gegen widersprüchliche Assimilationsschemata (ebd.) mit der hier vertretenen evolutionären Linguistik nicht vergleichbar, die aber auch nicht als voraussetzungslos verstanden werden soll. Das eigentliche Problem liegt in der Frage, wie genetische Prägungen sich mit der individuellen Entwicklung unter dem Einfluss der Umwelt in Beziehung setzen lassen⁵. Piaget (1974) stimmt mit Einschränkungen

⁵ Die Entwicklung des Silbensystems, wie sie in dieser Arbeit angenommen wird, könnte u. U. mit dem Konzept Piagets der Akkommodation und Assimilation verwechselt werden, da in beiden theoretischen Ansätzen der Gesichtspunkt der Anpassung eines Systems eine Rolle spielt. Im Gegensatz zu Piaget ist die

der Theorie Lamarcks zu, da "die Aquisition neuer Eigenschaften [...] für ihn im wesentlichen durch die prägenden Einflüsse der Außenwelt bestimmt [war], vor allem durch die Art, wie die Lebewesen dadurch ihre Gewohnheiten modifizierten" (Piaget 1974: 85). Er sieht jedoch diese Hypothese als zu einseitig an, "weil sie Einflüsse vernachlässigt, die sich für das Verständnis der Beziehungen zwischen dem Organismus und seiner Umgebung als wesentlich erwiesen haben" (ebd.) und weist (ebd.) auf Parallelen zwischen Lamarck und dem Behaviorismus hin, dessen Schema des "stimulus-response" er durch ein Schema ersetzen möchte (ebd.), in dem zwischen "stimulus" (ebd.: 86) und "response" (ebd.) die "Assimilation des Stimulus an ein gewisses Reaktionsschema" (ebd.) eingeschoben wird (ebd.), "das die Antwort auslöst" (ebd.). "Damit der Stimulus eine Reaktion auslöst, muß das Subjekt bzw. der Organismus überhaupt dazu fähig sein, und die Frage ist somit die, wie der Organismus diese Fähigkeit erlangt, die auf dem Gebiet der Embryonalentwicklung von Waddington als <<Kompetenz>> bezeichnet wurde und die sich als Empfindlichkeit gegenüber <<Induktoren>> definiert." (ebd.) Da die lebenden Organismen unterschiedliche genetische Voraussetzungen haben und deshalb auch ein Schimpanse weder Opernsänger noch Klaviervirtuose werden kann, liegt die Schlussfolgerung einer weitgehenden genetischen Determinierung nahe, die diese Fähigkeiten bestimmt. Auf den Spracherwerb bezogen kann diese Annahme jedoch für sich weder die phylogenetische Veränderung des Menschen noch die Differenzen und Differenzierungen auf der ontogenetischen Ebene erklären. Schemata oder Regeln sind nicht die Voraussetzung einer Entwicklung, auch wenn sie wie von Piaget über die Prozesse der Assimilation und Akkomodation als veränderbar gedacht werden, sondern ihre Folge⁶. Es ist ein wesentlicher Unterschied, ob die kognitive Entwicklung als eine Folge von Änderungen von Schemata bestimmt wird, wobei die früheren in die folgenden assimiliert werden, oder ob angenommen wird, dass jedes kognitive, lautliche etc. „Ereignis“ durch die Veränderung von Schnittmengen zu einer Verschiebung innerhalb eines Systems führt. Deshalb wird in der hier vertretenen Theorie angenommen, dass es das Prinzip des Lebens- und Reproduktionswillens ist, das die Entwicklung des Menschen vorantreibt, wobei sich Struktu-

evolutionäre Phonologie nicht teleologisch und bedarf auch nicht der Annahme von Regeln oder Schemata als Voraussetzung für Entwicklungsprozesse.

⁶ Vgl. dazu: "[...] die elementarsten Formen der Gewohnheit stammen immer aus einer Assimilation der neuen Elemente an die früheren Schemata, im gegebenen Fall an die Reflexschemata. Wichtig ist jedoch einzusehen, daß die Ausdehnung des Reflexschemas durch Einverleibung eines neuen Elements die Entstehung eines höheren Schemas zur Folge hat (die Gewohnheit), dem sich das einfache Schema (der Reflex) integriert. Die Assimilation eines neuen Elements an ein früheres Schema impliziert also ihrerseits die Eingliederung dieses Schemas in ein anderes höherer Ordnung." (Piaget 1976: 86).

ren aus Schnittmengenbildungen ergeben. Diese Annahme hat den Vorteil, dass alle Regeln a priori überflüssig werden. Piaget hat allerdings auch nicht den Anspruch eine linguistische Theorie zu entwickeln, sondern eine psychologisch ausgerichtete Theorie der Entwicklung der menschlichen Urteilslogik in Anknüpfung an Kants Logik (Bugge 1997: 14ff.). Sprachentwicklung mit der Entwicklung des Denkens und einer Urteilslogik gleichzusetzen, übersieht auch die Bedeutung der den Urteilen vorangehenden Prozesse in der Sprachentwicklung, die beim Kinde in der Produktion von Wörtern oder "Satzwörtern" bestehen, da in diesen ersten Wörtern wahrscheinlich verschiedene satzwertige Intentionen gebündelt sind. Ihre Produktion durch das Kind ist in dem Sinne mit logischen Urteilen gleichzusetzen, dass alleine durch ihre Existenz vom Kind eine ursprüngliche kausale Beziehung zwischen sich und der Außenwelt - Dingen, anderen Personen, dem eigenen Körper etc. - hergestellt wird. Die Urteilslogik des Kindes kann nur seinem Entwicklungsstand entsprechen. Deshalb ist sie angemessen. Eine Unterscheidung zwischen "egozentrischer" (Piaget 1972: 21ff.) Sprache des Kindes in den ersten Lebensjahren und "sozialisierter" (ebd.) Sprache in der weiteren Entwicklung, wie sie Piaget (1972) vornimmt, ist statisch und übersieht den "egozentrischen" (ebd.) Charakter jeder Sprache, da sie in ihrer Form und Logik einer speziellen Ontogenese entsprechen muss und deshalb immer adäquat ist. Die Illusion, die dieser Unterscheidung zu Grunde liegt, hat ihre Ursache in der geringeren Schnittmenge zwischen Kindersprache und Erwachsenensprache, die zur Postulierung einer höher bewerteten Sprache und Logik führt. Indem Piaget (1976) das "anschauliche Denken" (ebd.: 180) des Kindes als "entstellenden Egozentrismus" (ebd.) wertet, den er einem Denken gegenüberstellt, das "zu einem objektiven System dezentriert wird" (ebd.), übersieht er die Relativität dieser Attribute und Bewertungen, die in einer systembezogenen Analyse obsolet werden. Hier liegt der gleiche methodische Fehler wie in den Diskurstheorien vor, der zu einer Bewertung der Kindersprache auf der Basis der Zielsprache führt, ohne die primäre Kraft des kindersprachlichen Systems zu berücksichtigen gegenüber der sekundären Wirkung der Zielsprache. Die philosophischen und psychologischen Dimensionen der Fragestellung sollen in diesem Zusammenhang nicht weiter diskutiert werden, man könnte jedoch vom Standpunkt der evolutionären Phonologie aus die Frage stellen, ob sich nicht analog zur Entwicklung des Silben- und Phonemsystems die Entwicklung und Veränderung der individuellen und gesellschaftlichen Urteilslogik ebenfalls über ein Schnittmengenmodell erfassen und beschreiben ließe.

2.2. Phonetische Steuerung des Lauterwerbs und phonetische Begründung der Diskurstheorien - zum Verhältnis von Phonologie und Phonetik

Es ist bereits darüber gesprochen worden, dass die Einführung des Diskurses als Input des Spracherwerbs und der Sprachveränderung verschiedene Folgen für die Sprachtheorie hat, wie die Annahme von Optimierungsstrategien als Ursache von Sprachveränderungen sowie die Vorstellung des Spracherwerbs als Lernprozess.

Eine mit der Gruppe der Diskurstheorien verwandte Vorstellung wird von der Theorie der Natürlichen Phonologie (Stampe 1979) formuliert. Die Darstellung folgt der knappen Zusammenfassung in Wurzel (2001: 13ff.). Siehe aber auch Stampe (1979). Sie nimmt an, dass die "phonologische Komponente [...] innerhalb des Sprachsystems die Funktion [hat], die sprachlichen Äußerungen an die Beschaffenheit der menschlichen Sprachorgane anzupassen, d.h. phonologische (artikulatorisch-perzeptiv) zweckmäßige, mit anderen Worten: natürliche sprachliche Formen zu erzeugen."(Wurzel 2001: 19) Der Gegensatz zwischen den als angeboren angenommenen "phonologischen Prozessen"(ebd.), die als unbegrenztes Potential phonologischer Kombinationen verstanden werden und die allen Kindern in gleicher Weise zur Verfügung stehen, und der tatsächlich feststellbaren Begrenzung dieser Kombinationen wird mit der Unterdrückung von phonologischen Prozessen im Verlauf des Spracherwerbs und der Adaption der in einer Sprache geltenden "phonologischen Regeln"(ebd.: 20) erklärt (ebd.: 19f.). Die notwendige Unterdrückung von phonologischen Prozessen führe aber zu "Belastungen der Sprachkapazität" (Wurzel 2001: 20), während die "Nichtunterdrückung" (ebd.) mit dem Abbau der Belastungen gleichzusetzen sei (ebd.). Phonologische Regeln, die auf phonologischen Prozessen beruhen, werden als funktionale Optimierungsprozesse verstanden, um "Belastungen der Sprachkapazität zu vermeiden" (Wurzel 2001: 19) und die "perzeptiven Charakteristika"(ebd.) zu maximieren (ebd.) und die "artikulatorischen Schwierigkeiten"(ebd.) zu minimieren (ebd.). Die Dualität von Phonologie und Phonetik wird jedoch aufgehoben durch die Annahme der Determinierung der phonologischen Regeln durch die Beschaffenheit der Sprachorgane (Wurzel 2001: 19). Das Problem unterschiedlicher und auch widersprüchlicher phonologischer Prozesse wird in dieser Theorie nur konstatiert, kann aber nicht eindeutig geklärt werden: "Aus dem Vorhandensein von Widersprüchen zwischen phonologischen Prozessen folgt, dass es keine in eine einheitliche Richtung weisende undifferenzierte phonologische Optimierungstendenz geben kann"(Wurzel

2001: 20) und ebenso wenig eine "unter allen Gesichtspunkten optimale phonologische Struktur"(ebd.), was als Patchwork von Optimierungstendenzen aufgefasst werden kann. Die Annahme einer absoluten, phonetisch basierten Schwierigkeitshierarchie von Lauten und Lautkombinationen steht u. a. in Verbindung mit der Beobachtung, dass Kinder zunächst scheinbar "einfache" Laute produzieren, die im Verlauf der Entwicklung durch "schwierige" ersetzt werden. Diesem Sachverhalt entspricht in der Natürlichkeitstheorie die Annahme phonologischer Prozesse die zu Gunsten phonologischer Regeln unterdrückt werden, was im Grunde der Umkehrung der Vorzeichen bei einer Zahl entspricht: Das Modell dieser Arbeit betont mit einem "Plus" den produktiven Aspekt der Lautentwicklung, während die Natürlichkeitstheorie die "unvollkommene" und "fehlerhafte" Lautproduktion des Kindes als Subtraktion der unerwünschten phonologischen Prozesse (Regeln) ansieht. Dadurch wird der Lauterwerb zu einem pädagogischen Lernproblem, und außerdem wird indirekt die Steuerung des Lauterwerbs in den artikulatorischen Bereich verlegt und damit zu einer phonetischen Steuerung.

Dagegen wird in dem in dieser Arbeit vertretenen Ansatz die unterschiedliche Qualität und Verfügbarkeit von Lauten nicht übersehen, aber sie wird in einen Entwicklungszusammenhang gebracht, der mit Prinzipien der Evolution vereinbar ist und auf pädagogische Kategorien verzichten kann. "Einfache" Laute werden in der Sprachentwicklung früh produziert. Es soll in Anlehnung an Gedanken von MacNeilage (1998: 503f.; 2008: 91ff.) angenommen werden, dass die ersten von Kindern produzierten Laute eine enge Beziehung zu solchen Lauten haben, die unmittelbarer Ausdruck von elementaren menschlichen Lebensäußerungen sind, also Lauten, die beim Essen, Trinken an der Mutterbrust, Schreien, Weinen, Lautäußerungen des Wohlbefindens, Gähnen, "Rufen" etc. entstehen. Diese Lautäußerungen decken ein Spektrum von Gefühlen und ersten kognitiven Erkenntnissen ab und sind von ihrem Umfang her vermutlich mit entsprechenden Signalkorpora in der Tierwelt vergleichbar. Können diese Laute deshalb als "einfach" bezeichnet werden?

Hier wird vorgeschlagen, den Begriff der "Einfachheit von Lauten" durch die Begriffe der onto- und phylogenetischen "Initiallaute" sowie "Folgelaute" zu ersetzen, die den Ausgangspunkt einer dynamischen Entwicklung und ihre Fortführung bilden. Die Beobachtungen des frühen Auftretens bestimmter Laute oder deren bevorzugte Kombinationen in Silben beruhen zwar auf der besonderen physiologischen Gestalt der menschlichen Artikulationsorgane und würden anders klingen, müsste der Mensch auf einer Blechgießkanne blasen, und ihre weite Verbreitung in vielen Sprachen der Welt ist auf die Gemein-

samkeit der oben angeführten Lebensäußerungen und Artikulationsbasis zurückzuführen, die sich als dominierende Schnittmengen präsentieren. Deshalb ist es nicht die "Einfachheit" von Lauten, die sie für die ersten Phasen der Sprachentwicklung prädestinieren, sondern ihr ontogenetischer Funktionszusammenhang. Der hier diskutierte Begriff der "Einfachheit" kann auch nicht mit dem in dieser Arbeit verwendeten Begriff "Ökonomie" oder "ökonomisch" in Verbindung gebracht werden. Der letztere bezieht sich auf die Nutzung des begrenzten physiologischen Raumes durch eine Abstandsregelung, die die Eindeutigkeit der erzeugten Signale sichert. In einem sich ständig in Veränderung befindlichen Silben- und Lautsystem gibt es deshalb keinerlei absolute Rangordnung von Lautqualitäten, sondern relative zeitweilige Systemverhältnisse.

Auch die in den Präferenzgesetzen (Vennemann 1988) postulierte Universalität muss im Spannungsverhältnis der universalen menschlichen Voraussetzungen der Sprachorgane zu den individuellen Sprachsystemen jedes einzelnen Menschen gesehen werden. Die große Ähnlichkeit der Sprachorgane aller Menschen bedingt entsprechende Lautbildungspotentiale und daraus resultierende Schnittmengen bei den gebildeten Lauten. Es ist ebenfalls davon auszugehen, dass es physiologisch bevorzugte Kombinationsmöglichkeiten von Lauten gibt, wie sie in der Forschung und insbesondere in den Präferenzgesetzen (Vennemann 1988) ja auch nachgewiesen worden sind. Die Annahme jedoch, Lautwandel sei phonetisch motiviert, verallgemeinert quantitativ auffällige Phänomene, ohne sie im Gesamtzusammenhang der Lautentwicklung und des Lautwandels zu sehen und ignoriert die geistige Seite der Sprache, die sich in der Entwicklung des Silbensystems und des auf ihm beruhenden phonologischen Systems erweisen. Die Feststellung von "Präferenzgesetzen" (Vennemann 1988) der Silbenphonologie ist deshalb an sich nicht zu bestreiten, aber sie können die eigentliche Ursache der Sprachentstehung und -veränderung nicht erklären, sondern nur partiell auffällige Systembewegungen abbilden.

Geht man hingegen davon aus, dass phonetisch motivierter Lautwandel die primäre Veränderung darstellt und somit sowohl die Ursache für eine gestörte Balance innerhalb des Lautsystems ist wie auch das Heilmittel, um die Funktionalität der Sprache wieder herzustellen, muss die eigentliche Ursache für diese Annahmen aus dem Sprachbereich in die Psychologie verlagert und den Sprechern/Hörern sich widersprechende Optimierungstendenzen unterstellt werden oder es müssen universell geltende Regelwerke angenommen werden. Die Fokussierung auf phonetische Regelmäßigkeiten erschwert die Beurteilung des Zusammenhangs zwischen phonetischen "Präferenzen" und der Ursache ihrer Steuerung. Die Annahme einer phonetischen Steuerung der Lautveränderung ist nicht nur frag-

lich, weil sie die evolutionäre Entwicklung und Veränderung des Menschen nicht beachtet, denn es ist auch von einer Veränderung der Artikulationsorgane im Laufe der Zeit auszugehen, die folglich keine Konstanten sind. Wiederum ist zu fragen, worin der Anstoß zu ihrer Veränderung zu sehen ist, und da der Mensch sich selbst verändert, indem er real und kognitiv von der Welt Besitz ergreift - in einem viel umfassenderen Sinne, als dies in der Tierwelt zu sehen ist - kann die Steuerung des kognitiven und des durch ihn bedingten sprachlichen Prozesses nicht in einem isolierten Organ gesucht werden. Es ist auch nicht die Hand, die zur Systematisierung und Erklärung menschlicher Bewegungen dienen kann, sondern ihre Bewegungen sind aus den Intentionen des Menschen selbst abzuleiten. Die Existenz phonetischer Regelmäßigkeiten wird also nicht bestritten, jedoch die Annahme, sie seien ursächlicher Motor von Veränderung. Dieser hingegen ist in der praktischen und kognitiven Auseinandersetzung des Menschen mit der Welt und damit auch mit sich selbst zu suchen, die über die Ausdehnung des Lexikons und Verschiebungen innerhalb des Silben- und damit Phonemsystems das gesamte Spannungsgefüge des Lautsystems in einem latenten Gleichgewicht halten zwischen der tendenziellen Begrenzung des artikulatorischen Raumes, der seine „ökonomische“ Nutzung fordert, das heißt eine möglichst dichte maximale Belegung des verfügbaren Raumes wegen der begrenzten Ressourcen, bei gleichzeitiger minimaler Differenzierung, das heißt eine möglichst geringe Belegung des verfügbaren Raumes zur Sicherung der Eindeutigkeit der Signale. In diesem Kontext können Präferenzen für die Existenz von Phonemen sowie für die Existenz von bevorzugten Phonemkombinationen als Anzeichen ökonomischer - auf die intensive Nutzung des verfügbaren Raumes gerichtete Prozesse gesehen werden. Die Annahme der Präferenzgesetze (Vennemann 1988) impliziert aber auch, dass die Frage der Differenzierung nicht gestellt wird. Entgegengesetzte oder widersprüchliche phonetische Regeln oder Präferenzen, die in einer vorwiegend phonetisch motivierten Theorie nur subjektiv psychologisch gedeutet werden können, erhalten hingegen ihren Sinn, wenn man sie im Kontext eines Spannungsverhältnisses zwischen maximaler ökonomischer Nutzung des begrenzten Raumes und seiner minimalen Differenzierung zur Sicherung der Eindeutigkeit sieht. In einem begrenzten Raum mit einem geringen Belegungsdruck können die Abstände größer sein; wenn der Belegungsdruck wächst, müssen die Abstände verringert und durch die Bildung von Kombinationsgruppen wiederum ausreichende Systemabstände gesichert werden. Deshalb steht die Überlegung in dieser Arbeit, dass bei der Besetzung der Kopf- und Kodakonstituenten in den Silben in den beiden in dieser Arbeit berücksichtigten Korpora zunächst die Möglichkeiten der Besetzung mit einfachen

Konsonanten ausgeschöpft werden, bevor Komplexe Verwendung finden, auch nicht im Gegensatz zu Beobachtungen in anderen Korpora und den regelmäßigen Entwicklungen in anderen Sprachen, da eine systembedingte Differenzierungsforderung auch abweichende Verteilungen ermöglichen kann. Aus dieser Vorstellung heraus lassen sich jedoch auch andere Phänomene schlüssig erklären, ohne eine Steuerung durch phonetische Regeln annehmen zu müssen.

Bei der Entwicklung der kindlichen Sprache und seines Lautsystems steht naturgemäß der Aspekt des Aufbaus eines Systems im Vordergrund. Da der phonetische Raum und sein phonologisches Pendant jedoch begrenzt sind, was sich in der Gefahr einer zunehmenden Unschärfe und fehlenden Eindeutigkeit der Signale im Systemzusammenhang auswirken muss, ist in allen Sprachen die Tendenz einer Gegenbewegung gegen die zunehmende Ausweitung von Phonembildungen oder Phonemgruppierungsbildungen zu beobachten, u. a. die "Vokalisierung von /r/ in der Koda, wodurch z.B. in Varietäten des Deutschen /ar/ mit /a:/ zusammenfällt und *warte* und *wate*, *Fahrt* und *fad*, *Bart* und *Bad/bat*, *Carter* und *Kater* usw. homophon werden" (Vennemann, persönliches Schreiben 2013). Da anzunehmen ist, dass es eine Silbenhöchstgrenze innerhalb eines variablen Bandbreitenbereichs gibt mit einer zuzuordnenden Phonemhöchstgrenze, erfolgt in dem über den kausalen Motor motivierten und sich über die Perzeption selbst steuernden Silben- und Phonemsystem eine ständige Veränderung, die nicht nur als Umgruppierung innerhalb der bestehenden Menge von Konstituenten zu verstehen ist, sondern auch als begrenzte Neubildungen, wie sie die vorstehenden Beispiele illustrieren. Ihnen ist die signaltechnische Redundanz der wegfallenden Phoneme gemeinsam, die diese phonologische Systemänderung ermöglicht, obwohl sie in Widerspruch zu bestehenden Einträgen im Lexikon geraten können. Die durch eine zulässige Silbenhöchstmenge bedingten Verschiebungen und Veränderungen innerhalb des Silbenkorpus und des damit über die Schnittmengenbildung in Verbindung stehenden Phonemsystems, die zur Sicherung der Eindeutigkeit ständig erfolgen, bilden morphologische Sedimentationen im Silbenkorpus und Lexikon, die wiederum die zukünftige Entwicklung der Korpora auf verschiedenen Sprachebenen beeinflussen.

3. Das Phonemerwerbsmodell dieser Arbeit

In diesem Kapitel soll das Phonemerwerbsmodell dieser Arbeit im Zusammenhang dargestellt werden. Im darauffolgenden Kapitel werden einzelne Merkmale und sich aus ihnen ergebende Implikationen genauer erläutert.

Die Grundfrage der Sprachwissenschaft seit der Antike⁷ ist die nach dem Ursprung der Sprache, ob die Sprachfähigkeit des Menschen angeboren (von Gott) oder erworben ist. Da inzwischen die evolutionäre Entwicklung des Menschen nicht mehr bestritten wird, müsste der evolutionäre Gesichtspunkt auch auf alle Eigenschaften, Fähigkeiten, Gestaltveränderungen des Menschen wie auch seinen Geist und seine Sprache anzuwenden sein. Auch die Sprachwissenschaft des 19. Jahrhunderts hat keine Antwort auf diese Grundfrage geben können und hat sie schließlich sogar verboten (Christiansen, Kirby 2003: 2; Neis 2010: 555). Doch eine Frage, deren Lösung unabdingbar ist, lässt sich nicht einfach abweisen, denn sie bricht sich auch unbewusst und ungewollt eine Bahn.

Obwohl z.B. die Entwicklung der Phonetik im 19. Jahrhundert seit Kempelen (1791) nicht mit der Frage nach dem Ursprung der Sprache verbunden war, hat sie, beginnend mit Kempelens (ebd.) Versuch, Sprachlaute künstlich zu reproduzieren, indirekt zur Lösung der Frage beigetragen. Die detaillierte Verknüpfung der Sprachphysiologie mit den gebildeten Sprachlauten rückte den sprechenden Menschen und seine Sprachorgane in den Mittelpunkt. Als später deutlich wurde, dass es keine unmittelbare Entsprechung zwischen physikalischen Signalen und den gehörten Sprachlauten gibt und sich ungefähr zeitgleich aus der Phonetik die phonologische Wissenschaft entwickelte, entstand ein Verständnis der Sprachlaute als geistige Entitäten ohne unmittelbare physikalische Entsprechung, deren Erzeugung auf Grund angeborener Regeln zu den Grundsätzen der lange einflussreichen Sprachtheorie Chomskys (Chomsky, Halle 1968) und seiner Anhänger gehörte.

Die Erforschung der Regelmäßigkeiten der Sprache, um ihre Entstehung, ihren Wandel und ihr Wirken zu verstehen, bedarf einer evolutionären Sichtweise, die konsequent danach fragt, wie es zur Entstehung und Veränderung von Strukturen kommen kann.

Diese Arbeit ist der Versuch, eine evolutionäre Perspektive anzuwenden und auf die Annahme angeborener Faktoren zu verzichten, weil Modelle, die von angeborenen Elementen ausgehen, in Gefahr geraten, die Frage nach dem Ursprung sprachlicher Prozesse aus dem Blick zu verlieren und sich in einer Zirkelschluss-Falle zu verfangen. Der Mensch hat sich im Rahmen der Evolution entwickelt, und die Annahme von angeborenen Merkmalen für ihn bedarf wiederum einer evolutionären Erklärung. Mit dem hier vorgestellten

⁷ Die Frage nach dem Ursprung der Sprache hat die Menschheit schon lange beschäftigt: So finden sich Abhandlungen des Themas in Legendenform in der Bibel (z.B. Turm von Babel)(Neis 2010: 552) und in der Antike sind Plato, Aristoteles und Epikur zu nennen, die sich mit dem Sprachursprung beschäftigten (ebd.: 552).

Phonemerwerbsmodell ist die Annahme verbunden, dass es in analoger Weise auch für die Evolution der Sprachen an sich angewandt werden kann. Dieses Modell versucht mit einem Minimum von als angeboren angenommenen, für die Sprache des Menschen als notwendig erachteten Konstrukten auszukommen und ist auf der Annahme gegründet, dass vor ihrer Postulierung eine genaue komparative Untersuchung zeigen sollte, ob beim Menschen beobachtete Eigenschaften nicht auch in der Tierwelt vorhanden sind.

Das Kind entwickelt sich geistig seit seiner Geburt oder sogar schon als Embryo, aus eigenem Interesse an seiner Umwelt. Es entwickelt sich aus eigener Kraft und die Entwicklung seiner Sprache ist auf seine eigene produktive Tätigkeit zurückzuführen und nicht auf die Instruktionen eines Lehrers, der ihm über Rückmeldungen von richtigem oder falschem Verhalten den Weg zur Sprache weist. Zunächst hört das Kind die akustischen, artikulierten Formen der Umgebung, dann artikuliert es während der Plapperphase selbst auf der Gefühlsebene. Jedoch gibt es etwa im Alter von 12 Monaten eine qualitative Veränderung bei der Verwendung der artikulierten Formen durch das Kind. Es erfährt die Artikulationen nicht mehr nur auf der Gefühlsebene, sondern verbindet diese mit entdeckten kausalen Zusammenhängen. Die erste solche Verknüpfung eines kausalen Zusammenhangs mit einer artikulierten akustischen Form stellt den Beginn des Spracherwerbs und die Ausbildung eines Lautsystems in der Form des Phonemerwerbs dar. Es wird angenommen, dass der Ursprung der Sprache in einer solchen Verbindung entstanden ist.

Sprache ist als Mittel der Kommunikation nicht nur ein Phänomen in der menschlichen Zirkulationssphäre analog zur Funktion des Geldes in der Wirtschaft, sondern ein Instrument und ein Begleiter der kognitiven Entwicklung jedes Menschen. Weil die Sprachentwicklung jedes einzelnen Menschen auf seinen individuellen Erfahrungen in seiner Auseinandersetzung mit der Natur, mit seiner sozialen Umwelt und mit sich selbst beruht, ist sie historisch, d.h. zeitbezogen und vor allem einmalig. Deshalb kann z.B. die Ausbildung von Creolesprachen, die voraussetzt, dass junge Menschen bis zum Alter von 12 Jahren in einer zweisprachigen Welt aufwachsen, auch als Wirkung der Doppelfunktion der Sprache für die Kognition und die Kommunikation gedeutet werden.

Die Annahme der kognitiven Funktion der Sprache ermöglicht es, die Sprachentwicklung und den Sprachwandel in seiner historischen Dynamik zu untersuchen und zu begreifen.

Die kognitive Funktion der Sprache, die ihren Ursprung in der Auseinandersetzung der Menschen mit der belebten und unbelebten Natur hat, ist die produktive Seite der Sprache, ihr Motor der Veränderung. Da auch in der Tierwelt diese Interaktion zwischen Sub-

jekten und der Natur stattfindet, kann man keinen generellen Unterschied zwischen Mensch und Tier, sondern nur einen graduellen annehmen. Auch Tiere kommunizieren für sie relevante Zusammenhänge durch Signale. So beispielsweise die Warnsignale vor Räubern. Der Mensch hat jedoch die Auseinandersetzung mit der belebten und unbelebten Natur vor allem durch die Anwendung von vielfältigen Werkzeugen immer stärker systematisiert, optimiert und perfektioniert, wobei er tief in kausale Zusammenhänge eingedrungen ist, die seine geistigen Fähigkeiten erweitert und die Notwendigkeit der gesellschaftlichen Kommunikation erhöht haben.

Nach der Realisierung einer initialen kausalen Relation in der ersten sprachlichen Lautform, die tatsächlich als Sprache gewertet werden kann, weil die beiden Aspekte der Kognition und der Kommunikation in ihr vereint sind, erforscht das Kind von nun an durch aktives Handeln kognitiv seine Umwelt, was ständig zur Entdeckung neuer kausaler Zusammenhänge führt. Dieser Vorgang soll als *kausaler Motor* bezeichnet werden, weil er als die Ursache und der Beginn des kindlichen Spracherwerbs und indirekt des Phonemerwerbs angesehen wird. Deshalb liegt in dem hier vorgestellten Modell der Schwerpunkt auf der produktiven Seite der Sprache und der Eigenständigkeit des Kindes.

Jeder einzelne Sprecher baut aus eigener Kraft, angetrieben durch den kausalen Motor sein eigenes Sprachsystem auf, das unterschiedlich große "Schnittmengen" auf verschiedenen sprachlichen Ebenen mit denen anderer Sprecher gemeinsam hat (vgl. Abb. 1.).

In dieser Arbeit wird allein die Formseite der vom Kind gebildeten Zeichen untersucht, weshalb im Folgenden statt vom Zeichen von der *Lautform* des Kindes gesprochen werden soll. Die Begrenzung auf die Formseite der vom Kind gebildeten Zeichen wird vorgenommen, weil es als unmöglich angesehen wird, zuverlässig die vom Kind mit einer Lautform intendierten Bedeutungen bzw. kausalen Zusammenhängen von einer Außenposition her zu erfassen. Denn die Bedeutungen eines Sprachteilnehmers bilden ein in sich kohärentes, mit der Lautformenseite seiner Zeichen interagierendes Bedeutungssystem, und anstatt dass unterschiedliche Sprachteilnehmer über eine Identität von Bedeutungen kommunizieren, erfolgt dieselbe lediglich über Bedeutungsschnittmengen, die aber von den Sprachteilnehmern oft als vollständige Identität angenommen werden (vgl. Abb. 1.). Weil es aus diesem Grund nahezu unmöglich ist, die genaue Bedeutung der Lautäußerungen eines Kindes zu bestimmen, wird mit der Grundannahme gearbeitet, dass alle vom Kind geäußerten Lautformen bedeutungsvoll sind und so Ausdruck der bereits geleisteten Arbeit des kausalen Motors. Der Begriff der *Lautform* für die Formseite des Zeichens soll darauf hinweisen, dass die Form des vom Kind verwendeten Zeichens arti-

kuliert wird und akustisch wahrnehmbar ist (und nicht etwa schriftlich vorliegt). Der Begriff impliziert zudem, dass die Lautformen vom Kind kognitiv zur Unterscheidung kausaler Zusammenhänge verwendet werden und sich auf diese Weise von nicht-sprachlichen akustischen Formen abgrenzen. Der Bestandteil *Laut* soll aber nicht bedeuten, dass das Kind seine Lautformen bereits aus Phonemen bildet oder dieselben einzeln für sich wahrnimmt.

Der kausale Motor wird auch als die treibende Kraft für die historische Entwicklung der Lautformenseite von Sprachen angesehen. Mit der zunehmenden Anzahl von durch den kausalen Motor geschaffenen kausalen Zusammenhängen steigt die Anzahl der erforderlichen Lautformen. Auch Tiere würden ihr Signalrepertoire mit der Vermehrung der für sie relevanten Zusammenhänge erhöhen.

Eine Reihe von Umständen spricht für die Annahme, dass das Kind zumindest in der Anfangszeit seine Lautformen als Silben produziert, perzipiert und speichert. Die Koartikulationsforschung hat nachgewiesen, dass die Silbe die stärkste Einheit der Koartikulation ist (Wester 2003: 233). Das bedeutet, dass Phoneme einer Silbe sich stärker gegenseitig beeinflussen als Phoneme verschiedener Silben. Das kann so verstanden werden, dass sie in gewisser Weise eine untrennbare Einheit bilden, wie es sich in der Vokalabhängigkeit der Konsonanten zeigt. Da die kindlichen Lautformen in den Anfangsphasen der Sprachentwicklung überwiegend in Silbenform oder in Reduplikationen vorliegen, wird dies ebenfalls als Argument verstanden, die Silbe zur Grundlage der Lautformenuntersuchung zu nehmen.

Der Bedarf an neuen Silben verlangt ihre hinreichende Eindeutigkeit und Unterscheidung. Es ist aber festzustellen, dass bei der Differenzierung der Silben ökonomische Prinzipien gelten, die eine maximale Wiederverwendung von Lautelementen fördern, soweit sie nicht der notwendigen Differenzierung im System widersprechen. Ein idealer Ausgleich zwischen den beiden widerstreitenden Funktionen der Ökonomie und der Unterscheidbarkeit von Silben wird in einem sich selbst über Abstände und Dichtezustände steuernden Prozess erreicht: Die bereits vorhandenen, detailliert und multimodal gespeicherten Silben bilden miteinander Schnittmengen, die vorläufig als identische Konstituenten von zwei Silben wie zum Beispiel das Phonem /b/ in den Silben /bam/ und /bit/ verstanden werden. Dieses Beispiel darf aber nicht so missverstanden werden, dass Vergleiche von Einzelsilben im Vordergrund stehen. Vielmehr geht es um die Schnittmengenverteilung in einem temporären Silbensystem. Es wird angenommen, dass diese un-

bewusste Selbststeuerung des Silbensystems über die Perzeption erfolgt, indem perzeptiv die eigene Lautproduktion und deren Wirken in der Kommunikation abgeglichen werden. Auf Grund der funktional bestimmten Bildung der Silben und den sich daraus ergebenden Schnittmengenrelationen zwischen ihnen, soll von einem *Silbensystem* gesprochen werden.

Alle natürlichen Systeme steuern sich selbst und sind wertneutral. Wenn wir als Beispiel einen Flusslauf nehmen, so wird er seinem Zweck, Wasser aus dem Binnenland in einen Binnensee oder ein Meer zu überführen, in jedem Fall gerecht. Dabei spielt es keine Rolle, ob er sich auf Grund der Bildung von Sandbänken in Nebenarme verzweigt, ob große Flächen durch ihn überflutet werden, ob er mäandert oder sich tief in ein Sandsteingebirge versenkt. Veränderungen in einem natürlichen Flusssystem können durch die Beobachtung regelmäßiger Veränderungstypen systematisiert und zu einem Prognoseinstrument entwickelt werden. Diese angenommenen regelhaften Veränderungstypen "steuern" den Prozess jedoch nicht, sondern es ist das Gesetz der Schwerkraft, das Wasser von einem höheren Punkt zu einem tieferen fließen lässt. Analog zu diesem Beispiel müssen auch individuelle Lautsysteme an sich und in ihrer Entwicklung als sich selbst steuernd und wertneutral gesehen werden. Sie erfüllen ihren Zweck, den begrenzten artikulatorischen und phonologischen Raum so zu gestalten und ständig zu ändern, dass bei maximaler Wiederholung und minimaler Differenzierung alle Silben und Wörter des Systems eindeutig sind. Gerät diese Eindeutigkeit zeitweilig durch ein Anwachsen des Lexikons in Konflikte mit der bestehenden Systemkonfiguration, wird eine Erweiterung und/oder Umstellung im System veranlasst. Dieser Prozess ist kein psychologischer oder teleologischer, weshalb er nicht als Optimierung verstanden oder überhaupt mit den Kategorien "gut" und "schlecht" oder ihren Ableitungen bewertet werden kann, obwohl dies nicht so verstanden werden soll, als wenn eine Beschreibung und Analyse regelmäßig auftretender Phänomene abgelehnt wird.

Für das Silbensystem werden neben dem Selbststeuerungsprozess eine Reihe weiterer Eigenschaften angenommen:

Analog dazu, wie in den verschiedenen Sprachen der Welt individuelle Lautsysteme entstanden sind mit jeweils individuellen Phonemverteilungen, wird eine individuelle systembedingte Phonemverteilung für jeden Sprecher einer Sprache angenommen, die zu verschiedenen Zeitpunkten nicht mit sich selbst identisch, sondern ständig in Veränderung begriffen ist.

Die nur im Systemzusammenhang verständlichen Relationen innerhalb eines Lautsystems stehen im Widerspruch zu ihrer Beschreibung mit den Symbolen des IPA, denn diese werden über physiologische Eigenschaften der Sprachproduktion definiert, die einen Systemzusammenhang nicht beachten und für universell gehalten werden. Dennoch lassen sich durch die Anwendung von Schnittmengenuntersuchungen von Silben, die unter Anwendung der IPA-Phonemsymbole aufgezeichnet wurden, Hinweise auf den eigenständigen Systemcharakter der Sprache und ihrer Entwicklung gewinnen.

Das Silbensystem ist adaptiv, indem ungefähre akustische Eindrücke der Zielsprache auf der Grundlage des eigenen Bildungspotentials für die Silbenbildung entlehnt werden. Dies ist möglich, weil von der Annahme einer wechselseitigen Beziehung zwischen Sprachperzeption und -produktion ausgegangen wird, wobei die Perzeption des Kindes, die in gewissem Maße durch seine eigenen Produktionsmöglichkeiten bestimmt und begrenzt wird, dennoch dieser Produktion leicht vorausseilt. Das Kind kann die "Laute einer fremden Sprache", der Zielsprache, nur auf der Basis seines eigenen Lautsystems "interpretieren", ebenso wie es jedem Sprecher einer Sprache ergeht, der mit vollkommen fremden Lauten einer fremden Sprache aus einem anderen Lautsystemzusammenhang in Kontakt kommt. Da das Kind diesen Lauteindrücken nicht nur kontemplativ gegenübersteht, sondern ihre Perzeption mit der Entwicklung seiner eigenen Kognition verbunden ist, wird darin der Grund für einen geringen Vorlauf der Perzeption gesehen. Dadurch entwickelt sich das Silbensystem des Kindes durch eine ständige Vergrößerung der Schnittmengen in Richtung der Zielsprache. Die im Silbensystem des Kindes objektiv nachzuweisenden Schnittmengen müssen einen kritischen Wert erreichen, damit sie dem Kind zu Beginn des Lesenlernens subjektiv in der Form von Einzellauten verfügbar werden können.

Die Individualität jedes Lautsystems beruht außerdem auf dem Ausgangspunkt seiner Entwicklung und den sich daraus ergebenden Entwicklungsschritten. Diese sind aber für zwei Sprecher nicht übereinstimmend, weil je nach dem Anfangszustand die weiteren Entwicklungen in gewisser Hinsicht determiniert sind, was aber nicht als Automatismus einer Entwicklung verstanden werden soll, sondern als Beeinflussung zukünftiger Entwicklungen durch das in der Vergangenheit Entstandene. Es ist ein Determinismus, wie man ihn auch für das Wachstum eines Baumes annehmen kann, der bei seinem Bemühen, zum Licht zu wachsen, in einer Auseinandersetzung mit seiner Umgebung und Umwelt

immer die für ihn möglichen und richtigen Schritte unternimmt hinsichtlich des Wachstums und der Richtung von Wurzeln und Ästen, die seine individuelle Gestalt prägen.

Eine Analogie zwischen der morphologischen Entwicklung eines Baumes und der sich ändernden Verteilungsstruktur eines individuellen Silbenkorpus und der sich aus ihm durch Schnittmengen ergebenden individuellen Phonemstruktur liegt in ihrer Determination. Diese wird nicht als absolute genetische Vorherbestimmung ihrer Gestalt oder Struktur verstanden, sondern als ständige spezifische Reaktion auf Herausforderungen unter der Voraussetzung bestehender Strukturen und Verteilungen: Zwar gibt es typische Erscheinungsbilder von Birken, Kiefern oder Eichen, die eine genetisch bedingte Einschränkung der Formentwicklung begründen: Aus einem Birkensamen kann sich keine Eiche entwickeln. Ihre Erscheinungsbilder verändern sich aber im Verlauf ihrer Genese und sie unterscheiden sich auch für die Vertreter bestimmter Spezies untereinander entsprechend ihres Standortes, ihres Nahrungs- und Wasserangebotes sowie des für sie bestimmenden Ökosystems.

Eine Birke, die ihre erste Entwicklung als Unterholz in einem Hochwald beginnt, hat eine andere Gestaltausprägung als ein allein wachsender Baum: Sie wächst in einem sehr schlanken Profil nach oben zu einer Lücke im Blätterdach, während der nachbarlose Baum von Beginn an auch in die Breite wächst. Ändern sich nach einer Reihe von Jahren die äußeren Bedingungen für beide Bäume, müssen sie ihre Strukturen anpassen.

Ebenso wie die Gestalt der Bäume ist auch die grundsätzliche Gestalt der Phonemkorpora genetisch durch die physiologischen Vorgaben der Artikulationsorgane begrenzt, obwohl diese Grenzen nicht immerwährend sind, sondern sich im Laufe der evolutionären Entwicklung verschieben.

Weil die Phonemsymbole auf individuelle Silbensysteme bezogen sind, die sich voneinander unterscheiden, und da sie als Silbenschnittmengen eines Lautsystems gedacht werden, soll davon gesprochen werden, dass Phoneme unterschiedlicher Sprecher nur Schnittmengen miteinander gemeinsam haben. Wie zu zeigen sein wird, lassen sich im Gegensatz zur Annahme eines zufälligen Phonemerwerbsprozesses Merkmale des Lautentwicklungsprozesses identifizieren, die man als *organisches Wachstum* bezeichnen kann. Dabei wird Bezug auf eine zweite Analogie zwischen Baum und Lautsystem genommen, die sich aus ihrer jeweiligen äußeren bzw. inneren Struktur ergibt. Die morphologischen Konstituenten Wurzelsystem, Stamm und Krone beim Baum sollen als eine gewisse Analogie zur Entwicklung der Silbenkonstituenten und ihres Systems gesehen werden. Dem Stamm entspricht die Silbenkonstituente des Silbenkerns, während Wurzel-

geflecht und Krone durch die Silbenränder repräsentiert werden. Diese Analogie ist wegen der vergleichbaren Eigenschaften der Konstituenten gewählt worden, da sowohl der Stamm wie der Silbenkern eher als Konstante in der Entwicklung dienen, während sich die Silbenränder und Wurzelsystem und Krone durch ihre individuelle Differenzierung auszeichnen.

4. Eigenschaften und Implikationen des Phonemerwerbsmodells

In diesem Kapitel werden einige Eigenschaften und Implikationen des Silbensystems näher erläutert und mit Beobachtungen und Modellen zu sprachlichem Verhalten verknüpft.

4.1. Der Kausale Motor

4.1.1. Das Erkennen kausaler Zusammenhänge durch das Kind

Untersuchungen zeigen, dass Kinder während des ersten Lebensjahres in zunehmendem Maße kausale Zusammenhänge erkennen können (Sommerville 2007: 49f.), wie z.B. die zielgerichtete Handlung von Personen, die nach Gegenständen greifen, was durch den Wunsch nach dem Besitz eines Gegenstandes ausgelöst wird (ebd.). Kinder beginnen aber auch physikalisch kausale Zusammenhänge zu verstehen, die in der Verwendung von indirekten Hilfsmitteln wie z.B. einer Harke zum Heranziehen eines zu weit entfernten Gegenstands bestehen (Meltzoff 2007: 44), nachdem ihnen die Verwendung des Gegenstandes demonstriert wurde. Dabei besteht ein Zusammenhang zwischen dem zunehmend höheren Alter der Kinder und ihrem Erfolg, das Hilfsmittel selbst korrekt zu verwenden. Der Interventionsansatz zur Kausalität (Woodward 2007) geht davon aus, dass Kinder die Ursachen kausaler Zusammenhänge durch aktive Handlung und Manipulation der entsprechenden Größen entdecken (Woodward 2007: 20), weil sich mit der Ursache auch ihre Wirkung verändern sollte (ebd.). Dabei können aber auch reine Beobachtungen von Handlungen anderer Personen (Klein-)Kindern beim richtigen Gebrauch von Hilfsmitteln weiterhelfen (Meltzoff 2007), jedoch zieht Meltzoff (2007: 45) aus dem oben beschriebenen Harkengebrauch den Schluss, dass "[o]ne possibility that arises from this work is that infants have to be "on the cusp" of solving the problem themselves to get the boost from seeing how someone else solves it" (Meltzoff 2007: 45). Die Nachahmung von Handlungen anderer scheint also nur möglich zu sein, soweit die eigene praktische Betätigung des Kindes schon ausreichend vorangeschritten ist. Das Kind, dessen Erkenntnisinteresse von seinen Bedürfnissen geleitet wird, entdeckt also in der aktiven intervenierenden Auseinandersetzung mit seiner Umwelt kausale Zusammenhänge.

Dass dieses Erkenntnisinteresse den Spracherwerb antreibt, könnte sich im Zusammenhang zwischen den Fähigkeiten des Gebrauchs von Hilfsmitteln ("means-end"-Beziehungen) durch Kinder und ihrem sprachlichen Fortschritt zeigen: So korrelierte der Umfang der Fähigkeiten beim Verwenden von Hilfsmitteln für bestimmte Zwecke von Kindern im Alter von 4, 8 und 12 Monaten mit dem Umfang ihres passiven und aktiven Vokabulars im Alter von 2 Jahren (Siegel 1981 in Blake 2000: 159) und unterscheidet sich für Kinder im Alter von 8 Monaten, die im Alter von 3 Jahren als sprachliche Spätentwickler bezeichnet werden können, von dem für Kinder mit normaler sprachlicher Entwicklung ebenfalls im Alter von 8 Monaten (Miller, Siegel 1989 in Blake 2000: 159). Dieser Zusammenhang kann so interpretiert werden, dass das durch die eigenen Bedürfnisse geleitete Erkenntnisinteresse und die damit verbundene Entdeckung kausaler Zusammenhänge die Sprachentwicklung antreibt und insbesondere den Lexikonumfang vergrößert, der wiederum einen Einfluss auf den Umfang des Silbensystems und das Phonemsystem hat. Die Abhängigkeit des Phonemsystems vom Lexikon⁸ ist schon vielfach beobachtet worden (Stoel-Gammon 2010: 16).

4.1.2. Der Kausale Motor, die Sprachentstehung und -entwicklung

Sowohl menschengeschichtlich als auch ontogenetisch wird das Entdecken kausaler Zusammenhänge und deren lautliche Artikulation als die Ursache der Entwicklung der menschlichen Sprache angesehen. Die Entdeckung kausaler Zusammenhänge ist dabei nicht beliebig, sondern wird von einem Erkenntnisinteresse geleitet, das auf die Befriedi-

⁸ Mit der Aussage der Abhängigkeit des Silben- und damit des Phonemsystems ist nicht eine Untersuchung der Äußerung *einzelner* Zielsprachlautformen gemeint (siehe: Diskurstheorien) und etwa daraus abgeleitete Gründe für die in den Diskurstheorien als falsch angenommenen Äußerungen des Kindes wie etwa die in Elsen (1994) formulierte Frage, ob eine unzureichende Konzeptentwicklung oder artikulatorische Schwierigkeiten für die vom Beobachter als Übergeneralisierung wahrgenommenen Abweichungen bei der Äußerung von Zielsprachwörtern durch das Kind eine Rolle spielen (ebd.).

Vielmehr beinhaltet die Konstatierung dieser Abhängigkeit Unterschiede hinsichtlich der globalen Verteilungslösungen von Konstituenten im Silbensystem, die in dieser Arbeit in Abhängigkeit vom Lexikonumfang gesehen werden (vgl. den Begriff des kausalen Motors): So sind beispielsweise 5 Bedeutungen, die das Kind durch Lautformen kodiert mit weniger Silben verbunden als 500 Bedeutungen. Entsprechend weist das Silben- (und damit das Phonem)system für diese beiden Lexikonumfänge ganz unterschiedliche Verteilungen auf.

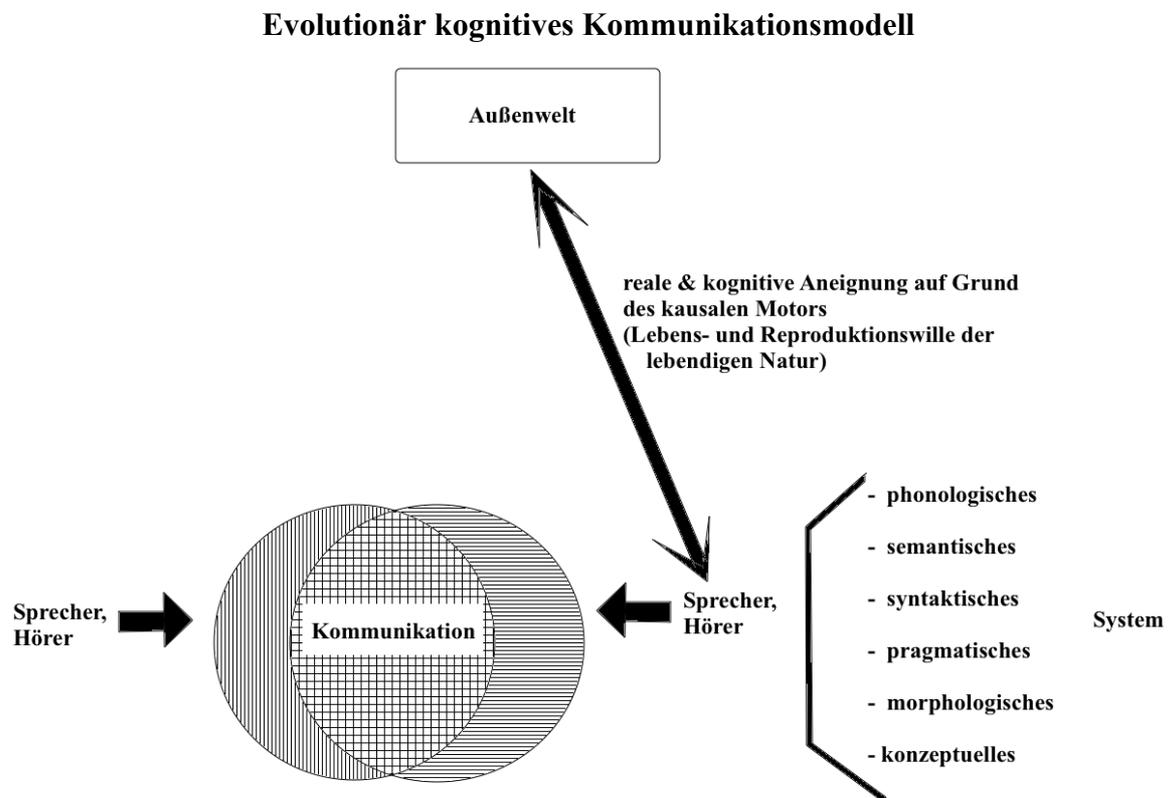
Diese Sicht wird durch das Postulat III von Stoel-Gammon (2010: 15) unterstützt, dass die Phonementwicklung und die Entwicklung des Lexikons einander angemessen ("commensurate" (ebd.)) sind, welches als Verallgemeinerung der Beobachtungen anzusehen ist, dass "children with large vocabularies have more advanced phonological systems than those with small vocabularies." (ebd.). Einige der von Stoel-Gammon (2010) genannten Unterschiede betreffen u.a. "[the] size and nature of the consonantal repertoire"(Stoel-Gammon 2010: 16) und können als Indiz für die auf die in dieser Arbeit angenommene Abhängigkeit des Silben- (und damit Phonemsystems) vom Lexikonumfang angesehen werden, die sich in Vergrößerungen der Silbenschnittmengen und der Art ihrer Verteilung zeigt.

gung menschlicher Bedürfnisse ausgerichtet ist. Der Begriff *kausaler Motor* verweist auf ein an Bedürfnissen orientiertes Erkenntnisinteresse, das zur Realisierung kausaler Zusammenhänge führt, die in sprachlichen Lautformen materialisiert werden.

Dieses Erkenntnisinteresse ist der primäre Motor der Sprach- und Denkentwicklung.

Wenn etwas entsteht, muss von irgendwoher ein Anstoß kommen. In diesem Sinne muss die phylogenetische Entstehung der Sprache auf eine Ursache zurückgeführt werden, die die Menschen von den Tieren unterscheidet. Es gibt Gründe für die Annahme, dass Sprache nicht nur eine kommunikative Funktion hat, sondern sich in Verbindung mit der kognitiven Entwicklung herausgebildet hat und sich verändert. Zweitens soll angenommen werden, dass die Kognition nicht nur ein Prozess ist, der im Gehirn vorhanden ist und sich im Gehirn abspielt - natürlich tut er das - sondern ein Prozess, der die Verbindung zwischen der äußeren Welt und dem Menschen herstellt und sein Weg ist, wie er sich die äußere Welt, zu der auch sein Ich gehört, aneignet. Der kognitive Prozess, der hier als "kausaler Motor" bezeichnet wird, materialisiert sich in der Sprache und führt zu ihrer Ausbildung und Veränderung. Der Vorteil dieser Hypothese ist, dass sie nicht auf Perpetuum-Mobile-Modelle, Zirkelschluss-Modelle sowie angeborene Eigenschaften oder Strukturen angewiesen ist, sondern dass sich sowohl die grundlegenden Elemente der Sprache wie auch die Ausbildung abstrakter Strukturen in Einklang mit Erkenntnissen der Evolution bringen lassen. Geht man davon aus, dass der Lebens- und Reproduktionswille der lebendigen Natur das Motiv der Evolution ist, so ist die kognitive Funktion der Sprache, also der Kausale Motor, der besondere Weg der geistigen Aneignung der Welt in Verbindung mit einer materiellen Auseinandersetzung mit dieser Welt.

Abbildung 1: Evolutionär kognitives Kommunikationsmodell



Diese Hypothese kann auf jede Annahme einer Bewusstheit oder bewussten Steuerung des Prozesses verzichten, womit natürlich nicht in Frage gestellt werden soll, dass Sprache selbst zum Gegenstand einer geistigen Auseinandersetzung oder des Versuchs von Beeinflussungen werden kann. Sie betont nur, dass im überwiegenden Maße Sprachentwicklung und -veränderung unbewusst verlaufen und sich sogar einer bewussten Beeinflussung entziehen. Wenn man weder mit einer bewussten Steuerung der Sprachentwicklung arbeiten will noch von der Annahme einer unbewussten Steuerung durch vorhandene Regelsysteme im Gehirn ausgehen will, stellt sich natürlich die Frage, wie die Sprachentstehung und -entwicklung unbewusst gesteuert wird?

Die Funktion der sprachlichen Signale wurde über den Kausalen Motor erklärt. Zur Eigenschaft jedes Signals gehört, dass es eindeutig sein muss. Diese Annahme wird auch auf die sprachlichen Signale übertragen, d.h. jedes einzelne sprachliche Signal muss sich durch eine besondere Eigenschaft von allen anderen unterscheiden, damit es sowohl für den einzelnen Menschen wie auch in der Kommunikation verwendbar wird. Kein Signal

kann mit einem anderen vollkommen identisch sein, ohne seine spezielle Aufgabe zu verlieren.

Wenn man nicht annehmen will, dass diese Lautäußerungen unvollkommene Reproduktionen der es beeinflussenden Zielsprache sind, muss die Frage beantwortet werden, wie es zur Ausbildung dieser unterschiedlichen Lautformen kommt. Zur Beantwortung dieser Frage soll von einer weiteren Hypothese ausgegangen werden. Es wird angenommen, dass der unbewusste Prozess der Sprachentwicklung sich zwar rezeptiv an der Zielsprache orientiert, dies aber auf der Grundlage seiner eigenen sich ständig verändernden Systemvoraussetzungen tut, zu deren Prinzipien neben der Notwendigkeit der Eindeutigkeit der Signale ihre Bildung in einem ökonomischen Zusammenhang gehören. Bei Schweden ist es ein beliebter Spaß Deutschen gegenüber, sie zu bitten, einmal das Wort für deutsch *Krankenschwester* auszusprechen. Als Deutscher interpretiert man diese Lautfolge auf der Grundlage des eigenen phonologischen Systems, in dem der schwedische postalveolare Frikativ /ɧ/ nicht enthalten ist. Sie ersetzen dieses Phonem durch den ihnen bekannten postalveolaren Frikativ, das "sch". Analog dazu kann angenommen werden, dass das Kind den Input der Zielsprache jeweils auf der Basis eines spezifischen Systemzustandes perzipiert. Daraus ergibt sich eine jeweils zeitweilige Begrenzung sowohl der Perzeptions- wie auch der Produktionsmöglichkeiten.

Eine weitere Einschränkung ergibt sich zwangsläufig aus der Begrenztheit des phonetischen Raumes, die den Rahmen bildet für seine ökonomische Nutzung. In allen Sprachen der Welt ist die Zahl der Vokale und Konsonanten sehr begrenzt und übersichtlich. Es soll angenommen werden, dass diese Mengenverhältnisse etwas mit der ökonomischen Ausnutzung eines begrenzten Raumes zu tun haben, wobei die entstehenden Laute sich nicht zu ähnlich sein dürfen, um ihre Eindeutigkeit zu wahren. Dann könnten die Phonemsysteme als jeweils besondere Verteilungslösung von Lauten in einem individuellen System betrachtet werden. Die Aufteilung eines Raumes ist eine geometrische Frage. Die Verteilung von Punkten in einem begrenzten Raum muss durch Abstände geregelt werden. Eine unbewusste Steuerung des Lauterwerbs kann deshalb als Gegensatz zwischen den beiden Polen der Eindeutigkeit der Signale und ihrer ökonomischen Erzeugung in einem begrenzten Raum gesehen werden.

Es soll davon ausgegangen werden, dass das Kleinkind rezeptiv Laute aus der Umgebungssprache identifiziert, soweit es zu ihrer Bildung selbst in der Lage ist. Da aber der Bedarf des Kindes an Signalen beschleunigt wächst, würde eine absolute Gleichsetzung von Rezeptions- und Produktionsfähigkeiten die Veränderung in der kindlichen Sprache

nicht erklären können. Deshalb wird angenommen, dass die Rezeptionsfähigkeiten des Kindes seinen Produktionsfähigkeiten ein wenig vorausgehen. Damit wäre auch erklärt, warum Kinder sich in eine bestimmte Sprache hineinentwickeln, und zwar auf der Grundlage ihrer eigenen sprachlichen Systemvoraussetzungen, die jeweils vollkommen einmalig sind und keinem anderen Sprachsystem eines anderen Menschen gleichen. Diese Hypothesen ermöglichen es, neben der Annahme der Individualität jedes Prozesses auch auf den semantischen Bezug zur Zielsprache zu verzichten und damit die kindliche Sprache als deformierte Variante der Zielsprache zu begreifen, die sich aus der unvollkommenen Anwendung angeborener Regeln ergibt. Stattdessen wird wiederum der Schwerpunkt auf die Autonomie der Systementwicklung gelegt mit der einzigen Einschränkung des oben erwähnten Rezeptionsvorlaufs, der partiell die Lautentwicklung steuert.

Ein an Bedürfnissen orientiertes Erkenntnisinteresse lässt sich auch für Tiere annehmen, da sie für ihre Lebenswelt relevante kausale Zusammenhänge mit entsprechenden Signalen wie z.B. die Warnrufe vor Räubern kommunizieren. Im Gegensatz zur Tierwelt hat sich der Mensch in zunehmendem Umfang mit kausalen Zusammenhängen v.a. durch die Herstellung von Werkzeugen vertraut gemacht, die für sein eigenes Denken und wegen der Relevanz der Werkzeuge für die Gruppe seiner Artgenossen einen sprachlichen Ausdruck verlangten. Die frühen Einsichten des Menschen in die mit der Werkzeugherstellung verbundenen kausalen Mittel-Zweck-Beziehungen zeigen sich auch in archäologischen Fundstücken von Werkzeugen, die in ihrer Form heutigen Werkzeugen entsprechen und damit einer bis jetzt nicht übertroffenen optimalen Form. Durch die Abstraktion der mit der Werkzeugherstellung verbundenen kausalen Zusammenhänge in sprachlichen Ausdrücken wurden diese kausalen Relationen wiederum für weitere kognitive Untersuchungen zugänglich. Auch die Zunahme der Gehirngröße des Menschen steht sicher im Zusammenhang mit der Entdeckung kausaler Zusammenhänge und der Sprachentwicklung und der damit einhergehenden Zunahme der Komplexität menschlicher Kognition.

Die Aktivität des Kindes nach seiner Geburt kann vor Allem als Vorbereitung der ersten kausalen Zusammenhänge verstanden werden, mit denen der Spracherwerb beginnt. Denn der initiale kausale Zusammenhang wird in einer bedeutungsvollen Lautform abstrahiert und für die Kommunikation und das Denken des Kindes zugänglich gemacht. Die weitere Entwicklung der Sprache des Kindes wird ab dem Zeitpunkt der initialen kausalen Relation, von der Entdeckung weiterer interessegeleiteter kausaler Zusammenhänge vorangetrieben und ebenso das Ausmaß seines Phonem- und Silbensystems. Aktivitäten, die das Kind beim Entdecken kausaler Zusammenhänge fördern, sollten auch sei-

ne sprachliche Entwicklung beschleunigen. Vor diesem Hintergrund kann auch die von Tomasello (2006: 144) aufgeführte Untersuchung gesehen werden:

Sie [Carpenter, Nagell und Tomasello 1998 in ebd.] stellten fest, daß Säuglinge, die im Alter von zwölf Monaten längere Zeit mit ihren Müttern bei Beschäftigungen gemeinsamer Aufmerksamkeit verbrachten, in diesem Alter und in den folgenden Monaten mehr Sprache verstanden und verwendeten. Außerdem stellten sie fest, daß Kinder von Müttern, die den Aufmerksamkeitsfokus ihrer zwölf Monate alten Kinder mit Worten begleiteten, in den folgenden Monaten ein größeres Vokabular beim Verstehen hatten (wobei die Auswirkungen auf die Sprachproduktion etwas später auftraten). (Tomasello 2006: 144)

Denn als ursächlich für den schnelleren Spracherwerb der Kinder könnte statt der von Tomasello (ebd.) als angeboren angenommen in Szenen der gemeinsamen Aufmerksamkeit wirkenden intentionalen Kognition, die sicherlich mit längeren Szenen der gemeinsamen Aufmerksamkeit von Säuglingen und Erwachsenen einhergehende verstärkte Entdeckung kausaler Zusammenhänge angesehen werden, die bei Kindern mit geringerer Interaktion wahrscheinlich geringer ausgebildet ist.

Der kausale Motor als treibende Kraft der Sprachentwicklung ist an verschiedene Voraussetzungen geknüpft. Zwar wird angenommen, dass es zwischen den menschlichen und tierischen Signalsystemen keinen prinzipiellen Unterschied gibt, jedoch hat sich der Mensch im Laufe seiner phylogenetischen Entwicklung sehr stark verändert, was sich in seiner Gestalt und in der Form und Größe seines Gehirns ausdrückt sowie der Form und Funktionsweise seiner Artikulationsorgane. Deshalb sind die kognitiven Fähigkeiten von Primaten zwar denen des Menschen in Ansätzen vergleichbar, aber es ist nicht zu erwarten, dass sie innerhalb weniger Generationen an die Fähigkeiten des Menschen anschließen könnten. Die Sprachentwicklung ist auch an soziale Beziehungen geknüpft, die zwar nicht ihre eigentliche Ursache sind, aber eine notwendige Voraussetzung, da Signalsysteme nur in Wechselbeziehungen eine Funktion erhalten. Damit sind sie nicht unwichtig, aber sie sind nicht die Hauptursache der Sprachentstehung und -entwicklung, die im Individuum gesucht werden muss und nicht in sozialen Beziehungen. Ebenso wie ein Baum sich aus einem Samen entwickelt, wozu ihn sein Lebens- und Reproduktionswille antreibt, und entsprechend der äußeren Bedingungen eine einmalige Gestalt annimmt auf der Grundlage seiner genetischen Voraussetzungen, ist die Sprachentwicklung auch an genetische Voraussetzungen gebunden, die sich in physiologischen Gegebenheiten ausprägen. Die für die Sprachentstehung wichtigste ist die Fähigkeit des Gehirns, nach ungefähr 12 Monaten Lebenszeit wiederholte Erfahrungen, die als Schnittmengen relevant

werden, spezifischen Lauten zuzuordnen und kommunikativ zu nutzen und kognitiv für den Ausbau eines logischen Systems. Die genetische Prägung des Baumes wie der menschlichen Sprachentwicklung sind die Keime einer Entwicklung, die sich nicht als "Klonen" einer vorgegebenen Struktur vollziehen, sondern immer als besondere Ausprägung einer besonderen Entwicklung, die deshalb auch auf die genetische Prägung zurückwirken und diese verändern. Der kausale Motor wird deshalb aus den genannten Gründen als Hauptursache der Sprachentstehung und -veränderung angenommen, aber nicht als alleinige Ursache, da er auf physiologische, kognitive und soziale Voraussetzungen angewiesen ist. Umgekehrt lässt sich jedoch feststellen, dass ohne die Annahme des kausalen Motors Sprache nicht alleine aus der Kommunikation entstehen kann noch aus der Physiologie.

4.2. Das Schnittmengenkonzept

In diesem Abschnitt soll das Konzept der Schnittmenge hinsichtlich des Silbensystems eines einzelnen Sprechers erläutert (4.2.1.) und in Bezug auf die Silbensysteme unterschiedlicher Sprecher (4.2.2.) betrachtet werden.

4.2.1. Schnittmengen im Silbensystem eines einzelnen Sprechers

Die Schnittmengen sind als Ergebnis und Indiz der sich selbststeuernden funktionalen Silbenbildung eines einzelnen Sprechers zu sehen, weil die Größe der Schnittmengen über eine Abstandsregelung festgelegt wird, die einen Ausgleich zwischen der Ökonomie ihrer Erzeugung und der Unterscheidbarkeit der gebildeten Silben schafft.

Die *Schnittmengen* von Silben werden in dieser Arbeit vorläufig als Übereinstimmungen in denselben Silbenkonstituenten angesehen. So hätten beispielsweise die Silben /gig/ und /gat/ hinsichtlich des Kopfes /g/ eine Schnittmenge. Es handelt sich aber um keine Identität, weil /g/ in unterschiedlichem Silbenkontext vorkommt. Die Schnittmenge wird dabei aber nicht als vorwiegend paarweiser Silbenvergleich verstanden, sondern als die Betrachtung der systemspezifischen Verteilungen der Schnittmengen im Gesamtzusammenhang.

4.2.1.1. Die funktionale Selbststeuerung der Silbenbildung und die damit verbundene Schnittmengenbildung als ein von Beginn des Phonemerwerbs an wirkender Prozess

Es gibt Annahmen (Stoel-Gammon 2010: 17), dass die Speicherungen der Wörter von Kindern sich während des Phonemerwerbs verändern (ebd.). Diese schließen aber nicht aus bzw. nehmen ausdrücklich an, dass anfänglich die kindlichen Äußerungen nur als holistische Einheiten gespeichert werden, d.h. nicht weiter zerteilt in Untereinheiten, und dass sich erst nach dem Erwerb einer bestimmten Anzahl von beispielsweise 50-100 Zielsprachwörtern eine Umstrukturierung der Speicherung der Wörter hin zu Segmenten einstellt (ebd.). Im Gegensatz dazu wird in dem hier vertretenen Konzept der Silbenschnittmengenbildung innerhalb eines organischen Wachstumsprozesses angenommen, dass es keine voneinander getrennten Phasen von strukturell unterschiedlichen Silbenspeicherungen gibt, in dem Sinne, dass zunächst Silben als Ganzheiten gespeichert werden und erst nach Erreichen einer ausreichenden Anzahl von Zielsprachwörtern durch das Kind eine Umstrukturierung der Speicherungen stattfindet, sondern es wird eine kontinuierliche Speicherung der Silben als Ganzheiten angenommen, aus denen sich über Schnittmengenbildung ein Lautverständnis (Phoneme, Grapheme) entwickelt. Dieser Speicherungs- und Schnittmengenbildungsprozess ist in quantitativer Hinsicht kontinuierlich, indem jede neue Silbe sich von allen anderen unterscheidet, während Schnittmengen zwischen diesen kontrastierenden Silben zu einer allmählichen Addition von Gemeinsamkeiten führen. Beispielsweise ist die Produktion und die akustische Ausprägung eines akustisch aufgezeichneten Phonems wie des Konsonanten /d/ vor dem Vokal /i/ eine andere als vor dem Vokal /o/ (Johnson, Mullenix 1997: 2), so dass die mit dem Phonem /d/ verbundenen Korrelate Unterschiede und Gemeinsamkeiten aufweisen, die zum einen zu einer Differenzierung der Silben /di/ und /do/, zum anderen aber zu einer quantitativen Verstärkung der /d/-Anteile führen.

4.2.1.2. Substanzbezogenheit der Schnittmengenbildung

Der Hinweis auf die Substanzbezogenheit der Schnittmengenbildung erfolgt v.a. vor dem Hintergrund, dass vielfach abstrakte Silbenstrukturen unabhängig von den die Silbe in abstrakten Transkriptionen bildenden Phonemen untersucht werden. So stellen beispielsweise Levelt. et al. (2000) die zwei von ihnen als universal angenommenen Erwerbseihenfolgen CV-CVC-V-VC-CVCC-VCC-CCV-CCVC-CCVCC und

CV-CVC-V-VC-CCV-CCVC-CVCC-VCC-CCVCC auf, ohne dabei zu beachten, dass die Entwicklung der abstrakten Silbenstrukturen nicht losgelöst von konkreten Einzelphonemen betrachtet werden kann, was dadurch bestätigt wird, dass viele der von Levelt et al. (2000) sehr spät aufgeführten abstrakten Silbenstrukturen auch schon sehr früh auftreten können (Lindner 2009: 171). Die Schnittmengenbildung erfolgt also nicht bezogen auf abstrakte Silbenpositionen, sondern auf das Vorhandensein konkreter messbarer Nuancen von Phonemen in unterschiedlichen Silbenkontexten wie beispielsweise den unterschiedlichen Ausprägungen des Phonems /d/ als Kopf der Silben /di/ und /do/ (Johnson, Mullenix 1997: 2). Diese Substanzbezogenheit der Schnittmengenbildung hat zur Folge, dass man von unterschiedlichen Graden der Schnittmengenbildung für unterschiedliche Phoneme zu einem Zeitpunkt für ein Kind sprechen muss. Es ist zu erwarten, dass Phoneme mit einem unterschiedlichen Ausmaß an Schnittmengenbildung zu Unterschieden im Sprachverhalten der Kinder führen. Dies könnte sich beispielsweise in der Präzision von Lautformen äußern, die nicht im Lexikon eines Kindes vorkommen⁹, aber über einen der beiden im folgenden Abschnitt vorgestellten Parameter der Nachbarschaftsdichte und der phonotaktischen Wahrscheinlichkeit sichtbar werden.

4.2.1.3. Hinweise auf die Schnittmengenbildung ohne direkte Einbeziehung des Bewusstseins

Neben der im folgenden Abschnitt beschriebenen phonematischen Bewusstheit, die bewusste Manipulationen von Teilen von Lautformen durch einen Sprachteilnehmer beinhaltet, gibt es auch Hinweise auf eine zunehmende Schnittmengenbildung, die indirekt aus dem Verhalten des Kindes ohne Anforderungen an seine bewusste Aufmerksamkeit abgeleitet werden können. Diese Hinweise sollen in den beiden folgenden Unterabschnitten in Form der Nachbarschaftsdichte und der phonotaktischen Wahrscheinlichkeit eines Wortes vorgestellt werden. Diese Sprachverhaltensparameter werden normalerweise auf das Wort bezogen, sollen hier aber als Ausdruck der Schnittmengenbildung im Silbensystem aufgefasst werden¹⁰.

⁹ auch *non words* genannt

¹⁰ In Abschnitt 5.2. wird eine Lexikon-Theorie vorgestellt, die von der Speicherung von Wörtern in Form von invarianten Phonemfolgen ausgeht. In diesem Zusammenhang wird auch das Verhältnis des Lexikons zur Silbe thematisiert.

4.2.1.3.1. Die Nachbarschaftsdichte eines Wortes

Ursprünglich hat Luce (1986) sein *Neighbourhood Activation Model (NAM)* entwickelt, um Beobachtungen über das Verhalten Erwachsener bei der Erkennung bzw. Identifikation gehörter Wörter zu erklären. Grundlegendes Konzept seines Modells ist die Nachbarschaft eines Wortes im Lexikon, wobei ein Wort ein *Nachbar* eines anderen Wortes im Lexikon ist, wenn es aus demselben durch Hinzufügen, Weglassen oder Ersetzen eines Phonems hervorgeht (Baese-Berk, Goldrick 2009: 529, Scarborough 2004: 21, Vitevitch 2002: 736). Z.B. sind die Wörter *scat* /skæt/, *at* /æt/, *hat* /hæt/, *cut* /kʌt/ und *cap* /kæp/ Nachbarn des Wortes *cat* /kæt/ (Vitevitch 2002: 736). Die Anzahl der Nachbarn eines bestimmten Wortes des Lexikons wird dann als die *Nachbarschaftsdichte*¹¹ dieses Wortes bezeichnet (Baese-Berk, Goldrick 2009: 529). Von einer hohen Nachbarschaftsdichte (*dense neighborhood*) spricht man, wenn ein Stimuluswort sehr viele Nachbarn hat, dagegen von einer geringen Nachbarschaftsdichte (*sparse neighborhood*), wenn das entsprechende Wort im Vergleich zu anderen Wörtern des Lexikons sehr wenige Nachbarn hat. Die Nachbarschaften von Wörtern können als Nachbarschaften von in den Wörtern enthaltenen Silben aufgefasst werden¹². Die Silbennachbarschaften sind wiederum Ausdruck der Wirkung des sich funktional selbst-steuernden Silbenbildungsprozesses, wobei die Verteilung der Schnittmengen den Ausgleich zwischen der Ökonomie und der Unterscheidbarkeit der produzierten Silben widerspiegelt. Die im Folgenden aufgeführten Nachbarschaftsdichteeffekte können dann als Hinweis auf die sich entwickelnde Schnittmengenverteilung aufgefasst werden: Auf der Grundlage des Konzepts der Nachbarschaft sagt das NAM aus, dass die richtige Erkennung eines Wortes die erfolgreiche Unterscheidung desselben von seinen Nachbarn im mentalen Lexikon erfordert (Luce 1986: 38) und dass dieselbe umso besser möglich ist, je häufiger es im Vergleich zu seinen Nachbarn im Lexikon verwendet wird und je weniger Nachbarn es hat (Luce 1986: 36).

¹¹ Die quantitative Definition der Nachbarschaftsdichte erfordert die Herstellung eines Verhältnisses. Die Schwierigkeit liegt jedoch nicht darin, das quantitative Verhältnis von Nachbarschaftsmengen zu einem Bezugswort herzustellen, sondern in der Definition der Ähnlichkeit, die Nachbarschaften begründet.

¹² ,insofern die oben angegebenen Operationen zur Bestimmung der Nachbarschaft sich auf ganze Silbenkonstituenten beziehen und damit komplexe Konstituenten nicht weiter unterteilen. Für eine Weiterentwicklung des Begriffs der Schnittmenge, die auf den speziellen Status der Komplexe eingeht, sind für die Nachbarschaftsbestimmung auch phonembezogene Operationen denkbar.

Außerdem muss der Zusammenhang zwischen einem Konzept und den mit ihm in Beziehung stehenden Silben(folgen) genauer geklärt werden. Denn für ein Konzept kann nicht nur eine einzige Silbenfolge geäußert werden, sondern die mit ihm verbundenen Silbensequenzen sind auch vom syntaktischen und pragmatischen Kontext abhängig.

Neben diesem Einfluss der Nachbarschaftsdichte auf die Worterkennung weisen auch andere Verhaltensweisen auf die nachbarschaftliche Organisation der Wörter im mentalen Lexikon von Erwachsenen hin: So wurde beispielsweise auch beobachtet, dass Wörter mit geringer Nachbarschaftsdichte einen kleineren Vokalraum aufweisen als solche mit hoher Nachbarschaftsdichte (Munson und Solomon 2004 in Baese-Berk, Goldrick 2009: 529f.). Auch bei Kindern scheint die Nachbarschaftsdichte einen Einfluss auf ihre Wortproduktionen zu haben: So wurde u.a. beobachtet, dass Wörter mit hoher Nachbarschaftsdichte präziser produziert werden als solche mit niedriger (Stoel-Gammon 2010: 23).

4.2.1.3.2. Die phonotaktische Wahrscheinlichkeit eines Wortes

Man spricht davon, dass ein Wort eines Lexikons eine *große phonotaktische Wahrscheinlichkeit* aufweist, wenn jedes seiner Phoneme an der entsprechenden Position in vielen Wörtern des Lexikons vorkommt (*segmentale Wortpositionshäufigkeit*) und zudem je zwei aufeinanderfolgende Phoneme dieses Wortes ebenfalls in vielen Wörtern des Lexikons vertreten sind (*Biphonhäufigkeiten*) (Vitevitch, Luce 1998: 326; 1999: 381; Vitevitch 2002: 742). So weist beispielsweise die Silbe *keek* /ki:k/, die nicht im Englischen als Wort verwendet wird, eine große phonotaktische Wahrscheinlichkeit auf, denn die Phoneme /k/, /i:/, /k/ kommen jeweils häufig im Kopf bzw. im Kern bzw. in der Koda im Englischen vor, repräsentiert in Form des Webster'schen Handwörterbuchs, und außerdem folgt das Phonem /i/ häufig auf das Phonem /k/ und das Phonem /k/ häufig auf das Phonem /i/ (Vitevitch und Luce 1998: 326). Untersuchungen haben gezeigt, dass nicht in der Zielsprache als Wörter verwendete Lautformen mit hoher phonotaktischer Wahrscheinlichkeit präziser produziert werden als solche mit geringerer (Stoel-Gammon 2010: 22f.). Insofern die phonotaktische Wahrscheinlichkeit sich auf das Vorkommen von Phonemen innerhalb von Silben beschränkt, kann auch in diesem Fall wie zuvor im letzten Abschnitt gefolgert werden, dass die unterschiedliche Präzision der Produktion von Silben durch das Kind als Anzeichen für eine zunehmende Schnittmengenbildung gewertet werden kann.

4.2.1.4. Bewusste Verfügbarkeit der Silbenschnittmengen: Von starker Silbenbindung der Schnittmengen hin zu phonematischen Einheiten.

Der Begriff der *phonematischen Bewusstheit* (*phonological awareness*(Koda 2007: 220)) ist ein Beispiel des Konzepts der *metalinguistischen Bewusstheit* (*metalinguistic aware-*

ness (Koda 2007: 220)), die die Fähigkeit beinhaltet, Sprachformen zu identifizieren, zu analysieren und zu manipulieren. (Koda 2007: 220) Der Begriff der metalinguistischen Bewusstheit kann sich auf alle sprachlichen Bereiche wie z.B. die Phonologie, Morphologie, Syntax oder Pragmatik beziehen, während der Begriff der phonologischen Bewusstheit nur den Bereich der Phonologie umfasst (ebd.). Der Begriff der *phonematischen Bewusstheit* ist bei der Untersuchung der Lesefähigkeiten von Kindern von Bedeutung, denn Studien zum frühen Lesenlernen zeigen, dass (i) die Sensibilität von Kindern für die Struktur gesprochener Laute direkt mit ihrer Fähigkeit, Wörter zu lesen und zu buchstabieren, in Verbindung steht, dass (ii) die Segmentierungsfähigkeit, d.h. die Unterteilung von Lautformen in Einzelsegmente ein guter Indikator für den zukünftigen Leseerfolg von Schulanfängern und älteren Schülern ist und dass (iii) das Training in phonologischer Bewusstheit¹³ den Leseerfolg signifikant verbessert (Koda 2007: 221). Im Zusammenhang mit dem in dieser Arbeit vorgeschlagenen organischen Schnittmengenkonzept der Silben deutet der Begriff der phonematischen Bewusstheit darauf hin, dass die in den Silben eines Kindes enthaltenen Lautelemente mit zunehmender Anzahl größere Schnittmengen bilden und so Teileinheiten für die gezieltere Beeinflussung durch das Kind zugänglich machen. Dies wird in der folgenden detaillierteren Beschreibung der die phonematische Bewusstheit bildenden Einzelfertigkeiten deutlich: Adams (1990 in Koda 2007: 221) unterteilt die phonologische Bewusstheit in fünf unterschiedliche Fertigungsgruppen, die entsprechend ihrer angegebenen Reihenfolge zeitlich erworben werden: Die erste Gruppe definiert er (ebd.) durch grundlegende perzeptuelle Fähigkeiten wie das Erinnern bekannter Reime wie z.B. dem von *sein* und *fein*, die zweite durch analytische perzeptuelle Fähigkeiten wie das Erkennen und Sortieren von Mustern von Alliterationen und Reimen. Beispielsweise enden *sein* und *fein* mit dem gleichen Laut, während *halt* und *heim* mit dem gleichen Laut beginnen. Die dritte Gruppe umfasst nach ihm (ebd.) sowohl das Bewusstsein für Teile innerhalb einer Silbe und Analysefähigkeiten, d.h. die Zerlegung einer Silbe in ihren Kopf und ihren Reim, die vierte Gruppe phonematische Analysefähigkeiten, d.h. die Durchführung einer vollständigen phonematischen Zerlegung, z.B. des Wortes *kann* in die Phonemfolge /k/+/a/+/n/ (ebd.). Schließlich beinhaltet die fünfte Gruppe die Fähigkeit, phonematische Manipulationen vorzunehmen, d.h. Wörter durch Löschung bzw. Auslassung, Einfügung und Umstellung von Phonemen zu erzeugen wie beispielsweise das Wort *fein* (/fam/) durch Weglassung seines ersten Pho-

¹³ s.u. die genauere Erläuterung der von diesem Begriff erfassten Einzelfertigkeiten

nems /f/ zu *ein* wird (Adams 1990 in Koda 2007: 221). Demgegenüber unterscheidet Yopp (1988 in Koda 2007: 221) nur zwei unterschiedliche Fertigungsgruppen innerhalb der phonologischen Bewusstheit, zum einen eine *simple awareness* (ebd.), worunter er das Erkennen von Silbenkonstituenten versteht (Koda 2007: 221) und zum anderen eine *compound awareness*, worunter er die Analyse und Manipulation von Folgen aus mehreren Phonemen versteht, was den beiden letzten Gruppen der zuvor vorgestellten Gruppierung von Adams (1990) entspricht (Koda 2007: 221). Die phonologische Bewusstheit wird als schrittweise zunehmende Manipulationsfähigkeit in Folge der zunehmenden Sensibilität gegenüber den Einzelphonemen angesehen (Koda 2007: 221) und könnte als Folge des zunehmenden Umfangs der Silbenschnittmengen eines Sprechers interpretiert werden. Sie sollte nach den hier gemachten Annahmen, dass ein positionsbezogenes Phonem sich als Schnittmenge des organischen Silbensystems ergibt, auf die einzelnen konkreten Phoneme bezogene Eigenschaften zeigen. So sollten Phoneme, denen eine größere Silbenschnittmenge zu Grunde liegt, leichter manipulierbar sein als solche Phoneme, die Schnittmenge weniger Silben sind. Der sich aus der Schnittmengendefinition der Phoneme ergebende zeitlich und quantitativ abgestufte Charakter der Phoneme sollte sich auch in entsprechend gestuftem Sprachverhalten sowohl des Kindes als auch der Erwachsenen zeigen. Für die letzteren deshalb, weil das organische Silbensystem dynamisch ist und gegenüber einem statischen Konzept ständigen Veränderungen unterliegt. Neben der Möglichkeit, die phonematische Bewusstheit als Ausdruck der fortschreitenden Silbenschnittmengen zu interpretieren, die die einzelnen Phoneme definieren, weist dieselbe auch auf den primären Status der Silbe gegenüber dem Phonem hin: Denn die Silbe und die mit ihr verbundenen Konstituenten sind früher als die Einzelphoneme im Verhalten der Kinder wiederzufinden. Kinder, die noch nicht zu lesen begonnen hatten, konnten in einer Studie von Treiman und Zukowsky (1991 in Koda 2007: 221f.) genauso gut wie Kinder, die bereits mit dem Lesen begonnen hatten, eine Silbe wie beispielsweise *ham* in unterschiedlichen Wörtern wie z.B. *hammer hammock* erkennen (ebd.). Dies galt aber nicht für ein gemeinsames Phonem wie z.B. /t/ in zwei Wörtern wie beispielsweise *steak* und *sponge*, wobei die des Lesens unkundigen Kinder viel schlechter abschnitten als die Kinder, die bereits mit dem Lesen begonnen hatten (Koda 2007: 221f.). Daraus schlossen Treiman und Zukowsky (1991) laut Koda (2007: 222),

that pre-reading children have little difficulty distinguishing syllables, and are capable of making phonological judgments based on the intra-syllabic, onset-rime distinction (the differentiation between the initial phoneme and the remainder within a syllable). More critically, however, children do

not become capable of breaking intra-syllabic units into smaller, phonemic, segments until they start reading independently. (Koda 2007: 222).

Dies zeigt die scheinbare Abhängigkeit der bewussten Verfügbarkeit von Phonemen vom Lesenlernen, weshalb Koda (2007: 222) ein grundlegendes phonologisches Verständnis von Wörtern, das als Vorläufer bzw. Grundlage des Lesens angesehen wird, von einer elaborierteren Fähigkeit unterscheidet, Phoneme zu analysieren und zu manipulieren (Koda 2007: 222). Die letztere Fähigkeit sieht sie als Folge des Lesenlernens an, bei dem die Entsprechung von Phonem-Graphem-Paaren gelernt werde (ebd.). Die Übereinstimmung der Zeitpunkte des Lesenlernens und der phonematischen Bewusstheit ermöglicht eine Täuschung hinsichtlich des kausalen Zusammenhangs zwischen beiden Prozessen.

Der Zugriff von Kindern auf Einzelkonsonanten z.B. in komplexen Silbenköpfen ist schon vor dem Beginn des Lesenlernens in der Form von Lautspielereien oder Nonsensereimen möglich und kann deshalb als eine dem Lesenlernen vorangehende Bewusstheit einzelner Phoneme auf der Grundlage ausreichend großer Schnittmengen im Silbenlexikon gesehen werden. Diese Schnittmengen ermöglichen zum Zeitpunkt des Lesenlernens die Entwicklung der phonematischen Bewusstheit und ihre Transformation in ein Graphem-Bewusstsein. Ohne die Existenz hinreichend großer Silbenschnittmengen kann ein phonematisches Bewusstsein nicht entstehen und auch nicht die Aneignung eines Graphemsystems. Interessanter Weise beginnt das Lesenlernen in alten Fibeln mit den Graphemen <i> und <a> als für die Kinder am einfachsten nachzuvollziehende Transformation eines Lautbewusstseins in die Schrift. Hierin spiegelt sich wohl eine intuitiv richtige pädagogische Erkenntnis der Schulmeister wider.

4.2.2. Schnittmengen zwischen den Silbensystemen unterschiedlicher Sprecher

Da das Silbenschnittmengensystem eines Sprechers dem organischen Wachstum mit den damit verbundenen funktionalen Anforderungen unterliegt und seinen Anfang mit der ersten Silbe nimmt und im Rahmen der Selbststeuerung durch die Einheit von Differenzierung und Ökonomie sich weiterentwickelt, ist es nicht weiter verwunderlich, dass die Produktionen eines Sprechers, die von einem anderen als mit seinen identisch angesehen werden, bei objektiv physikalischen Messungen starke Abweichungen zeigen können. Dies ist bereits für unterschiedliche Sprecher ein und derselben Sprache der Fall wie z.B. in der ausgedehnten Streuung ein und desselben Vokals unterschiedlicher Sprecher hinsichtlich des ersten und zweiten Formanten sichtbar wird (Johnson, Mullenix 1997: 3).

Aber auf Grund der noch ausreichend großen Schnittmengen zwischen den Silben der unterschiedlichen Sprecher der gleichen Sprache ist eine erfolgreiche Kommunikation gewährleistet. Anders sieht es aus, wenn die Sprecher unterschiedlichen Sprachen angehören, wie durch das im Folgenden vorgestellte Modell von Best (1995) zur Wahrnehmung fremdsprachlicher Phoneme durch einen Sprecher deutlich gemacht werden soll. Bests (1995) Modell der Phonemassimilation ist auch deshalb von Bedeutung, weil das Verhältnis des kindlichen Silbensystems zu dem eines Sprechers der Zielsprache umso mehr dem der Silbensysteme zweier Sprecher jeweils fremder Sprachen gleicht, je mehr das Kind am Beginn seines Worterwerbs steht. So kann ein Fremder, der nicht zu den Bezugspersonen eines Kindes gehört und somit nicht meint, mit den Eigentümlichkeiten der Lautformenseite des Kindes vertraut zu sein, nur ungefähr die Hälfte von dem verstehen, was ein sich normal entwickelndes zweijähriges Kind äußert (Coplan, Gleason 1988 in Stoel-Gammon 2010: 4).

4.2.2.1. Bests (1995) Modell der Phonemunterscheidungsfähigkeit

Best (1995) präsentiert mit ihrem *Perceptual Assimilation Model (PAM)* ein Modell, wie erwachsene Sprachteilnehmer nicht in ihrer Sprache vorkommende Phoneme einer ihnen weitgehend unbekanntem Sprache wahrnehmen und wie sich ein analoges Verhalten bei Neugeborenen entwickelt. Weil Bests (1995) Modell der unterschiedlichen Assimilationsfälle eine Verallgemeinerung von empirischen Beobachtungen darstellt, bleiben diese Assimilationsklassen auch ohne die weitergehenden, fragwürdigen Annahmen des Modells gültig und können als Indizien für die Schnittmengenbildung gewertet werden. Best (1995) geht davon aus, dass die Phoneme als physikalische Invarianten im Sprachsignal direkt vorhanden sind, die wiederum den Gesten der Artikulatorischen Phonologie entsprechen sollen (ebd.). Der Sprecher bzw. Hörer einer Sprache versuche dann während des Wahrnehmungsvorgangs, im Sprachsignal die Gestenkonstellationen ihrer Sprache als akustische Invarianten wiederzufinden. Wie im Abschnitt 4.3. erläutert wird, ist es bisher noch nicht geglückt, Invarianten im akustischen Sprachsignal ausfindig zu machen. Vielmehr müssen die Phoneme als der Bezug von perzipierten Silbeneinheiten oder -segmenten durch einen Hörer auf sein Silbensystem verstanden werden. Best (1995) versucht mit ihrem Modell der Phonemassimilation sowohl das Wahrnehmungsverhalten von Phonemen einer (weitgehend) unbekanntem Sprache durch Sprecher einer anderen Sprache als auch die Veränderung der Wahrnehmung von Lautunterschieden kleiner

Kinder seit ihrer Geburt zu erklären. Dazu nimmt sie zunächst den folgenden Mechanismus der (Nicht-) Assimilation von Phonemen durch einen Sprecher einer Sprache an:

Er erkennt ein Phonem entweder als einen sprachlichen Laut oder nicht. Im erstgenannten Fall des Erkennens eines sprachlichen Lautes in dem Phonem einer anderen Sprache erfolgt nun, sofern dieses einem Phonem seiner Muttersprache ausreichend ähnlich ist, eine Kategorisierung desselben als in die native Phonemkategorie gehörig, der es am ähnlichsten ist. Bei unzureichender Ähnlichkeit hingegen wird das gehörte Phonem zwar als ein sprachlicher Laut angesehen, aber in keine muttersprachliche Phonemkategorie eingeordnet. Im Fall der Einordnung eines Phonems in eine native Kategorie unterscheidet Best (1995: 261) wiederum drei Fälle: Erstens den, dass das assimilierte Phonem identisch mit dem Phonem der Muttersprache ist, in dessen Kategorie es einsortiert wurde, und somit bereits ein Phonem der Muttersprache ist, zweitens den, dass das Phonem einem nativen Phonem ("native invariant") ziemlich ähnlich ("reasonable similar") ist und damit ein akzeptables Vorkommnis dieses nativen Phonems (Best 1995: 261) darstellt und drittens den Fall, dass das wahrgenommene Phonem dem assimilierten nativen Phonem nur wenig ähnlich ist, aber erkennbare Unterschiede bestehen und somit ein abweichendes Vorkommnis dieser Kategorie sei (Best 1995: 261). Abhängig davon, wie ähnlich zwei fremdsprachige Phoneme muttersprachlichen Phonemen sind, und damit, wie gut und zu welchen Phonemen sie assimiliert werden, macht Best (1995: 262f.) hinsichtlich ihrer Unterscheidbarkeit die folgenden sechs Vorhersagen: Sie sind erstens ausgezeichnet unterscheidbar, wenn sie in zwei unterschiedliche Phonemkategorien der Muttersprache einsortiert werden können, zweitens gut bis sehr gut unterscheidbar, wenn eines von ihnen in eine native Phonemkategorie eingeordnet werden kann und das andere als nicht-sprachlicher Laut eingestuft wird, drittens gut bis sehr gut auseinanderzuhalten, wenn sie beide als nicht sprachliche Laute wahrgenommen werden, viertens schlecht oder gemäßigt, wenn sie als sprachliche Laute wahrgenommen werden, aber keines der beiden in eine muttersprachliche Phonemekategorie einsortiert werden kann und schließlich fünftens schlechte Unterscheidbarkeit im Fall, dass beide dem gleichen Phonem der Muttersprache zugeordnet werden, sofern sie beide als gleich gute Vorkommnisse desselben betrachtet werden und sechstens bessere, wenn dies nicht der Fall ist, sie also als unterschiedlich gute Exemplare des nativen Phonems betrachtet werden, in dessen Kategorie sie der Sprachteilnehmer beide einordnet (ebd.).

Für Best (1995) sind entgegen den Annahmen des organischen Silbenwachstums die Phoneme keine dynamischen Zustandsindikatoren des Silbensystems eines Sprechers,

sondern invariant im Sprachsignal vorhandene Hinweise auf die Artikulationsgesten. Deshalb entwickelt das Kind nach ihrer (Best 1995) Auffassung kein eigenständiges System, sondern muss die Gesteninvarianten im Sprachsignal lediglich entdecken (Best 1995). Diesen Entdeckungsvorgang der die gestischen Konstellationen im Sprachsignal repräsentierenden akustischen Invarianten stellt sich Best (1995) als einen durch perzeptuelles Lernen ermöglichten vor. Dadurch sammle das Kind zunehmend mehr Erfahrungen mit dem akustischen Signal und lerne die akustischen Invarianten der Gestenkonstellationen der Muttersprache in hierarchischer Reihenfolge entsprechend ihrer Abstraktheit zu erkennen. Die gegen Ende des ersten Lebensjahres abnehmende Fähigkeit des Kindes, nicht in seiner Muttersprache vorkommende Phoneme zu unterscheiden, sieht Best (1995) dadurch als erklärt an, dass das Kind durch zunehmende Erfahrung der Muttersprache die Gesteninvarianten im Signal immer besser erkenne, während dies für nicht in der Muttersprache vorkommende Phoneme nicht der Fall sei.

Das Phonemassimilationsmodell von Best (1995) stellt eine Verallgemeinerung des Verhaltens von Hörern bei der Wahrnehmung fremder Phoneme dar, sollte aber nicht im Kontext von Invarianten im akustischen Sprachsignal betrachtet werden, sondern dem des Bezugs des durch den Hörer wahrgenommenen Sprachsignals auf sein organisch wachsendes Silbensystem und auf die sich aus diesem ergebenden Phoneme. Das in Bests (1995) Modell zusammengefasste Assimilationsverhalten veranschaulicht dennoch die Schnittmengenbildung zwischen den Silbensystemen der Sprecher unterschiedlicher Sprachen, was insbesondere für den Phonemerwerb auch von Bedeutung ist, weil das Verhältnis des kindlichen Silbensystems zu dem der Zielsprache umso mehr dem zweier Fremdsprachen entspricht, je weiter sich das Kind am Anfang seines Phonemerwerbs befindet. Dieses sich verändernde Verhältnis zwischen Kind- und Zielsprache veranschaulicht auch, dass der Begriff der Fremdsprache ein fließender ist, denn durch den Schnittmengenbegriff werden feine Schattierungen zwischen Silbensystemen erfasst, es gibt keine absoluten Grenzen oder Zugehörigkeiten.

4.3. Phoneme als phonetisch-physikalische Invarianten gegenüber denselben als illusionäre Einheiten

Die Entwicklung von Phonemsystemen ist nur in dem Sinne auf die phonetischen Grundlagen zurückzuführen, als sie die materielle Basis für die Artikulation bilden. Phoneme entstehen jedoch als illusionäre Einheiten im Laufe der ontogenetischen und phylogenetischen Sprachentwicklung durch das Wirken des sich selbst steuernden Silbenbildungs-

wachstums. Versuche, für einzelne Phoneme Invarianten in den akustischen Sprachsignalen zu finden, sind bisher nicht erfolgreich gewesen (Jusczyk 2000: 46), und haben vielmehr das Ausmaß des varianten Verhältnisses zwischen dem akustischen Signal und der für es wahrgenommenen Phonemfolge eines Sprechers verdeutlicht. Im folgenden Abschnitt soll an einigen Beispielen das Verhältnis zwischen der Wahrnehmung einer Phonemfolge und objektiv messbaren Eigenschaften des akustischen Signals veranschaulicht werden. Anschließend wird die *motor theory* der Sprachwahrnehmung von Liberman et al. (1967, 1985) als Versuch dargestellt, die invariant wahrgenommenen Phoneme als die Berechnungen einer angeborenen, neuronalen Struktur aufzufassen.

4.3.1. Das Verhältnis zwischen einem messbaren akustischen Sprachsignal oder den artikulatorischen Bewegungsabläufen und einer durch einen Sprachteilnehmer wahrgenommenen Phonemfolge

Die einfache Eins-zu-Eins-Entsprechung zwischen einem akustischen Sprachsignal und der Folge von Symbolen des für alle Sprachen verwendeten IPA, die ein Sprachteilnehmer nach Einweisung zur Notation dieses Sprachsignals verwenden kann, existiert nicht. So gibt es einerseits unterschiedliche akustische Realisierungen eines Phonems oder einer Folge von Phonemen durch verschiedene Sprecher (Johnson, Mullenix 1997: 2f.): Beispielsweise zeigen die Produktionen eines Vokals durch verschiedene Sprecher derselben Sprache eine ausgedehnte Streuung hinsichtlich der Messung ihres ersten und zweiten Formanten (Johnson, Mullenix 1997: 3). Andererseits kann die Produktion eines Phonems durch einen Sprecher in unterschiedlichen Phonemkontexten sehr unterschiedliche akustische Realisierungen aufweisen (Johnson, Mullenix 1997: 3). Dies wird auch als *lack of invariance problem* (ebd.: 4) bezeichnet: Mit dieser Veränderlichkeit eines Phonems im Kontext anderer Phoneme befasst sich die Koartikulationsforschung (Morton, Tatham 2006: 22), die u. a. entdeckt hat, dass z. B. in der Äußerung /əku/ die Lippenrundung für den Vokal /u/ schon während des vorhergehenden /k/ oder Schwa beginnt, was zeigt, dass es keine objektiv messbaren, kontextfreien Phoneme innerhalb einer Äußerung gibt (Morton, Tatham 2006: 25). Hierbei wird auch deutlich, dass das Sprachsignal ein Kontinuum ist, in dem mehrere als diskret gehörte Phoneme nicht tatsächlich als mit einem Abstand aufeinanderfolgende diskrete Einheiten vorhanden sind. Dieses Phänomen benennen Liberman et al. (1967: 436) als *the disappearance of phoneme boundaries* (ebd.) und ziehen die Schlussfolgerung aus ihrem Versuch, eine akustisch vorliegende

CV-Silbe zwischen dem Konsonanten und dem Vokal so zu trennen, dass sowohl ein reiner Konsonant als auch ein solcher Vokalteil erhalten wird:

We cannot cut either the /di/ or the /du/ pattern in such a way as to obtain some piece that will produce /d/ alone. If we cut progressively into the syllable from the righthand end, we hear /d/ plus a vowel, or a nonspeech sound; at no point will we hear only /d/. (Lieberman et al. 1967: 436).

Die Wahrnehmung von Phonemen ist zudem keine rein akustisch fixierte Einheit: Denn Hörer können für einen akustischen Stimulus in Abhängigkeit von dem Vokal, der mit dem akustischen Stimulus zeitlich nachfolgend kombiniert und ihnen präsentiert wird, unterschiedliche Konsonanten wahrnehmen (Lieberman et al. 1952 in Jusczyk 2000: 45).

4.3.2 Phoneme als gehörte Gesteninvarianten in der *motor theory* der Sprachperzeption mit Hilfe eines angeborenen Gestenübersetzungsmoduls

Die *motor theory* der Sprachperzeption (Lieberman et al. 1967: 2f.) nimmt an, dass der Mensch über ein spezialisiertes Modul im Gehirn verfügt, welches die von einem Sprecher produzierten und von einem anderen Sprachteilnehmer gehörten akustischen Signale automatisch, ohne Zwischenschritte über beispielsweise das auditive Abbild, direkt in invariante Bewegungsbefehle übersetzt (ebd.), die auch beabsichtigte phonetische Gesten genannt werden. Dabei wird angenommen, dass dieses Modul jedem Menschen angeboren ist, also nicht erlernt werden muss, sich also nur phylogenetisch entwickelt hat, aber ontogenetisch keinen Veränderungen unterliegt. Vom evolutionären Standpunkt aus erhebt sich die Frage, wie sich ein neuronales Modul entwickelt haben soll, das im Gehirn des heutigen Menschen angeboren ist und während des Lebens keinen Veränderungen unterliegt. Da dazu von Liberman et al. (1985) keine Angaben gemacht werden, handelt es sich um eine sog. *black box*, der andere konstruktivere theoretische Ansätze vorzuziehen sind. Nach der Konzeption dieses Moduls ist Sprachwahrnehmung also keine allgemeine auditive Fähigkeit des Menschen, sondern ein auf die Sprachproduktion und -perzeption spezialisierter Vorgang. Dabei entsprechen diese Bewegungsbefehle bzw. beabsichtigten Gesten herkömmlichen phonetischen Merkmalen wie dem Zurückziehen der Zunge, Lippenrundung oder Kieferanhebung (Lieberman et al. 1985: 2). Ein Phonem ergibt sich dann als eine bestimmte Menge dieser Gesten (ebd.) wie z.B. für das Phonem /m/ aus einer Geste für einen labialen Plosiv oder für /b/ aus derselben und zusätzlich der der Velumabsenkung (ebd.). Es wird angenommen, dass diese Gesten die elementaren

Einheiten sowohl der Sprachperzeption wie auch der Sprachproduktion sind (Liberman et al. 1985: 2), wie dies auch in der Artikulatorischen Phonologie der Fall ist. Die Wahrnehmung einer produzierten Äußerung wird in dieser *motor theory of speech perception* in der Wahrnehmung eines speziellen Musters von produzierten, intendierten Gesten des Sprechers (Liberman et al. 1985: 3) gesehen, dessen Tätigkeit selbst sich wiederum auf Grund des den Gesten zugeschriebenen grundlegenden Status auf die Produktion derselben beschränkt (Liberman et al. 1985: 3). Das Objekt der Sprachwahrnehmung ist also ein invariantes Bewegungsobjekt (Liberman et al. 1985: 7). Liberman et al. (1985: 7) machen darauf aufmerksam, dass ihre Theorie der Sprachwahrnehmung nicht nur bedeute, dass Sprachperzeption und -produktion auf der Grundlage der gleichen Einheit funktionieren, sondern sie (ebd.) machen deutlich, dass die Theorie einseitig auf die Bewegungen der Sprache konzentriert sei (ebd.). Dies ist v.a. vor dem Hintergrund ihrer erfolgreichen Versuche zu verstehen, im akustischen Signal Invarianten für die Phoneme auszumachen, was wie im vorhergehenden Unterabschnitt exemplarisch illustriert, problematisch und sicher unmöglich ist. Somit nehmen Liberman et al. (1985: 6) an, dass das oben erwähnte Modul zur Übersetzung von gehörter, akustisch kontinuierlicher Sprache in zeitliche Gestenzusammenstellungen die systematischen Regelmäßigkeiten der Koartikulation von Gesten nutzt (Liberman et al. 1985: 6), die sich in der Akustik widerspiegeln würden. Für diese Annahme ist aber notwendig, dass die Sprache wirklich auch in der Produktion von einzelnen Gesten besteht und nicht etwa der Schwerpunkt auf der Produktion ganzer Silben liegt, wie es in dem hier vertretenen Ansatz des organischen Silbenwachstums angenommen wird und es Ergebnisse der Koartikulationsforschung nahelegen (Wester 2003: 233). Dies soll natürlich nicht bedeuten, dass die Phoneme nicht mit zunehmender Entwicklung des Lexikons eine gewisse Stärke erlangen, aber sie bleiben vom organisch-dynamischen Silbensystem eines einzelnen Sprechers abhängige Größen.

Liberman et al. (1985: 10ff.) bringen eine Reihe von Argumenten für ihre Konzentration auf das angeborene Gestenübersetzungsmodul wie etwa das Fortbestehen des Hörens bestimmter Phoneme in einem synthetisierten akustischen Signal auch bei Weglassung einiger akustischer Merkmale oder den Ausgleich von verminderten akustischen Merkmalen durch andere in einer Art "trading relation" (ebd.), die noch einmal darauf hinweisen, dass ihre Theorie als Reaktion auf das Nicht-Auffinden der akustischen Invarianten aufzufassen ist. Die Veränderlichkeit der Phoneme im Sprachfluss, die sich aus der Logik der Annahme der Produktion invarianter Einheiten wie den Phonemen ergibt, sehen Liberman et al. (1985: 14f.) auch als Argument für ihre Theorie an, denn diese Veränder-

lichkeit würde systematische Hinweise auf die Erzeugung durch die invarianten Gesten geben und so deren Zuordnung zum akustischen Signal unterstützen bzw. erleichtern (ebd.). Eine erfolgreiche Sprachwahrnehmung hat dann stattgefunden, wenn es mittels des angeborenen Gestenübersetzungsmoduls möglich war, aus einer Sequenz von zeitlich parallel angeordneten Gesten die Oberflächenform aus einer angenommenen Gestenfolge abzuleiten und sie so als mit diesen Gesten erzeugbar zu bewerten (Lieberman et al. 1985: 19). Dies ist im Wesentlichen ein Verhältnis zwischen einer zu Grunde liegenden Phonemfolge (*underlying form*) und der Oberflächenform (*surface form*) wie sie in der generativen Phonologie angenommen wird, auch wenn hier statt der invarianten Phoneme die Gesten fungieren und statt der Ableitungsregeln das phonetische Interpretationsmodul. Dies formulieren Liberman et al. (1985: 19) auch im Zusammenhang mit der Frage nach der Unterscheidbarkeit von nicht sprachlichen akustischen Signalen und solchen, die zur Sprache gehören:

We conclude that acoustic patterns are identified as speech by reference to deep properties of a linguistic sort: if a sound can be 'interpreted' by the specialized phonetic module as the result of linguistically significant gestures, then it is speech; otherwise, not. (In much the same way, grammatical sentences can be distinguished from ungrammatical ones, not by lists of surface properties, but only by determining whether or not a grammatical derivation can be given.) Of course, the kind of mechanism such an 'interpretation' requires is the kind of mechanism the motor theory presumes. (Lieberman et al. 1985: 19).

Die Beantwortung der Fragen, die sich aus den "Veränderlichkeiten" im akustischen Sprachfluss ergeben, wird also in den Bereich des Angeborenen verschoben. Anstatt die Sprachwahrnehmung des Kindes im ersten Lebensjahr als einen Vorgang der Erfahrung mit der Zielsprache zu begreifen, sehen Liberman et al. (1985: 24) die Fähigkeiten von sehr jungen Kindern, eine große Anzahl phonetischer nativer und nicht-nativer Kategorien zu unterscheiden (Eimas, Siqueland, Jusczyk and Vigorito (1971) in Liberman et al. (1985: 24)) als Beweis dafür an, dass Gesten und die mit ihnen verbundene Kopplung von Perzeption und Produktion über das neuronale Gestenübersetzungsmodul angeboren sind und dass auf Grund der abnehmenden Unterscheidungsfähigkeit der nicht-nativen phonetischen Kategorien durch Kinder (Werker, Tees, 1984 in Liberman et al. (1985: 24)), diese die Unterscheidungsfähigkeit für die nicht in der Zielsprache verwendeten Gesten verlieren (Lieberman et al. 1985: 24). Somit wird der Phonemerwerb als Verlernen von angeborenen Gestenmodulen mit unklarem Entwicklungsursprung angesehen, anstatt als aktiver Entwicklungsprozess in Auseinandersetzung mit der Zielsprache.

4.3.3 Phoneme als Silbensystemgrößen in Form von Silbenschnittmengen

Das Problem, die Invarianten von Phonemen im Sprachsignal finden zu müssen, beginnt mit der Annahme, dass jeder Sprecher bereits über Phoneme als diskrete Einheiten verfügt. Es wird dabei nicht in Betracht gezogen, dass Phoneme das Ergebnis einer Entwicklung sind. So ergibt sich das "lack of invariance problem" (Johnson, Mullenix 1997: 4), das bedeutet, dass sich das akustische Sprachsignal nicht in die gehörte Abfolge aus voneinander getrennten akustischen Einheiten zerlegen lässt, sondern sich Veränderlichkeit zeigt wie in Abschnitt 4.3.1. beispielhaft angeführt. Eine scheinbare Lösung dieses Varianzproblems ist mit der *motor theory* der Sprachperzeption von Liberman et al. (1985) gegeben worden, indem sie ein angeborenes neuronales Übersetzungsmodul annehmen. Damit ist aber die Erklärung der Sprachperzeption in einer *Black Box* verpackt worden, die eine Entwicklung nicht erklärt. Demgegenüber werden in diesen Ausführungen die Silben als Grundeinheiten angesehen¹⁴. Eine der Eigenschaften von Sprache allgemein und der Silbe im Besonderen ist ihre Unterscheidbarkeit und damit Eindeutigkeit. Keine sprachliche Lautform darf im Allgemeinen mit sich selbst identisch sein, wenn unterschiedliche Bedeutungen mit ihr verknüpft sind. Die bekannten Ausnahmen bestätigen die Regel.

Das Kind beginnt mit einer ersten Silbe und erweitert sein Silbensystem durch das organische Wachstum, dass die Silbenbildung über die Prinzipien von ökonomischer Erzeugung und differenzierender Eindeutigkeit steuert. Dabei können bereits existierende Silbenschnittmengen nicht grenzenlos weiter ausgedehnt werden, weil auf Grund der notwendigen Unterscheidbarkeit gewisse Dichtewerte nicht überschritten werden dürfen. Das Silbensystem des Kindes muss dem ständigen Druck nachkommen, neue Silben für die Bedeutungsunterscheidung zu bilden, dadurch ist es u.a. in ständiger Veränderung und versucht den Ausgleich zwischen Ungleichgewichten zu schaffen. Die Phoneme sind dabei nicht die bereits vorhandenen Grundeinheiten, sondern vom Silbensystem abhängige Größen, die den Schnittmengen entsprechen, die die Silben miteinander bilden. Verschieben sich die Silben im detaillierten Speicherraum, so verändert sich auch die Qualität der Phoneme. Da die Bildung neuer Silben sich aus dem bereits gebildeten System ergibt, ist die Dynamik in diesem Sinne unter Beachtung der funktionalen Anforderungen determiniert, und wegen der Kontinuität des Silbenraums ist die Existenz von je zwei

¹⁴ Siehe auch Kapitel 5.

identischen Phonemsystemen unterschiedlicher Sprecher unwahrscheinlich. Die Invarianz der Phoneme zeigt sich also nicht in ihrer Gegebenheit als unveränderliche Einheiten im akustischen Sprachsignal, noch in Form von angeborenen, sprachspezifischen neuronalen Rechenmoduln, sondern in dem unveränderlichen Eindruck, den ein individueller Sprecher erhält, wenn er das Sprachsignal wahrnimmt und dieses in Bezug auf sein eigenes organisches Silbensystem und die sich daraus ergebenden Phoneme interpretiert.

4.4. Sprachuniversalien und das organische Silbenwachstum

Sprachstrukturen werden als *universal*¹⁵ bezeichnet, wenn sie für alle Sprachen der Welt zutreffen (Mairal, Gil 2006: 1). Im Folgenden sollen vier verschiedene Auffassungen des Begriffs der Universalität in der Sprache kurz vorgestellt werden.

4.4.1. Universalität als angeboren postulierte Sprachstruktur

Chomsky und Halle (1968: 5) nehmen für alle Sprecher aller Sprachen angeborene, binäre distinktive Merkmale an. Dies begründen sie mit dem bekannten "poverty of the stimulus"-Argument (Lust 2006: 47), nach dem das Kind auf Grund eines Mangels an ausreichenden Daten der Zielsprache diese nicht zuverlässig aus dem Input lernen könne (Chomsky, Halle 1968: 4). Jedes Kind verfügt aus ihrer Sicht also von Beginn an über die durch die binären distinktiven Merkmale festgelegten Phoneme und deshalb sind die Phoneme universell. Die Begründung des angeborenen Charakters der binären distinktiven Merkmale und damit der endlich vielen durch sie definierten Phoneme ist aber angesichts starker Verarbeitungskapazitäten wie etwa des statistischen Lernens (Thiessen 2009) nicht haltbar, das dem Kind zuverlässig Regelmäßigkeiten der Zielsprache zeigt. Viel wichtiger ist aber die Konzeption des Phonemerwerbs an sich: Die Tätigkeit des Kindes auf eine minimale Extraktionsprozedur mit Hilfe angeborener grammatischer Module zu begrenzen, mit denen die wahrscheinlichste Grammatik durch das Kind gefunden werden kann (Chomsky, Halle 1968), unterschätzt die Eigenständigkeit des Kin-

¹⁵ Die hier gegebene Definition von Sprachuniversalien wird auch mit dem Attribut *absolut* versehen (Bybee 2006: 108), während von diesen sog. *relative* Sprachuniversalien abgegrenzt werden (ebd.), die mit Hilfe des Begriffes der Markiertheit definiert werden (ebd.), der von je zwei Sprachstrukturen x, y wie z.B. nasalen und oralen Vokalen (ebd.) x als markierter als y bezeichnet (hier: nasale Vokale (x) sind markierter als orale (y)), wenn x u.a. in insgesamt weniger unterschiedlichen Sprachen vorkommt als y (ebd.: 181), x seltener in einem ausgewählten Text einer Sprache, die beide Strukturen x und y aufweist, wiederholt wird als y (ebd.) und zudem die Existenz von x in einer bestimmten Sprache die von y impliziert bzw. erfordert (ebd.). Relative Universalien sind nicht so umfassend wie absolute Universalien (Bybee 2006), weil sie sich nicht auf alle Sprachen der Welt beziehen müssen, sondern nur auf die, die die beiden Elemente x und y einer Markiertheitsrelation enthalten (ebd.).

des und begrenzt die Erkenntnismöglichkeiten zum Verständnis des eigenen Systems des Kindes. Diese Sichtweise ist an absoluten Maximalwerten orientiert, die zwischen der als fehlerhaft verstandenen Sprache des Kindes und der als richtig verstandenen Sprache der Erwachsenen (Zielsprache) bestehen (vgl. Abschnitt 2.1.), und sie geht von der Annahme statischer Beziehungen aus, die die offensichtliche Dynamik des kindlichen Spracherwerbs durch angeborene statische Regeln erklären soll. Fasst man dagegen das kindliche Sprachsystem und insbesondere sein dynamisch sich wandelndes und entwickelndes Phonemsystem als ein offenes dynamisches System auf, entfallen die Probleme, die Annäherung beider Sprachniveaus aneinander zu erklären, da sie zueinander in einer fließenden, über Schnittmengen definierten Beziehung stehen. Denn es geht nicht vorwiegend darum zu erklären, warum das Kind die "richtige" Grammatik der Zielsprache nicht in ihrer vollen Komplexität anwendet, sondern zu zeigen, wie es aktiv auf der Grundlage seiner Voraussetzungen Lauteindrücke der Zielsprache in sein eigenes System integriert, sich so aber der Zielsprache allmählich annähert. Dabei darf nicht vergessen werden, dass es für den Phonem-erwerb kein statisches Endziel gibt, weil auch das Silbensystem eines erwachsenen Sprechers ständig in Bewegung ist. So erreichen amerikanisches Englisch lernende Kinder beispielsweise erst nach durchschnittlich sieben Jahren etwa im Alter von acht Jahren ein passives Vokabular von 10000 Wörtern (Anglin 1989: in Steol-Gammon 2010: 19). Da ein erwachsener Sprecher über ein erheblich größeres Lexikon verfügt, und die Silbenbildung vom Lexikon angetrieben wird, ist somit der Phonemerwerb auch mit 8 Jahren noch nicht abgeschlossen. Für das Deutsche werden beispielsweise als theoretische Obergrenze 38.000 unterschiedlich Silben angenommen (Schweitzer, Möbius 2004). Also bedarf das Kind einer langen Zeit zur Erreichung eines ausreichenden Schnittmengenumfangs mit der Zielsprache, damit man von einem Phonemniveau eines Erwachsenen sprechen kann. In dieser Zeit erkundet das Kind zudem intensiv seine Umgebung und seine Kognition beeinflusst den Spracherwerb und die Sprachentwicklung. Mit der Annahme angeborener distinktiver Merkmale und damit auch angeborener Phoneme kann der Prozess des Phonemerwerbs deshalb nicht verstanden werden.

4.4.2. Universalität als Optimierung von Sprachformen entlang ordinaler Qualitätsskalen

Die grundlegende Annahme dieses Ansatzes postuliert, dass die Veränderung von Sprache immer eine Optimierung bezüglich der Qualitätsskala eines sprachlichen Parameters z.B. der Silbenstruktur oder der Wortlänge (Vennemann 1988: 1ff.) darstellt, die als

Sprachuniversalie angesehen wird. Die Veränderung eines Sprachparameters wird demnach mit der Verbesserung der entsprechend der Qualitätsskalen dieses Parameters schlechteren (zeitlich) vor den besseren Strukturen (ebd.: 2) gleichgesetzt. Werden jedoch umgekehrt zunächst die besseren vor den schlechteren sprachlichen Einheiten eines Parameters verändert, so wird der Grund für diese Durchbrechung der Qualitätsskalen in der Optimierung eines anderen sprachlichen Bereiches gesehen (ebd.: 3f.). Die Silbenpräferenzgesetze (Vennemann 1988) entsprechen nun Qualitätsskalen des Sprachparameters Silbenstruktur und werden somit selbst im oben genannten Sinn als universal betrachtet (ebd.: 4). Sie machen Aussagen zu der historischen Veränderung konkreter Silben und beinhalten u.a. die Silbenmerkmale der Anzahl von Phonemen einer Silbenkonstituente, die der konsonantischen Stärke von Einzelphonemen, die des konsonantischen Stärkeverhältnisses der Phoneme oder unterschiedlicher Konstituenten und die der Anzahl der Moren einer Silbe (ebd.). Neben vielen weiteren Präferenzgesetzen besagt beispielsweise das "Head law" (Vennemann 1988: 13), dass der Kopf einer Silbe umso präferierter ist, je weniger sich die Anzahl seiner Phoneme von Eins unterscheidet und je größer sowohl das Verhältnis der konsonantischen Stärke des ersten Phonems zu der des Kerns dieser Silbe ist als auch die konsonantische Stärke des ersten Phonems für sich betrachtet (ebd.). Historische Veränderungen von Silben, die vor der Veränderung von durch die Präferenzgesetze als schlechter eingestuften Silben erfolgen¹⁶, widersprechen den Präferenzgesetzen dabei nicht, weil sie als die Optimierung eines anderen sprachlichen Parameters angesehen werden (s.o.). Die Grundannahme ist demnach, dass ein Sprachparameter wie die Silbenstruktur eine absolute inhärente phonetische Qualität hat. Dies erscheint aus dem Blickwinkel der im Wesentlichen gleichen physiologisch-anatomisch-neuronalen Ausstattung des Menschen plausibel. Denn der durch die Artikulatoren erzeugbare akustische Raum und dessen Verhältnis zur Perzeption sollten für alle Menschen im Wesentlichen gleich sein. Deshalb sagt Vennemann (1988: 4) auch, dass in einer "sufficiently rich phonetic theory" (ebd.) eine Begründung für diese Gesetze gefunden werden müsse.

Die Präferenzgesetze (Vennemann 1988) sind als umfassende Verallgemeinerungen der historischen Veränderungen von Silben und der typologischen Verteilung der in den Gesetzen integrierten Silbenmerkmale und -relationen anzusehen. Es ist zu fragen, ob die Präferenzgesetze zutreffende Verallgemeinerungen und Rationalisierungen phänomenaler Erscheinungen von Silbenstrukturen darstellen, die aber durch ihre ausschließlich phone-

¹⁶ wie beispielsweise in Auer (1994: 56ff.) aufgeführt

tische Begründung den inneren Systemzusammenhang der Sprachentwicklung und des Sprachwandels nicht vollständig aufklären können (vgl. auch Abschnitt 2.2.). Die Nutzung des artikulatorischen Raumes und der Artikulatoren in verschiedenen Sprachen ist sehr unterschiedlich. Wenn eine phonetische Logik hinter den Präferenzgesetzen stände, wäre zu erklären, warum sie nicht universell ist. Die Sprachveränderung wird aber nicht direkt durch die physiologische Seite der Sprache angetrieben, sondern durch den inneren Systemcharakter der Sprache gesteuert, der auf Unterscheidung und Eindeutigkeit gerichtet ist unter ökonomischer Nutzung der physiologischen Ressourcen.

4.4.3. Universalität als allgemeine während des Sprachgebrauchs operierende neuronale Verarbeitungsmechanismen

Bybee (2008) macht darauf aufmerksam, dass durch Sprachvergleich gefundene Verallgemeinerungen oft Ausnahmen haben (ebd.: 112) und es deshalb das "typical dilemma" (ebd.) sei, dass es kaum wirkliche Universalien gebe (ebd.: 108), das heißt Beschreibungen von sprachlichen Regelmäßigkeiten, die für alle Sprachen der Welt gelten. Sie sieht aber hingegen die während der Sprachverarbeitung wirkenden neuronalen Mechanismen als die wahren Universalien an, weil diese immer in jeder Sprache beim einzelnen Sprachgebrauch wirkten und gemeinsam zu häufig beobachteten Sprachveränderungsabfolgen führten wie beispielsweise /p/ > /pf/ > /f/ > /h/ > /ø/ (ebd.: 108f.), die sie auch als "common paths" (ebd.: 109) bezeichnet. Als Beispiele solcher Mechanismen führt sie u.a. die Wiederholung von Wörtern an, die zu deren Automatisierung und wiederum gestischer Reduktion führten (ebd.: 120), außerdem die Kategorisierung phonetischer Varianten auf der Grundlage phonetischer Ähnlichkeit (ebd.) oder die Verstärkung phonetischer Unterschiede (ebd.). Es ist auf Grund der vergleichbaren Physiognomie und neuronalen Ausstattung von einem allen Menschen in gleicher Weise theoretisch maximal verfügbaren Raum auszugehen, der u.a. mit den artikulatorischen, auditiven und neuronalen Fähigkeiten der Menschen verbunden ist und der die Grundlage der Sprachentwicklung und -veränderung bildet. Aber diese Grundlage muss deshalb nicht zwingend die Veränderung der Lautformen verursachen. Ebenso wenig wie Werkzeuge allein ein Produkt herstellen können, für die es eines Konzepts und einer Intention und Steuerung bedarf, sind die artikulatorischen "Werkzeuge" für sich oder in einer direkten neuronalen Rückkopplung die Urheber der Sprachentwicklung und -veränderung (vgl. auch Abschnitt 2.2.).

In Bybees (2008) Annahmen bleibt auch die Frage offen, wie die Unterschiedlichkeit von Sprachen mit ihren Mechanismen zu erklären ist. Da kein innerer Zusammenhang zwi-

schen den einzelnen von ihr aufgeführten Verarbeitungsmechanismen erkennbar wird, stehen diese in einem zufälligen Verhältnis und die Unterschiedlichkeit der Sprachen wäre damit ebenfalls ein zufälliges Resultat. Die häufig beobachteten "common paths" (ebd.) wie die oben mit /p/ beginnende Phonemveränderung sind das Ergebnis eines von außen auf verschiedene Sprachen angewandten Vergleichs, der rein phänomenal ist und keine Erklärung aus der inneren Systemstruktur der Sprachen anstrebt. Wegen der gemeinsamen physiologischen und neuronalen Grundlage aller Menschen und weil sich auch annehmen lässt, dass es bei allen Unterschieden zwischen Sprachen und Sprechern bevorzugte Wege der Nutzung des sprachphysiologischen Raumes gibt, sind Schnittmengen im phänomenalen Bereich verschiedener Sprachen zu erwarten, die aber nichts über die Systemstruktur der einzelnen Sprachen aussagen.

4.4.4. Universalität als Gesetze des organisch wachsenden Silbensystems

In dem hier vorgestellten Phonemerwerbsansatz wird davon ausgegangen, dass jeder Sprachteilnehmer sein eigenes Silbensystem besitzt, das mit denen anderer nur über Schnittmengen in Beziehung steht. Die Entwicklung des Silbensystem findet zwar auf der allen Menschen gemeinsamen physiologisch-neuronalen Grundlage statt, ist aber nicht durch diese bedingt, sondern durch die Anforderung des Silbensystems, das nicht angeboren ist, sondern entsteht und in dem die Phoneme als Schnittmengen der Silben eine sekundäre Größe sind, die sich aus den Silben mit primärem Status erst ergeben. Ursächlich für die Sprachentwicklung wird der Silbenbedarf des kausalen Motors angesehen: Mit der steigenden Anzahl von zur Bedeutungsunterscheidung erforderlichen Silben, steigen auch die Anforderungen an das Silbensystem. Die Ökonomie des Systems erfordert ein Anknüpfen an bereits gebildete Silbenbestandteile, d.h. möglichst große Schnittmengen, während die notwendige Eindeutigkeit der Silben durch Differenzierung mittels unterschiedlicher Kombinationen oder der Einführung neuer Silbenelemente gewährleistet wird, was sich tendenziell in geringeren Schnittmengen ausdrücken kann. Das Universale zeigt sich hierbei in der Individualität jedes Sprach- und Silbensystems jedes einzelnen Sprechers, das mit denen anderer Sprecher nur Schnittmengen bildet, so dass die dynamisch sich verändernden Phoneme, die die Sprache und das subjektive Empfinden einer Sprachgemeinschaft bilden, durch diese "fließenden" Schnittmengen verändert werden.

4.5. Ausmaße der Eigenständigkeit des kindlichen Phonemerwerbs und sein Verhältnis zur Zielsprache

In diesem Abschnitt soll zunächst für einige Phonemerwerbsansätze aufgezeigt werden, in welchem Umfang sie dem Kind ein eigenes Phonemsystem zugestehen und welches Verhältnis dasselbe zur Zielsprache einnimmt.

4.5.1 Vorschläge anderer Ansätze zum Umfang einer kindbezogenen systematischen Komponente und deren Zielsprachverhältnis

Die Ansätze zur Erklärung bzw. Beschreibung des Phonemerwerbs können zum einen danach klassifiziert werden, inwieweit sie dem Kind gegenüber der Zielsprache eine eigene systematische Komponente zugestehen und zum anderen, wie stark sich diese für unterschiedliche Kinder unterscheiden darf bzw. wie kindindividuell sie konzipiert ist.

So gibt es zum einen Ansätze, denen ihre statische Sicht des Wesens des Phonems ermöglicht, die kindlichen Äußerungen aus den bereits mental vorhandenen Phonemfolgen durch bestimmte Einschränkungen abzuleiten, denen das Kind unterliegen soll. In solchen Ansätzen kommt das dynamische Wesen des Phonemerwerbs nicht zur Geltung, in dem sich Produktion und Perzeption miteinander entwickeln und die Phoneme, aufgefasst als Systemgrößen des individuellen Silbensystems, das Ergebnis des bis zu einem Zeitpunkt erzeugten Silbensystems sind. In solchen Ansätzen wie den beiden im Folgenden präsentierten wird v.a. das Abweichen des Kindes von der Zielsprache betont.

So nimmt Smith (2010) einen Standpunkt ein, der davon ausgeht, dass das Kind auf die Phonemsequenzen Zugriff hat, die ein Erwachsener der entsprechenden Sprache in Zielsprachäußerungen hören würde. Es gibt also keine Entwicklung der Phoneme beim Kind auf der mentalen Ebene, weil diese bereits vorhanden sind. Das Systematische der kindlichen Äußerungen sieht Smith (2010) lediglich im Fehlen von Phonemen in Produktionen von zielsprachlichen Äußerungen des Kindes: "The child has no system of his own, except in so far as this reflects the absence of a characteristic of the system he is exposed to." (Smith 2010: 103). Dieses Fehlen erklärt er sich durch die fehlerhafte Steuerung von Artikulatoren über ein neuronales Netz, welche erst noch erlernt werden müsse (ebd.). So sagt er zu der späten Produktion von Velaren in den Äußerungen eines von ihm untersuchten Kindes: "Rather than being a characteristic of an autonomous system this should be viewed as the side-effect of the inability to control the articulation of a gesture which was appropriately mentally represented [...]" (Smith 2010: 103). Smith (2010) sieht das

System des Kindes lediglich in artikulatorischen Koordinationsschwierigkeiten repräsentiert, die zur Nicht-Produktion von ansonsten dem Kind zugänglichen Phonemen führen. Sein (Smith 2010) Ansatz kommt somit über eine Klassifikation von Abweichungen der kindlichen Äußerungen in Fehlergruppen nicht hinaus. Dabei werden die kindlichen Äußerungen in solche klassifiziert, die zu denen der Zielsprache über Regeln in Beziehung gesetzt werden wie u.a. der Löschung, der Assimilation und der Substitution eines Phonems (Lust 2006: 158). Auf diese Weise wird ein System von Differenzverhältnissen mittels verschiedener Regelarten zwischen den kindlichen Lautformen und den Zielsprachlautformen konstruiert. Der Blick auf die Regelmäßigkeiten und systematischen Zusammenhänge innerhalb der Äußerungen des Kindes, unabhängig von den möglicherweise für das Kind als Vorbild dienenden Zielsprachlautformen, fehlt dagegen (vgl. Abschnitt 2.1). Stampes (1979) Ansatz ist mit dem von Smith (2010) vergleichbar hinsichtlich der Annahme des mentalen Vorhandenseins der Phoneme der Zielsprache sowie auch der Begründung, warum diese nicht entsprechend der Phonemsequenzen von Zielsprachlautformen geäußert werden. Er postuliert für alle Kinder aller Sprachen sog. *phonologische Prozesse* und exakt zur Verfügung stehende Phonemfolgen (Stampe 1979: 1), die beide angeboren seien (Stampe 1979: 45 , 47) und eine reale neuronale Grundlage haben sollen (Stampe 1979: 44): Die phonologischen Prozesse operieren dabei auf diesen Phonemfolgen und ersetzen Phoneme aus als schwieriger angenommenen Phonemklassen durch welche aus als einfacher eingestuften Phonemgruppen (ebd.). Sie beruhen also auf einem Konzept von Schwierigkeitsklassen der Phoneme (ebd.). Beispielsweise klassifiziert Stampe (1979: 1) stimmhafte Plosive als schwieriger als die stimmlosen Entsprechungen, weshalb er einen Prozess annimmt, der die Erstgenannten durch die Zweitgenannten ersetzt (ebd.). Stampe argumentiert dabei mit den inhärenten Schwierigkeiten, die mit der Erzeugung der stimmhaften Plosive verbundenen sein sollen: So erfordern stimmhafte Plosive einen stärkeren Luftstrom aus den Lungen zur Stimmerzeugung durch die Glottis, der aber gleichzeitig durch den kompletten Verschluss von Mund und Nasenraum nach außen zunehmend behindert wird (ebd.). Dies sei dagegen bei stimmlosen Plosiven nicht der Fall (ebd.). Neben dem Ersetzen von stimmlosen durch stimmhafte Plosive kann das Problem der Aufstauung von Luft und damit der Behinderung der Stimmbildung auch durch das Öffnen des Velums, das Herabsenken der Glottis und Öffnen des Mundes gelöst werden (ebd.). Der Phonemerwerb eines Kindes bestehe nun darin, die Prozesse zu verlernen bzw. zu unterdrücken, die die schwierigeren Phoneme der Zielsprache durch einfachere ersetzen und somit zu den "Abweichungen" der kindlichen Äußerungen von

Zielsprachformen führen (Stampe 1979: 21). Stampe (1979: 14) vertritt die Ansicht, dass für alle angeborenen phonologischen Prozesse letztlich eine phonetische Erklärung gegeben werden kann (ebd.). Aus diesem Grund und wegen der Konzeption eines Teils der phonologischen Prozesse als Substitutionen von schwierigeren durch einfachere Phoneme wird als Hauptursache für die beobachteten Abweichungen der kindlichen Lautformen von den ihnen zugeordneten Zielsprachformen physisch-artikulatorisches Unvermögen angenommen und somit eine dem Ansatz von Smith (2010) vergleichbare Erklärung angeboten. Wie auch bei Smith (2010) bleibt hier die Hauptfrage, wie das Kind zu Phonemfolgen der Zielsprachwörter kommt.

Während die beiden zuletzt vorgestellten Ansätze dem Kind bis auf die Annahme von Schwierigkeiten, die mit der Artikulation der bereits vorhandenen Phonemfolgen verbunden seien, keine weiteren Systemeigenschaften zuschreiben, hebt sich der Ansatz von Jakobson (1962) dadurch ab, dass er im Phonemerwerb des Kindes mehr Regelmäßigkeiten erkennt. So beinhaltet sein Modell des Erwerbs der distinktiven Merkmale zum einen eine zeitliche Komponente und zum anderen durch die mit den distinktiven Merkmalen verbundene Klassifikation der Phoneme eine strukturelle Komponente, die zudem über die vorwiegend akustische Definition der distinktiven Merkmale und das von ihm postulierte Konzept des maximalen Kontrastes (Jakobson 1962: 375) die Funktion der Unterscheidbarkeit von Lautformen beinhaltet. Jedoch bezieht Jakobson (1962) seinen Ansatz auf alle Kinder aller Sprachen und berücksichtigt somit nicht individuelle Entwicklungen einzelner Kinder. Die Überlegungen von Piske (2001) und Vihman und Croft (2007) schreiben dem Kind hingegen im Gegensatz zu den bisher beispielhaft unter dem Aspekt dieses Abschnitts vorgestellten Ansätzen eine stärkere eigene Komponente in Form von sog. Schablonen (Vihman, Croft 2007) oder sog. Mustern (Piske 2001) zu. Diese werden jeweils als häufige Kombinationen von Merkmalen in den Lautformen eines einzelnen Kindes bestimmt (Vihman, Croft 2007, Piske 2001) und sind somit individuell. Diese Überlegungen stehen nicht im Widerspruch zur Annahme einer evolutionären Sicht der Sprachentwicklung, weil sie den Schwerpunkt auf systematische Aspekte in der Sprache des Kindes verlagern und seine Eigenständigkeit nicht ignorieren und auch dem einzelnen Kind eine individuelle Sprachstruktur zugestehen. Dagegen sehen beide Ansätze die kindlichen Lautformen zu stark auf Zielsprachformen bezogen, indem sie ihre Bildung unmittelbar aus der Zielsprache ableiten (vgl. Abschnitt 2.1.). Ausführlicher werden die Ansätze dieses Abschnitts (bis auf den von Stampe (1979)) im Kapitel 6 vorgestellt und diskutiert.

4.5.2. Verhältnis des kindlichen Silbensystems zur Zielsprache im Ansatz des organischen Silbenwachstums

In dem hier vertretenen Ansatz des organischen Silbenwachstums entlehnt das Kind nur insofern Lauteindrücke, als dieselben seinem System angemessen sind. Angemessen bedeutet dabei, dass das Kind in der Regel nur Silben produziert, die im Zusammenhang des organischen Wachstums seines Silbensystems möglich sind. Dies beinhaltet, dass sie alle damit implizierten Regelmäßigkeiten und Zusammenhänge erfüllen müssen. So stellen die zu einem Zeitpunkt von einem Kind geäußerten Silben die Ausschöpfung der artikulatorischen Möglichkeiten dar, die aber gleichzeitig mit ihrer zunehmenden Ausnutzung erweitert werden. Außerdem erfüllen diese Silben als Ergebnis des organischen Wachstums die funktionalen Anforderungen nach Eindeutigkeit durch Unterscheidbarkeit von allen anderen Silben und die der Ökonomie, die sich in den Besonderheiten der Schnittmengenbildung zeigt. Die Entwicklung des kindlichen Sprachsystems allgemein und des in der Silbenbildung materialisierten Lautsystems ist also weder vollkommen autonom, da es sich nach seinen perzeptiven Möglichkeiten und auf der Grundlage seiner jeweiligen produktiven Fähigkeiten am Zielsprachsystem orientiert, noch ist es zufällig, da es sich nach inneren Gesetzmäßigkeiten entwickelt analog zum organischen Wachstum in der belebten Natur.

4.6. Das Verhältnis der Perzeption und der Produktion im Phonemerwerb

In diesem Abschnitt sollen die perzeptiven Fähigkeiten des Kindes ab dem Zeitpunkt seiner Geburt dargestellt und anschließend nach dem Verhältnis derselben zur Produktion gefragt werden.

4.6.1. Die perzeptiven Fähigkeiten des Kindes seit seiner Geburt

Kinder können sehr früh direkt nach der Geburt oder einige Monate später eine große Menge von Phonemkontrasten in ihrer eigenen wie auch in einer fremden Sprache unterscheiden (Jusczyk 2000: 50ff.). So erkannten beispielsweise Englisch lernende ein- und viermonatige Kinder die im Englischen vorkommenden Unterschiede zwischen den CV-Silben /ba/ vs. /pa/ mit stimmhaftem Plosiv gegenüber stimmlosem Plosiv (Eimas et al. 1971 in Jusczyk 2000: 50f.) und neben diesen viele weitere Kontraste innerhalb von Silben (Jusczyk 2000: 51ff., Curtin, Hufnagle 2009: 108ff.). Zudem hielten Kanadisches Englisch lernende Kinder im Alter von ein bis vier Monaten den nicht im kanadischen

Englisch vorkommenden Kontrast nasal vs. oral für den Vokal /a/ in Silben wie *pa* auseinander (Trehub 1976 in Jusczyk 2000: 55).

Auf Grund solcher sehr frühen Lautunterscheidungsfähigkeiten von Kindern, insbesondere derjenigen, die direkt nach der Geburt getestet wurden, ist die Frage, ob diese Fähigkeiten angeboren sind oder durch Erfahrung gelernt werden, noch nicht entschieden worden (Jusczyk 2000). Sollten sie angeboren sein, so bleibt aber fragwürdig, ob es sich um sprachspezifische Verarbeitungsvorgänge handelt, wie sie beispielsweise in der *motor theory* der Sprachperzeption (Lieberman et al. 1985) angenommen werden. Gegen diese Auffassung spricht u.a., dass der Mensch die kategoriale Wahrnehmung auch beim Hören musikalischer Stimuli zeigt (Fitch 2010: 99) und dass Tiere ebenfalls sowohl ihre eigenen stimmlichen Äußerungen als auch die des Menschen kategorial wahrnehmen (Fitch 2010: 99). Außerdem scheinen Kinder noch im Mutterleib die Akustik der Umgebung zu lernen, da sie beispielsweise bei der Geburt Geschichten, die von der Mutter während ihrer Schwangerschaft vorgelesen wurden, solchen vorziehen, die während dieser Zeit nicht zu hören waren und somit neu sind (Curtin, Hufnagle 2000: 108). Es ist somit anzunehmen, dass die perzeptiven Fähigkeiten des Kindes entweder das Ergebnis vorgeburtlicher Hörerfahrungen sind oder aber allgemeine Perzeptionsmechanismen darstellen, die sie mit anderen Tieren teilen. Der Einfluss der Zielsprache macht sich besonders gegen Ende des ersten Lebensjahres bemerkbar, indem die Unterscheidungsfähigkeit für nicht-native Kontraste zunehmend verlorengelht wie beispielsweise den oben genannten der Nasal-Oral-Unterscheidung von Vokalen durch Kinder mit kanadischem Englisch als Zielsprache (Curtin, Hufnagle 2009: 110). Dies kann zum einen als Resultat der Erfahrung mit der Zielsprache interpretiert werden, welche zu einer Intensivierung der Wahrnehmung der in der Zielsprache vorkommenden Lautregelmäßigkeiten führt (Kuhl 2000: 11853).

Eine weitere Erklärung könnte darin gesucht werden, dass Kinder in den ersten Monaten nach der Geburt eine größere Offenheit und Sensibilität gegenüber verschiedensten Lauten besitzen könnten, die aber mit dem Beginn des Sprechens und damit dem Beginn der Bedeutsamkeit von Sprachlauten insofern eingeschränkt wird, als die Sprachlautrezeption nun immer stärker über das eigene Lautsystem des Kindes interpretiert und gefiltert wird. Die Schwerpunkte in seinem Lautsystem lassen sich dabei über die Schnittmengenbildungen identifizieren.

4.6.2. Die Produktion des Gehörten durch das Kind

Um das Verhältnis der Produktion des Kindes zu seiner Perzeption zu klären, muss beantwortet werden, wie das Kind gehörte sprachliche Äußerungen perzeptiv speichert und wie es bei der Produktion auf diese zurückgreift. Zunächst sollen hier einige theoretische Ansätze vorgestellt werden, die auf der Annahme des direkten Wortlernens basieren und auf dieser Grundlage das Verhältnis zwischen der Perzeption und der Produktion eines Kindes zu klären versuchen. In diesem Zusammenhang soll der eigene Standpunkt entwickelt werden. Diese Ansätze haben gemeinsam, dass sie das direkte Lernen von Wörtern der Zielsprache durch das Kind stark in den Mittelpunkt rücken, anstatt einen Rückgriff des Kindes auf allgemeine, ihm durch die Perzeption zur Verfügung stehende Regelmäßigkeiten zu berücksichtigen. Die Annahme einer unmittelbaren Beziehung zwischen den Lautformen des Kindes und denen in der Zielsprache ist nicht unproblematisch¹⁷, da der Spracherwerb des Kindes als regelhafter Mechanismus zur Tilgung von Fehlern verstanden wird, wobei der Aspekt eines eigenständigen sprachlichen Produktionssystems verloren geht (vgl. Abschnitt 2.1.).

Smith (2010: 105) nimmt an, dass dem Kind für seine Produktionen exakt die Phonemfolgen von Äußerungen der Erwachsenen dienen, wie sie ein geschulter Hörer mittels des IPA's transkribieren würde. Dies begründet er (Smith 2010: 105f.) u.a. mit der Tatsache, dass das von ihm beobachtete Kind die englischen Wörter *trane* und *crane* auf die gleiche Weise ausspreche, aber die zu ihnen gehörigen Referenten richtig durch Auswählen derselben identifiziere (ebd.). Dies ist aber kein zwingender Grund, um die Speicherung von durch das Kind gehörten Lautformen als Phonemsequenzen anzunehmen. Sofern die Phoneme nicht als angeboren angesehen werden, wie Smith (2010: 3) dies tut, sind sie im hier vorgestellten Ansatz Schnittmengen aus den bereits vom Kind geäußerten und wahrgenommenen Silben, die detailliert und multimodal gespeichert werden. Phoneme werden in diesem Ansatz als eine Größe des organischen Silbensystems eines einzelnen Sprechers verstanden, die nicht unabhängig von demselben existieren, sondern einen Systemzustand andeuten. Somit sind sie das Resultat der Entwicklung des Silbensystems des Kindes, welches über den kausalen Motor angetrieben wird, der ständig Lautformen zur Konzeptunterscheidung erfordert. Man könnte die oben angeführten Perzeptionsexperimente mit sehr jungen Kindern oder auch Neugeborenen als Argument

¹⁷ Siehe dazu die Abschnitte 4.7. und 2.1.

für den aufgeführten Standpunkt von Smith (2010) ansehen, dass die Speicherung der Zielsprachäußerungen in Form von Phonemfolgen abläuft. Doch eine Unterscheidung von Minimalpaarsilben wie beispielsweise /ba/ und /pa/ durch ein einmonatiges Kind muss nicht bedeuten, dass das Kind die Phoneme /b/ und /p/ hinsichtlich der Opposition stimmlos-stimmhaft unterscheidet (Eimas et al. 1971 in Jusczyk 2000: 50f.), sondern dies könnte ebenso als Beweis dienen, dass das Kind dazu in der Lage ist, die entsprechenden Silben auseinanderzuhalten (Jusczyk 2000: 70f., Drescher 2009: 204f.), wobei einschränkend in Frage gestellt wird, dass das Kind in diesem Alter überhaupt Sprachdifferenzen unterscheidet.

Ausgehend von seiner Annahme, dass das Kind die zielsprachlichen Äußerungen in Phonemsequenzen aufnimmt (Smith 2010: 105), fasst Smith (2010: 109) die Äußerungen des Kindes lediglich als das Ergebnis eines unreifen neuronalen Mechanismus auf, der mit den korrekt gegebenen Phonemfolgen operiert und die Artikulatoren falsch steuert (ebd.). Dies ist aber eine nicht haltbare Vorstellung des Verhältnisses von Produktion und Perzeption des Kindes. Denn sie wird dem Wesen des Phonems als dynamische Schnittmengengröße des Silbensystems nicht gerecht, weil sie dasselbe als bereits vorhanden annimmt. Diese Annahme ermöglicht Smith (2010: 109) erst, die Produktionen des Kindes durch einen die Artikulatoren falsch steuernden neuronalen Produktionsmechanismus zu erklären (ebd.). Sein Modell wird auch dem Produktions-Perzeptions-Verhältnis des Kindes beim Spracherwerb nicht gerecht, weil es nicht alleine die Produktion ist, die sich entwickelt, sondern auch die Perzeption. Beide entwickeln sich miteinander, wobei in dem Ansatz dieser Arbeit angenommen wird, dass die Perzeption der Produktion zeitlich gering vorausläuft, was durch die im vorhergehenden Abschnitt angeführten Perzeptionsexperimente von Kindern, die noch nicht sprachlich artikulieren, unterstützt wird. Ebenso bleibt bei Smith (2010) ungeklärt, wie sich ein neuronaler Steuerungsmechanismus für die Artikulatoren entwickelt haben könnte und wie das Kind die Kenntnis des Phonemsystems erwirbt, die er voraussetzt. Das anzunehmende Verhältnis von Produktion und Perzeption wird auf andere Weise in dem von Vihman (2009: 20ff.) konzipierten "artikulatorischen Filter" (ebd.) modelliert, nach dem "the child's familiarity with his or her own vocal production patterns makes sequences in input speech that are like those patterns particularly salient." (Vihman 2009: 21). Dabei bezieht Vihman (2009: 22) diesen Effekt durch den von ihr sog. artikulatorischen Filter auf das Lernen zielsprachlicher Wörter, indem sie annimmt, dass gerade solche Wörter der Zielsprache gelernt werden, die den individuellen Produktionen des Kindes aus der Plapperphase besonders ähnlich sind

(ebd.). Die Auswahl des Kindes von perzeptiven Eindrücken für seine eigene Produktion sollte aber nicht so sehr auf die einzelnen zielsprachlichen Worteinheiten beschränkt sein (vgl. Abschnitt 2.1.), sondern mehr davon ausgehen, dass das Kind auf Grund seiner Silbenunterscheidungsfähigkeiten und Lernmechanismen wie dem statistischen Lernen (Thiessen 2009) aus dem zielsprachlichen Signal extrahierte Regelmäßigkeiten in seinen Produktionen als ungefähre akustische Richtungen einfließen lässt. Dabei wird *ungefähre akustische Richtungen* als Begriff verwendet, um zu betonen, dass das Kind nicht auf Phonemfolgen zurückgreift, sondern auf die sich aus dem organischen Silbensystem und den Silbenschnittmengen ergebenden Regelmäßigkeiten. Die Zuordnung von Bedeutungen zu den vom Kind geäußerten Lautformen ist nämlich problematisch, wie im folgenden Abschnitt diskutiert wird, und im Extremfall gibt es das Phänomen der Protowörter (Piske 2001: 77ff.), die dem Kind in Einzelfällen als Wörter ohne zielsprachlich semantisch und lautliches Pendant zugestanden werden.

4.7. Problematische Aspekte der Zuordnung von kindlichen Lautformen zu Zielsprachwörtern

Es wird in dieser Arbeit angenommen, dass alle Lautformen des Kindes ab dem Beginn der eigentlichen Sprachproduktion bedeutungsvoll sind¹⁸. Im Folgenden soll der Begriff der *Bedeutung* für Konzepte und Wortbedeutungen verwendet werden, weil durch den Begriff des Konzeptes die Existenz von etwas Bedeutungsvollem vorausgesetzt wird, das wiederum die Wortbedeutung festlegen soll, aber die Entstehung dieses Bedeutungsvol-

¹⁸ Die in Diskurstheorien zum Spracherwerb vorzufindende Vorgehensweise, dass kindliche Lautformen zu Zielsprachformen in Beziehung gesetzt und die als "echte Fehler" (Elsen 1999: 22, 207) wahrgenommenen Abweichungen des Kindes z.B. durch nicht korrekte Arbeitsweisen eines neuronalen Netzwerkes (vgl. Stemberger (1992), Elsen (1999), u.a. Abschnitte 2.1.) erklärt werden (vgl. auch Kapitel 2), verleitet dazu, unverständliche Äußerungen des Kindes als Äußerungen ohne Bedeutung oder "Nonsenssätze und Ausdrücke" (vgl. Elsen 1999: 136, 255) aufzufassen. Die in Elsen (1999: 136, 255) angeführten Beispiele von Formulierungen des Kindes, die ohne Sinn sein sollen (Punkt "H Nonsenssätze und -ausdrücke" (ebd.) des Anhangs (ebd.)), beziehen sich auf "formal korrekte Sätze, die aber inhaltlich verworren sind, daneben Nonsensfüllsel wie *(a)popo* und Phantasiewörter und schließlich eine selbstkreierte Redewendung (ä.)." (Elsen 1999: 255). Vom Beobachter angenommene inhaltliche Verworrenheit der Äußerungen des Kindes, nicht einzuordnende Lautformen ("Nonsensfüllsel" (ebd.)) oder Eigenkreationen werden in dem auf die Produktivität des Kindes und nicht die Erreichung von zielsprachlichen Normzuständen gerichteten Ansatz dieser Arbeit nicht als bedeutungslose Spielereien oder Unsinn angesehen, sondern als bedeutungsvoll. Auf eine konkrete Bestimmung von Bedeutungen einzelner Lautformen muss allerdings verzichtet werden, da die semantischen Schnittmengen zwischen der kindlichen Sprache und der Zielsprache nicht bestimmt werden können. Lautspielereien und „unsinnige“ Lautformen können aber auch ein Ausdruck des sich entwickelnden phonematischen Bewusstseins auf der Grundlage hinreichend großer Silbenschnittmengen sein (Vgl. dazu die Ausführungen zum Verhältnis der phonematischen Bewusstheit und dem Lesenlernen am Ende von Abschnitt 4.2.1.4.)

len nicht klärt¹⁹. Es wird davon ausgegangen, dass jeder Sprachteilnehmer sein eigenes, individuelles System von Bedeutungen hat, dass sich von dem aller anderen Sprachteilnehmer unterscheidet. Daraus folgt insbesondere, dass die Bedeutungen der Wörter des Kindes keine Identitäten, sondern nur Schnittmengen mit für sie angenommenen Zielsprachwortbedeutungen haben. Deshalb sollen im Gegensatz zu den meisten theoretischen Ansätze zur Erklärung des Phonemerwerbs die Lautformen nicht nach konkreten Zielsprachwörtern gruppiert werden. Im folgenden, ersten Abschnitt werden beispielhaft Untersuchungen angeführt, die als Hinweis auf das Vorliegen dieses eigenen und individuellen Bedeutungssystems gewertet werden. Ansätze zum Spracherwerb, die jeweils mehrere kindliche Lautformen nach für sie angenommenen Bedeutungen gruppieren, sollen in Analogie zu den in Kapitel 2 hinsichtlich der Formseite vorgestellten Diskurstheorien als *bedeutungsbezogene Diskurstheorien* zum Spracherwerb bezeichnet werden.

Die mit diesen Theorien verbundene Problematik, die u.a. aus den im ersten Abschnitt vorgestellten Annahmen begründet ist, wird im zweiten Abschnitt diskutiert. Im dritten und letzten Abschnitt werden schließlich einige Argumente für den Wortansatz zum Phonemerwerb kritisch besprochen.

4.7.1. Hinweise auf das eigene und individuelle Bedeutungssystem eines Kindes

Es wird angenommen, dass die Bedeutungen der Wörter unterschiedlicher Sprachteilnehmer keine Identitäten, sondern nur Schnittmengen mit den Wörtern anderer Sprachteilnehmer gemeinsam haben, weil davon ausgegangen wird, dass jeder Sprachteilnehmer

¹⁹ So kommt es u.a. zum sog. "symbol grounding problem" (Sanford 2006: 154), das beinhaltet, dass die Konzepte, durch die die Semantik von Wörtern bzw. Lautformen begründet werden soll, selbst wieder durch Wörter festgelegt werden (ebd.) wie dies für die herkömmlichen in Geeraerts (2010) präsentierten Semantiktheorien der Fall ist. Die Herkunft dieser die Konzepte definierenden Wörter bleibt dann aber letztlich ungeklärt.

Es gibt Versuche, dieses "symbol grounding problem" zu lösen: Entsprechend dem Begriff der Simulation versteht ein Sprecher Wörter seiner Muttersprache, indem er neuronale Zustände aktiviert, wie sie auch bei der tatsächlichen Wahrnehmung der durch diese Wörter bezeichneten Objekte hätten gemessen werden können (Borghini 2010: 1). Diesen Gedanken integriert Barsalou (1999) in seine Theorie der Bildung von Konzepten. Konzepte werden durch das Zusammentreffen mit konkreten Objekten oder das Ausführen wirklicher Handlungen als sog. Rahmen gebildet und durch das Wirken der selektiven Aufmerksamkeit im Zusammenspiel mit neuen konkreten Objekten oder Handlungen weiter spezifiziert (Barsalou 1999:583f., 586, 590f.). Durch diese Simulations-Definition von Konzepten können sie zwar zum einen individuell sein und müssen nicht mehr über sprachliche Symbole erklärt werden, aber ihre genaue Entstehung bleibt ungeklärt und der für die Bedeutungen eines Individuums anzunehmende Systemcharakter unberücksichtigt. Es bleibt zu untersuchen, ob die Verwendung des Konzeptbegriffs für die Beschreibung des Bedeutungssystems eines Individuums ebenso wenig hilfreich für die Erklärung der Entstehung der Wortbedeutung ist wie die Verwendung des Musterbegriffs für die Erklärung der Phonementwicklung (siehe Kapitel 6.).

über ein individuelles und eigenes Bedeutungssystem verfügt. Die beiden im Folgenden aufgeführten Studien werden als Hinweise auf dieses individuelle Bedeutungssystem beim Kind gewertet, auch wenn die Untersuchung des Bedeutungssystems mit Hilfe von Merkmalen und Konzepten als unpassend angesehen wird (s.o.): Kinder teilen mit den von ihnen verwendeten Wörtern ganze Objektbereiche systematisch nach für sie bedeutungsvollen Merkmalen auf und bilden eigene Konzepte, so beispielsweise die Abgrenzung von "ei- oder kugelförmige[n] Nahrungsmittel[n] einer bestimmten Größe" (Elsen 1995: 235) mit dem Wort *Ei* gegen Objekte von gleicher Gestalt, die aber nicht essbar sind, mit dem Zielsprachwort *Ball* durch ein 10 Monate altes, Deutsch lernendes Kind (ebd.: 235ff.). Die Konzepte müssen dabei wie dem von "ei- oder kugelförmige[n] Nahrungsmittel[n] einer bestimmten Größe"(ebd.: 235) nicht in der Zielsprache durch ein Lexem kodiert sein (ebd.: 238). Die Unterscheidung der von Kindern adaptierten Zielsprachwörter nach eigenen, für sie bedeutsamen Merkmalen, zeigt sich in einer Studie mit Mandarin Chinesisch lernenden 3, 5 und 7 Jahre alten Kindern (Saji et al. 2011):

Das chinesische Mandarin hat ca. 20 verschiedene Wörter für die Handlung, die im Englischen normalerweise mit *carry* (dt. tragen) ausgedrückt wird (ebd.: 48). Diese chinesischen Wörter werden nach der Art, wie ein Objekt getragen wird, voneinander unterschieden (ebd.), so beispielsweise danach, ob etwas mit beiden Händen, auf dem Rücken, auf dem Kopf oder unter dem Arm getragen wird (ebd.: 50). Dabei werden im Chinesischen diese verschiedenen Varianten für das Wort *tragen* mit bestimmten Objektkategorien häufiger verwendet als mit anderen (ebd.: 49), obwohl der Gebrauch der *tragen*-Wörter in allen Situationen stattfinden darf, in denen ein Objekt in Relation zum durch das *tragen*-Verb bezeichneten korrekten Körperteil transportiert wird (ebd.). So ist z. B. der Begriff *Schale* in diesem Zusammenhang sehr häufig mit der Variante des Verbs *tragen* verknüpft, die mit der Vorstellung eines Transports von Gegenständen auf dem Kopf (ebd.) verbunden ist. In ihrem Experiment zeigten Saji et al. (2011: 54), dass die Kinder im Gegensatz zu den Erwachsenen die *tragen*-Verben anstatt nach der in der Zielsprache verwendeten Körperteildifferenzierung nach der häufig mit ihnen verwendeten Objektkategorie differenzierten (ebd.). Das Kind produziert also Lautformen, die denen von Zielsprachwörtern ähnlich sind. Die kindlichen Produktionen weichen aber semantisch und lautlich von der Zielsprache ab, weil sie einen eigenen Status haben, der sich aus der Stellung der Produktionen in ihrem System auf der Laut- und Bedeutungsebene ergibt. Hinsichtlich der Bedeutungsseite zeigt sich dieser in der Notwendigkeit der Beschreibung der Adaption der Zielsprachwörter durch das Kind mit für die Zielsprache ungewöhnlichen

Merkmale (z.B. "ei- oder kugelförmige[n] Nahrungsmittel[n] einer bestimmten Größe"(Elsen 1995: 235)). Die Produktionen des Kindes werden im Ansatz dieser Arbeit dagegen als Indizien für ein eigenes, individuelles Bedeutungssystem gesehen, das nur Schnittmengen mit denen anderer Sprachteilnehmer gemeinsam hat.

4.7.2. Bedeutungsbezogene Diskurstheorien zum Spracherwerb und problematische Aspekte der Gruppierung jeweils mehrerer kindlicher Lautformen nach Zielsprachwörtern

Im vorangehenden Abschnitt wurden Beobachtungen aufgeführt, die als Hinweise auf ein eigenes und individuelles Bedeutungssystem eines Kindes gewertet werden. Aus dieser Annahme ergibt sich die Folgerung, dass Bedeutungen unterschiedlicher Sprachteilnehmer nur Schnittmengen miteinander gemeinsam haben können, aber keine Identitäten aufweisen. Deshalb ist eine Vorgehensweise unzulässig, die mehrere unterschiedliche Lautformen des Kindes nach einem für sie angenommenen Zielsprachwort gruppiert, weil sie implizit die Identität von Bedeutungen unterschiedlicher Sprachteilnehmer unterstellt, die nach den vorhergehenden Ausführungen nicht gegeben sein kann. Oben wurden Ansätze zum Spracherwerb, die jeweils mehrere kindliche Lautformen nach für sie angenommenen Bedeutungen gruppieren, als *bedeutungsbezogene Diskurstheorien* zum Spracherwerb bezeichnet. Die Analogien zu den im Kapitel 2 hinsichtlich der Formseite definierten Diskurstheorien zum Spracherwerb zeigen sich in der auch hier durchgeführten Betrachtung der bedeutungsbezogenen Abweichungen von Lautformen des Kindes von denen von Zielsprachwörtern, als deren Realisierungen sie interpretiert werden:

So nimmt das Kind in der sog. *Überdehnung* systematisch auf Objekte in der Umgebung Bezug, die über die Extension des Zielsprachwortes hinausgehen (Rothweiler, Meibauer 1999: 17), indem es z.B. alle Vierbeiner mit *Hund* (ebd.) bezeichnet. Bei der sog. *Unterdehnung* wird demgegenüber nur auf eine kleinere Teilmenge der Extension des Zielsprachwortes verwiesen (ebd.), so z.B. bei der Verwendung des Wortes *Hund* nur auf braune Hunde (ebd.). U.a. mit Hilfe solcher Klassifikationen der bedeutungsbezogenen Abweichungen des Kindes wird versucht, die Bedeutungsentwicklung des Kindes zu beschreiben und zu erklären. Der Erwerb von zielsprachlichen Wörtern durch das Kind ist aber mehr als nur ein allmählicher Angleich von einzelnen Wörtern des Kindes an eine für sie angenommene Zielsprachwortbedeutung und -form. Denn die Produktionen des Kindes müssen als eigenständig innerhalb eines individuellen Silben- und Bedeutungssystems aufweisend aufgefasst werden (s.o).

Diese Möglichkeit wird Kindern jedoch bei der Untersuchung ihrer Äußerungen nur in sehr geringem Maße in Form der sog. *Protowörter* (Piske 2001: 78, Stoel-Gammon 2010: 2f.) zugestanden. Diese werden durch einen stabilen Wort-Bedeutungs-Zusammenhang identifiziert (Stoel-Gammon 2010: 2f.) und es kann ihnen keine Zielsprachlautform zugeordnet werden, die als Vorlage bei der Produktion dienen könnte (Piske 2001: 78). So produzierte ein von Piske (ebd.) untersuchtes Kind /naŋnaŋ/, "wenn [...] [es] ausdrücken wollte, daß sich jemand beeilen sollte" (ebd.). Die meisten der vom Kind produzierten Lautformen werden jedoch in der Mehrzahl der theoretischen Ansätze zur Erklärung des Phonemerwerbs zu einem Zielsprachwort in Beziehung gesetzt. Damit eine vom Kind produzierte Lautform "als eine Variante eines bestimmten Zielworte[s] gewertet" (ebd.) wird, muss sie erstens eine gewisse Ähnlichkeit zu der Lautform des Zielsprachwortes aufweisen (Stoel-Gammon 2010: 3), zweitens in einem für das Zielsprachwort passenden Kontext verwendet werden (ebd.) und drittens relativ stabil in der Aussprache sein (ebd.). Auf Grund dieser Kriterien werden verschiedene Lautformen des Kindes in der Regel einem einzigen Zielsprachwort zugeordnet. So werden in einem bekannten Beispiel von Ferguson und Farwell (1975 in Piske 2001: 61) die folgenden zehn Wörter eines englischsprachigen Kindes als Ausspracheversuche des englischen Wortes *pen* aufgefasst und zusammengruppiert:

Tabelle 1: Zehn dem Zielsprachwort *pen* zugeordnete kindliche Lautformen (Ferguson, Farwell 1975 in Piske 2001: 61)

/mãǎ/	/ỹÑ/	/dɛďň/	/hɪn/	/m̃bõ/
/p^hɪn/	/t^hŋ t^hŋ t^hŋ/	/ba^h/	/d^hauň/	/buã/

(entnommen aus Piske 2001: 61 in abgeänderter Form)

Neben der Missachtung des bereits angesprochenen eigenen, individuellen Bedeutungssystems soll im Folgenden auf weitere problematische Gesichtspunkte dieser Gruppierungen hingewiesen werden: Erstens führt eine solche Gruppierung dazu, dass in der Untersuchung die verschiedenen kindlichen Lautformen nicht mehr weiter als einzelne Lautformen betrachtet werden, sondern dass für sie beispielsweise neuromotorische Routinen angenommen werden. Im Gegensatz dazu wird in dem Ansatz dieser Arbeit jede Lautform als eigene sprachliche Identität betrachtet und berücksichtigt, insofern sie neue Silben enthält, weil diese Rückschlüsse auf das Silbensystem zulassen. Zweitens sind die Kriterien für die Wortidentifikation sehr ungenau und unterliegen letztendlich der Ent-

scheidung des Untersuchenden. Stabilität könnte sich auf den zeitlichen Rahmen und die Ähnlichkeit von Lautformen beziehen, aber das zugelassene Ausmaß derselben bleibt im Verborgenen.

4.7.3. Aus der Gruppierung von kindlichen Lautformen sich ergebende Phänomene als Gründe für die Annahme des Wortes als grundlegender Einheit im Phonemerwerb

Das Paradoxe an den Argumenten für den Ansatz, das Wort oder das artikulatorische Muster als die grundlegende Einheit im Phonemerwerb annehmen zu müssen, liegt darin, dass sie sich gerade aus dem Vorgehen ergeben, möglichst alle vom Kind geäußerten Lautformen zu einer der zielsprachlichen Lautformen in Beziehung setzen zu wollen. Dadurch entsprechen einer Zielsprachform mehrere verschiedene kindliche Lautformen. Dieses Verhältnis wird dann u.a. als "phonetische Veränderlichkeit" der kindlichen Lautformen bezeichnet und so interpretiert, dass das Kind zwar die Zielsprachform zu äußern bestrebt sei, aber diese nur ungenau realisiere. Diese "Ungenauigkeit", die eigentlich als produktive Wortbildungsleistung des Kindes aufgefasst werden sollte, kann dann auf Grund der Unterschiedlichkeit der zusammengruppierten kindlichen Lautformen nur noch im Rahmen des Wortes stattfinden, wie die Logik der folgenden für das Wort als grundlegende Einheit angeführten Argumente zeigt: So nennt Piske (2001: 61ff.) die auch in anderen wortbezogenen Ansätzen ausgemachten Phänomene der sog. *prosodischen Variabilität* und der sog. *lexikalisch begrenzten segmentorientierten Kodierung*. Unter dem erstgenannten Begriff wird das Phänomen verstanden, dass das Kind scheinbar für ein und dasselbe Zielsprachwort mehrere verschiedene Lautformen äußert, die alle jeweils Merkmale des Zielsprachwortes enthalten (ebd.). Diese Übereinstimmungen werden so interpretiert, dass das Kind die Merkmale des Zielsprachwortes schon in einer Art Sammelkasten (*box*) gespeichert hat, diese aber noch nicht in der richtigen Reihenfolge "als stabile Kette von Segmenten" (Piske 2001: 62) wiedergeben könne (ebd.), d.h. das Wort wird als Kodiereinheit angenommen. Diese Beobachtung von nicht in der "richtigen" Reihenfolge erscheinenden Phoneme ergibt sich aber erst aus der Gruppierung unterschiedlicher Lautformen des Kindes nach einem einzigen für sie angenommenen Zielsprachwort. Unter der sog. "lexikalisch begrenzte(n) segmentorientierte(n) Kodierung" (Piske 2001: 62ff.) versteht man das Phänomen, dass das Kind ein Phonem wie z.B. /b/ an gleicher Position in unterschiedlichen Zielsprachwörtern wie *Bauch*, *Banane*, *Ball* und *bitte* (Piske 2001: 63) unterschiedlich wiedergibt und man somit nicht eine bloße Substi-

tutionsregel für ein solches Phonem aufstellen kann, wie es in generativen Beschreibungen des Spracherwerbs normalerweise geschieht. Deshalb geht man davon aus, dass das Kind die Phoneme nicht wortübergreifend, sondern bezogen auf individuelle einzelne Wörter kodiert und sieht dies als Argument für das Wort als grundlegende Kodiereinheit im Spracherwerb an (ebd.). Erst durch die Zuweisung der kindlichen Lautformen zu Zielsprachlautformen kann behauptet werden, das Kind variere ein Zielsprachphonem an identischen Positionen in unterschiedlichen Zielsprachlautformen durch verschiedene Phoneme. Es soll hier nicht behauptet werden, das Kind probiere nicht auch verschiedene Artikulationen für einen akustischen Gehöreindruck eines Zielsprachwortes aus. Bei Lautproduktionen des Kindes, die vom Beobachter als Versuch verstanden werden, Zielsprachwörter zu artikulieren, ist ein phonematisches Bewusstsein oder als dessen Vorstufe eine phonematische Sensibilität zumindest für Teilbereiche des phonologischen Systems jedoch eine Voraussetzung. Daraus kann jedoch kein Argument für den Wortansatz abgeleitet werden, da es sich als Phänomen erst aus der Gruppierung der kindlichen Lautformen ergibt.

4.8. Die Rolle der Funktionalität in Lindbloms (1999, 2000) Emergenzphonologie

Lindblom (1999, 2000) stellt mit seinem Ansatz der sog. "Emergenzphonologie" (engl. "emergent phonology" (Lindblom 1999)) einen Phonemerwerbsansatz vor, in dem der Begriff der Funktionalität von zentraler Bedeutung ist, die vor allem in der Minimierung von Energie gesehen wird, die zur Erzeugung von Sprache notwendig ist (ebd.). Das Phonem wird nicht als angeboren verstanden, sondern als während des Spracherwerbs entstehend (Lindblom 2000: 302). Voraussetzung für das Erlernen der Zielsprache durch die Kinder ist das Vorhandensein einer ausreichenden Menge unterschiedlicher Lautformen, die vom Kind perzeptiv als phonetische Ganzheiten in einem Exemplarspeicher gesammelt werden und entsprechend ihrer Ähnlichkeit wie im Clustering-Vorgang des maschinellen Lernens zusammengruppiert würden (Lindblom 2000: 303f.). Dies gewährleistet dann die Filterung der Phoneme bzw. ihrer phonetischen Korrelate durch das Kind (ebd.). Da die Zielsprache ausreichend voneinander unterschiedene Lautformen entwickelt habe (Lindblom 2000: 304), ergäben sich durch die Gruppierung die phonetischen Kategorien der Zielsprache (Lindblom 2000: 303f.). Die Modalitäten dieser Gruppierung bleiben jedoch unklar.

Die Energieminimierung spielt²⁰ für den Phonemerwerb des Kindes hingegen in der Lindblomschen (2000) Emergenzphonologie eine erheblich größere Rolle. Ein "criterion of minimum energy consumption" (Lindblom 2000: 302) soll zum einen eine Grundlage für die konkreten Äußerungen von Sprachteilnehmern und insbesondere die des Kindes sein und die Formen der Äußerungen mitbestimmen (Lindblom 2000: 304ff.) und zum anderen die Speicherung von Lautformen regeln (Lindblom 2000: 308ff.). Lindblom (2000: 306) setzt voraus, dass sich Bewegungen jeder Art hinsichtlich ihrer energieminimierenden Tendenz vergleichen lassen (ebd.). Deshalb führt er (Lindblom 2000: 305; 1990: 413f.) eine Untersuchung der Bewegung von Pferden an, bei der sie auf einem Laufband laufen mussten, dessen Geschwindigkeit ständig erhöht wurde (ebd.). Gleichzeitig wurde ihr Energieverbrauch als Funktion der Geschwindigkeit aufgezeichnet, der eine Reihe lokaler Minima aufwies (ebd.), die mit den Gangartenwechseln des Pferdes korrespondierten, das vom Schritt zum Trab und dann zum Galopp wechselte und damit in der Tendenz den Energieverbrauch minimierte (ebd.).

Analog hat Lindblom (2000: 306) bei der Modellierung der Kieferszillation eine Kurve des Energieverbrauchs mit einem Minimum berechnet (ebd.). Auch bei phonologischen Assimilationen sieht Lindblom (1983: 237) das Prinzip des minimalen Energieverbrauchs verwirklicht (ebd.). Beim Phonemerwerb soll dieses Energieminimierungsprinzip für das Kind nun die folgende Funktion erfüllen: Gehöreindrücke lassen sich auf unterschiedliche Weisen artikulieren, auch "degree of freedom problem" (Lindblom 2000: 305) genannt, was Lindblom (2000: 306) an einem Ausgleich ("trade off" (Lindblom 2000: 306)) unterschiedlicher Artikulatoren illustriert (ebd.: 306): Findet beispielsweise eine Behinderung des Kiefers in seiner natürlichen Bewegung für ein Phonem statt, so kann die Zunge gegensteuern, um den angestrebten Gehöreindruck trotzdem zu erreichen (ebd.: 306). Es wird angenommen, dass ein allgemeines evolutionäres Prinzip der Energieminimierung für Bewegungen und speziell in Bezug auf die Artikulation von Sprache dem Kind hilft, aus den vielen Möglichkeiten einen Gehöreindruck durch Bewegungen der Artikulationsorgane zu realisieren ("degree of freedom problem" (Lindblom 2000: 305)), die richtigen auszuwählen, d.h. die in der Zielsprache auch vorkommenden (Lindblom 2000: 307f.).

Denn das Kind artikuliere energieminimierend (Lindblom 2000: 307f.) und werde sodann durch die Gehöreindrücke der Zielsprache bestärkt (ebd.), die auch hauptsächlich ener-

²⁰ im Gegensatz zur Unterscheidbarkeit

giesparenden Artikulationen entsprechen (ebd.), weil die Zielsprache eine evolutionäre Adaption an diese Energieminimierung sei (ebd.: 308). Nähme man an, dass es wie beim Pferd drei ideale Bewegungsarten gibt oder auch nur einen energieminimierenden Oszillationsmodus des Kiefers, so bleibt dennoch unklar, wie diese energiesparenden Modi oder das Energieminimierungsprinzip die Äußerungen des Kindes im Detail bedingen könnten. Neben der Justierung der richtigen, in der Zielsprache verwendeten Artikulationen, soll das Energieminimierungsprinzip auch für die Entstehung der Phoneme ursächlich sein (Lindblom 2000: 308ff.). Die Speicherung von Lautformen ohne die Berücksichtigung von gemeinsam in ihnen enthaltenen Elementen bedeute einen größeren stoffwechselfmäßigen Energieaufwand, als wenn Lautformen in segmentierter Form gespeichert würden (Lindblom 2000: 309ff.). Wegen des Prinzips des minimalen Energieaufwands werde deshalb die Wiederverwendung von Teilen von Lautformen favorisiert (Lindblom 2000: 308ff.), was zur Segmentierung der Lautformen, d.h. ihrer Zergliederung in Phoneme führe (ebd.). Auch hier bleibt wie bei den Aussagen zu den Artikulationen die genaue Bildungsweise der Phoneme infolge ihrer Speicherung auf der angenommenen Grundlage der Stoffwechselenergieminimierung unklar.

Lindblom lehnt zwar (2000: 302) die Annahme von angeborenen Phonemen ab und versucht ihre Entstehung aus allgemeinen Prinzipien der Evolution herzuleiten (ebd.), aber im Grunde werden Mittel und Zweck vertauscht: Ein Pferd läuft vorrangig, weil es durch die Bewegung vorankommt (um ein Ziel zu erreichen), wobei sein Lauf dabei funktional energiesparend ist, damit es sich nicht über die Maßen erschöpft.

Entsprechend ist die Hauptfunktion der Sprache (d.h. ihr Zweck) ihre Verbindung mit dem Denken und der Kommunikation des Menschen, wobei ihre Funktionalität (nur) eine Begleiterscheinung ist. In der belebten Natur lassen sich vielerlei Beispiele für funktionale und besonders energieminimierende Bauweisen von z. B. Vögeln und Fischen finden, die das Ergebnis evolutionärer Entwicklungen sind. Es wäre jedoch irreführend, die Entwicklung dieser Arten nur aus der Perspektive eines funktionalen Aspektes erklären zu wollen. Die von Lindblom (ebd.) angenommenen Funktionen können nicht als Selbstzweck angesehen werden. Zudem läuft auch die Messung der Energieminimierung Gefahr, nach willkürlichen Kriterien zu erfolgen. So nimmt Lindblom (1983) für die bei der Assimilation als eingespart gedachte Energie die Minimierung des zurückgelegten Weges der Artikulatoren an, denn "assimilations were analyzed as reductions of distance in the articulatory space" (Lindblom 1983: 239). Nach Lindblom (1990) wird der Energieaufwand beim Sprechen durch die beiden Kriterien der Deutlichkeit und dem Streben nach

einem möglichst geringen Energieaufwand gesteuert. In der von ihm sog. "Hypospeech" (ebd.) verkörpert sich das zuletzt genannte Prinzip, da es wegen der undeutlichen²¹ Aussprache mit einem minimalen Kraftaufwand gleichgesetzt wird, während die deutliche Aussprache der "Hyperspeech" (ebd.) zu einem größeren Energieaufwand tendiert.

Lindblom (1990) setzt voraus, dass Sprecher ihre Äußerungen entlang eines eindimensionalen Energiekontinuums variieren und dabei nur so deutlich sprechen, wie es das Verständnis ihrer Äußerungen durch den Hörer verlangt (Lindblom 1990: 418). Das Energiekontinuum soll dabei dem bei einer Äußerung durch die Artikulatoren zurückgelegten Weg entsprechen²² (Lindblom 1983: 223; 1990: 428), das zwischen dem Optimum der Energieersparnis in der undeutlichen Sprache ("Hypospeech"(ebd.: 418)) und dem Minimum der Energieersparnis in der überdeutlichen Sprache, der "Hyperspeech" (Lindblom 1990: 418), verläuft. Auch hier gilt die Kritik, dass ein allgemein in der Evolution nachzuweisendes Phänomen, das als Begleiterscheinung von Entwicklungsprozessen auftritt, nicht zum Motor dieser Entwicklung erklärt werden kann. Problematisch ist an dieser Definition der verbrauchten Energie ebenso, dass das individuelle System eines einzelnen Sprechers keine Berücksichtigung findet, sondern allein äußere Merkmale betrachtet werden, die auf Einzeläußerungen des konkreten Sprechakts bezogen sind. Deshalb ist zu erwarten, dass Äußerungen von Sprechern das über solche Definitionen fixierte Energieminimierungsprinzip durchbrechen, weshalb Lindblom (2000: 303) auch an anderer Stelle einräumt, "[...] sociocultural evolution sometimes opposes the effects of the above-mentioned factors [i.e. minimum energy principle]" (Lindblom 2000: 303). Die Untersuchung der Sprache über eine Funktion der Energieminimierung kann also erst nach Bestimmung des Systems als solchem erfolgen, woraufhin eine "systembezogene Funktionalität" entsprechend adäquat festzulegen ist.

²¹ Lindblom (2000) nimmt für eine deutliche Aussprache einen höheren Energieaufwand an als für eine undeutliche Aussprache.

²² Diese Definition der verbrauchten Energie über den von den Artikulatoren zurückgelegten Weg legt Lindbloms (1983, 1990) Auffassung der Beeinflussung von Konsonanten im Kontext von Vokalen und der Vokalreduktion (Lindblom 1990: 428) als einem über eine Verkürzung der Artikulationswege verwirklichten energieminimierenden Verhalten nahe: So wird die Vokal-Konsonant-Koartikulation als eine Bewegung charakterisiert, bei der "a fraction of the maximum possible" (Lindblom 1990: 428) ausgeführt werde (aus der Sicht des Sprechers unter Wahrung des Hörerinteresses der Wahrnehmbarkeit) (ebd.).

5. Die Silbe als eine grundlegende Einheit

In diesem Kapitel wird zunächst die Bedeutung der Silbe in der Sprachverarbeitung vorgestellt und im anschließenden Abschnitt ihre Rolle im Modell des Lexikonzugriffs bei der Wortproduktion von Levelt et al. (1999) diskutiert. Anschließend soll im dritten Abschnitt die Unterschiedlichkeit der Silbenkonstituenten hinsichtlich der für sie angenommenen verschiedenen Funktionen beim Phonemerwerb präsentiert werden. Im vierten und letzten Abschnitt dieses Kapitels werden einige theoretische Definitionsversuche der Silbe besprochen und Vorschläge entwickelt, welche Merkmale eine theoretische Definition der Silbe u.a. beinhalten sollte.

5.1. Die Silbe in der Sprachverarbeitung

Historisch deuten die Silbenschriften, aus denen erst die Alphabetschriften hervorgegangen sind (Ladefoged, Johnson 2011: 243) auf die primäre Bedeutung der Silben gegenüber den einzelnen Phonemen hin, die aber erst ab einer gewissen kritischen Dichte der Silbenschnittmengen abhängig vom Silbensystem wahrnehmbar sind. Auf diesen kritischen Wert deutet auch die Tatsache hin, dass Silbenkerne schon vor dem Lesenlernen in Abfragen der phonematischen Bewusstheit verfügbar sein können, während dies für die Konsonanten erst nach dem Beginn des Lesenlernens zutrifft (Smith & Tager-Flusberg 1980 in Piske 2001: 50). Auch die frühere Verfügbarkeit von Silbenkonstituenten verglichen mit der der Phoneme bei Überprüfungen von Kindern hinsichtlich ihrer phonematischen Bewusstheit verweist auf die primäre Rolle der Silben gegenüber den Phonemen²³. Da bekannt ist, dass das Kind bei seinen sprachlichen Artikulationen auf Produktionen aus dem vorsprachlichen Zeitraum zurückgreift (Piske 2001: 23ff.), ist dieser auch für den Phonemerwerb von Bedeutung. In der vorsprachlichen Phase gehen den durch die Beteiligung von Artikulatoren zusätzlich zur Glottis produzierten Äußerungen, die silbenähnlich sind (Oller, Griebel 2008), auch solche voraus, die nur durch die Stimmritze produziert werden, vokalähnlich sind und auch "quasivowels" (Oller, Griebel 2008: 147) genannt werden (ebd.). Das Kind beginnt mit ihrer Produktion am Anfang seines ersten Lebensmonats (ebd.). Die "quasivowels" können wegen der Ruhelage der übrigen Artikulatoren auch als reine Stimmgebungsereignisse angesehen werden (ebd.).

²³ Siehe dazu auch Abschnitt 4.2.

Während des zweiten oder dritten Lebensmonats beginnt das Kind neben der Variation der "quasivowels" (ebb.: 147f.) und der Veränderung der Tonhöhe mittels des Kehlkopfs (ebd.: 148) mit der Produktion silbenähnlicher Äußerungen, auch "gooing" (ebd.: 147) genannt, die durch die Bewegung des Zungenrückens und dessen Kontaktbildung mit verschiedenen Mundraumstellen hervorgerufen werden (ebd.: 147). Weil die Bewegung des Dorsums und die mit ihm verbundenen Kontaktstellen im Mundraum nicht vorhergesagt werden können, wird das "gooing" (ebd.: 147) auch als unentwickelte Silbenbildung angesehen (ebd.: 148). Diese "unentwickelten" (ebd.: 147) silbenähnlichen Äußerungen werden als Weiterentwicklung der zeitlich früheren Quasivokale betrachtet (ebd.: 148). Nach einem Zwischenzeitraum, für den die Bildung von Kategorien für Äußerungen des Kindes angenommen wird (ebd.: 148ff.), folgt auf die zuvor beschriebene Produktion der silbenähnlichen, aber sehr veränderlichen "gooing"-Äußerungen (ebd.: 147), die Äußerung von sog. "canonical syllables" (ebd.: 150ff.) zwischen dem fünften und zehnten Lebensmonat (ebd.: 150). Diese können als weitergehende, komplexere Entwicklungsstufe der vorhergehenden Äußerungsarten angesehen werden, indem sie eine koordinierte Stimmgebung, eine Artikulation oberhalb der Glottis und zudem die Kategorienbildung von Äußerungen beinhalten (ebd.: 151). Sie haben also u.a. die folgenden Eigenschaften: Sie wirken auf Grund des Rückgriffs auf die gebildeten Kategorien wie systematisch erzeugt und weisen einen vokalähnlichen und mindestens einen seitlich beigefügten konsonantenähnlichen Laut auf (ebd.: 150). Wegen ihrer systematischen CV(C)-ähnlichen Gestalt gleichen sie echten Silben einer Sprache (ebd.: 150) und werden deshalb auch als "well-formed" (ebd.) oder kanonisch bezeichnet (ebd.). Neben der zuvor beschriebenen Produktion von silbenähnlichen stimmlichen Äußerungen zeigen sehr junge Kinder auch perzeptive Sensitivität gegenüber Silben, während diese für einzelne Phoneme nicht gegeben zu sein scheint: Neugeborene, die höchstens 2 Monate alt sind, konnten in einem Habituerungs-Saugratten-Experiment von Bertoncini und Mehler (1981) Unterschiede, die durch die Vertauschung der beiden äußersten Konsonanten einer Phonemefolge entstanden, erkennen, wenn es sich bei derselben Folge um CVC-artige handelt (ebd.). Dies war ebenfalls für eine Minderheit der Untersuchungsgruppe für VCCCV-Sequenzen der Fall, aber nicht für die Gruppe, der eine CCC-Folge ganz ohne Vokal präsentiert wurde. Daraus schließen Bertoncini und Mehler (1981: 259), "that the syllable is the natural speech processing unit, while the phoneme is a classificatory unit available to speakers who have mastered the alphabetic system of reading, and writing."(ebd.).

Zusätzlich zu der eben vorgestellten produktiven und perzeptiven Verfügbarkeit der Silbe für sehr kleine Kinder, scheint sie im Zusammenhang mit der Sprachverarbeitung fortgeschrittener, erwachsener Sprachteilnehmer auch hinsichtlich des Abrufs von Wörtern aus dem Lexikon gespeichert zu werden, wobei jedoch Uneinigkeit besteht, auf welcher Stufe der Sprachverarbeitung die Silbe anzusiedeln ist, was näher im folgenden Abschnitt im Zusammenhang mit dem Sprachverarbeitungsmodell von Levelt et al. (1999) thematisiert wird: Die Schlussfolgerung aber, dass die Silbe hinsichtlich der mit der Produktion von Sprache verbundenen Abläufe gespeichert werden muss, legt der sog. "Silbenhäufigkeitseffekt" (Schiller 2008: 159ff.) nahe, bei dem Sprachteilnehmer ein anderes Sprachverhalten zeigen bezogen auf häufige im Vergleich zu dem Verhalten für seltene Silben (Schiller 2008: 159ff.): So präsentierten Cholin et al. (2006: 214ff.) holländischen erwachsenen Sprechern zwei Gruppen von CVC-Silben, von denen die eine nur häufige und die andere nur seltene CVC-Silben enthielt (ebd.). Diese CVC-Silben entsprachen keinen Wörtern des Holländischen und die Sprecher mussten in einer Lernphase die einzelnen CVC-Silben mit einem von zwei Punkten auf einem Bildschirm assoziieren, um sie in der Testphase bei Anzeige des entsprechenden Punktes zu nennen (ebd.). Dabei zeigte sich, dass die häufigen CVC-Silben signifikant schneller genannt wurden als die seltenen CVC-Silben (ebd.: 218f.). Weil angenommen wird, dass nur gespeicherte mentale Einheiten einen Häufigkeitseffekt zeigen können, ziehen Cholin et al. (2006: 219) daraus den Schluss, dass Silben in phonetischer Form im Gedächtnis fixiert sein müssen und sprechen deshalb auch von einem sog. mentalen Silbenspeicher (Cholin et al. 2006: 219).

5.2. Levelt et al.s (1999) Modell des Lexikonzugriffs während der Sprachproduktion

Das Modell des Abrufs von Lexemen aus dem mentalen Lexikon von Levelt et al. (1999) soll hier vorgestellt werden, weil es die Silbe als eine zentrale Einheit ansieht, jedoch in einem umgekehrten Verhältnis zu den Phonemen, wie es in dieser Arbeit vorgeschlagen wird, das die Silben als grundlegend ansieht und die Phoneme als sich aus dem Silbensystem ergebende, sekundäre Einheiten. In seiner Konzeption erinnert das Modell von Levelt et al. (1999) durch die hierarchisch tief angeordnete Rolle, die es der Silbe als phonetische Realisierung der auf höherer Planungsebene festgesetzten Phonemfolgen zuweist, an die phonetischen Implementierungsregeln der generativen Phonologie (Chomsky, Halle 1968: 44). Diese übersetzen die invarianten, angeborenen Phoneme der "underlying form" (ebd.) lediglich in phonetische Oberflächenformen und führen so aber zu einer "discrepancy between underlying forms and their phonetic realizations" (ebd.).

Levelt et al. (1999) nehmen ein hierarchisches Netzwerkmodell mit fünf Ebenen an, bei dem auf der obersten Ebene die Aktivierung eines Konzepts wie beispielsweise das für das englische Verb *escort* zur Auswahl eines sog. Lemmas des Lexikons wie hier *<escort>* führt (Levelt et al. 1999: 3). Dieses Lemma wird auf der nächsten und dritten Ebene durch entsprechende Morpheme wie beispielsweise das Suffix *<ing>* für die Verlaufsform oder andere für den syntaktischen Zusammenhang erforderliche Morpheme modifiziert (Levelt et al. 1999: 4ff.). Erst auf der vierten Ebene werden für das Lemma Phoneme abgerufen, die gleichzeitig verfügbar sind und hinsichtlich ihrer Reihenfolge indiziert sind wie hier die Folge */ε₁,s₂,k₃,ɔ₄,r₅,t₆/* (ebd.). Außerdem wird auf dieser Ebene neben weiteren Verrechnungsschritten eine sog. metrische Schablone abgerufen, die für das Lemma *escort* u.a. die Anzahl seiner Silben wie hier zwei in Form von Silbenplatzhaltern angibt (ebd.). Es findet aber noch keine Zuordnung der Phonemfolge zu den Silbenplatzhaltern statt. Anders ausgedrückt, wird keine Speicherung der Lautformen der Lexeme über Silben angenommen. Dies begründen sie (Levelt et al. 1999: 5) damit, dass ein Lemma wie *escort* die Silbifizierung wie *e-scort* bei Suffixierung mit *ing* nicht beibehält, sondern zu *e-scor-ting* verändert (ebd.) und dass die Silbifizierung sich über Wortgrenzen im Satz wie in *"He'll escort us"* (hier: *e-scor-tus*) erstrecken kann (ebd.). Dieses Resilbifizierungsphänomen erfordere deshalb eine Speicherung der Lexeme als Phonemfolgen (Levelt et al. 1999: 5). Erst bei der Verbindung der Lexeme im Satzzusammenhang würden die zu ihnen gehörigen Phonemfolgen auf die Platzhalter der metrischen Schablonen nach allgemeinen Silbifizierungsprinzipien wie beispielsweise der Onset-Maximierung verteilt (ebd.). So ergeben sich mit Phonemen gefüllte Platzhalter, die auch "phonologische Silben" genannt werden (ebd.). Diese werden auf der fünften Ebene des Modells in sog. phonetische Silben übersetzt, die in einem Silbenspeicher aufbewahrt werden und schließlich die Artikulation bzw. Produktion einer konkreten Äußerungen ermöglichen sollen (ebd.). Levelt et al. (1999: 1f.) nehmen zwar für ihr Modell eine Verträglichkeit mit dem Phonemerwerb an, aber es ist die Frage, wie es während des Phonemerwerbs zu den Phonemen kommt, die nur durch die distinktiven Merkmale spezifiziert sein sollen (Levelt et al. 1999: 21). Eine Speicherung von Lexemen in Form abstrakter invarianter Phonemfolgen, wie sie von Levelt et al. (1999) vorgeschlagen wird, steht dem in dieser Arbeit vorgeschlagenen Ansatz entgegen, weil die Silben als relevant für die Entwicklung und Dynamik des Phonemsystems angesehen werden. Denn durch die Bildung neuer Silben in einem sich über Abstände selbst steuernden Prozess, der einen optimalen Ausgleich zwischen der Ökonomie der Bildung und der Unter-

scheidbarkeit von Silben erreicht, entstehen erst die Phoneme. In dem Ansatz dieser Arbeit sind die Silben essentiell für das Lexikon, weil über sie eine Unterscheidung der bereits gespeicherten Lautformen erfolgt. Im Ansatz von Levelt et al. (1999) hingegen sind sie nur das Ergebnis eines Übersetzungsvorgangs der invarianten Phonemfolgen in an sich ebenfalls invariante Gestenpartituren, die den kontinuierlichen Sprachfluss erzeugen sollen. Bei ihrem Argument gegen eine Speicherung von Lexemen im Lexikon in Form konkreter Silben beachten Levelt et al. (1999: 5) zudem nicht, dass das Kind zunächst nur einzelne Wörter äußert und erst in der Folge mehrere Wörter zu Sätzen kombiniert (Bloom 1995: 7). Für die Einzelwörter kann das Resilbifizierungsargument gegen die Speicherung von Lexemen über Silben aber nicht gelten. Von diesem Punkt ausgehend müsste untersucht werden, auf welche Weise eine Resilbifizierung im Satzzusammenhang erfolgen könnte. Dementsprechend könnte es sein, dass das Kind für seine Lexeme aus der Ein-Wort-Phase zunächst ihre Silbifizierung im Satzzusammenhang beibehält. Die Äußerung einer anderen Silbenfolge für eine Kombination von Lexemen im Satzzusammenhang als diejenige, die für die isoliert geäußerten Lexeme genannt werden, ist sicher ein Phänomen²⁴, das erst in einem größeren Silbensystem entsteht und auf syntaktische Einflüsse zurückzuführen ist, die in diesem Zusammenhang nicht berücksichtigt werden. Ähnliche Überlegungen könnten auch die Veränderung der Silbeneinteilung von Lexemen bei ihrer Teilnahme an Wortbildungsprozessen wie der Derivation von *un-der-stand* zu *un-der-stan-der* (Levelt et al. 1999: 20) erklären. Als zweites Argument für die Speicherung von Lexemen in Form von Folgen invarianter Phoneme geben Levelt et al. (1999: 20) die häufige Vertauschung von Einzelphonemen bei Versprechern durch erwachsene Sprachteilnehmer an (ebd.). Denkbar wäre aber auch hier eine Vertauschung von Silben, deren Transkriptionen sich lediglich in einem Phonem unterscheiden. Dies scheint durch den Hinweis von Levelt et al. (1999: 20) unterstützt zu werden, dass nicht nur Einzellaute ersetzt werden, sondern auch ganze Konsonantenkomplexe vom Wortanfang, die einem Silbenkopf entsprechen (ebd.).

5.3. Funktionen der Silbenkonstituenten während des Silbenwachstums

Jede Silbe bildet für sich eine untrennbare Einheit²⁵, die wiederum mit allen anderen Silben in einem Gesamtsystemzusammenhang steht, aus dem sich die Phoneme als illu-

²⁴ hier als "Resilbifizierung" bezeichnet

²⁵ In Abschnitt 5.1. wurde bereits auf die Bedeutung und intuitive Verfügbarkeit der Silbe im Sprachverhalten hingewiesen.

sionäre Einheiten herausbilden. Man könnte die Silbe mit der für sie gehörten Phonemfolge gleichsetzen, doch die Gruppierung mehrerer benachbarter Phoneme, auch als "silbische Konstituenten" (Davis 2006: 326) bezeichnet, können in einer korrelativen Beziehung zueinander stehen. Deshalb sagen Pike und Pike (1947) auch, dass "[t]he structure of these syllables does not consist of a series of sounds equally related, like beads on a string, but is rather like an overlapping series of layers of bricks." (Pike, Pike 1947: 78). Jede Silbe hat einen Kernbereich, der auch als "(Silben-)Nukleus" oder "(Silben-)Kern" bezeichnet wird und möglicherweise zwei Randbereiche²⁶, wobei der links vom Nukleus befindliche als "(Silben-)Kopf" oder "(Silben-)Anfangsrand" und der rechts vom Nukleus befindliche als "(Silben-)Koda" oder "(Silben-)Endrand" bezeichnet wird (Davis 2006: 326; Bosch 2010: 782ff., Blevins 1995: 212ff.). Durch die Untergliederung der Silbe in die Konstituenten des Kopfes, Kernes und der Koda werden einzelne Phoneme, die gemeinsam eine dieser Konstituenten bilden, zusammengruppiert. So bilden die zwei Phoneme /t/ und /r/ in der Silbe /traust/ der flektierten Verbform *trauen* die Konstituente des Kopfes, die beiden Vokale /a/ und /u/ den Silbenkern (bzw. Nukleus) und die beiden Konsonanten /s/ und /t/ die Silbenkoda. Die terminologische Untergliederung der Silbe in die Silbenkonstituenten ist durch das unterschiedliche Verhalten der Silbenkonstituenten gerechtfertigt, das sich u.a. in Beobachtungen von phonembezogenen Kombinationsbeschränkungen (sog. *phonotaktischen Restriktionen*) zeigt, die für alle Phoneme einer Konstituente, aber nicht für Phonempaare unterschiedlicher Konstituenten einer Silbe gelten müssen: Eine Folge aus zwei Plosiven wie in den Wörtern *Abt* (Vater 1992: 104), *nackt* (ebd.), *Haupt* (ebd.), *Akt* (ebd.), *Smaragd* (ebd.) kommt nur in der Koda vor (immer mit dem Konsonanten /t/ als zweitem Phonem (Vater 1992: 104f.)), während als Silbenkopf der Konsonantenkomplex /pt/ lediglich in aus Fremdsprachen übernommenen Namen wie *Ptolemäus* (Vater 1992: 104) vertreten ist (ebd.). Im Deutschen können im Onset wortinitialer Silben keine homorganen Obstruent + Sonorant-Folgen wie "/pm/, /bm/, /fm/, /vm/, /tn/, /dn/, /tl/, /kŋ/, /gŋ/, /xŋ/" (ebd.) vorkommen (Vater (1992: 105f.)), während Folgen aus Obstruent + Sonorant wie /mp/ wie in *Lump* (ebd.: 106) und /nf/ wie in *fünf* (ebd.), /nd/ wie in *Hand* (ebd.) oder /ŋk/ wie in

²⁶ Das Problem, die Konstituenten innerhalb der Silbe terminologisch fundiert festzulegen, ist direkt mit der Definition der Silbe verbunden, die die Voraussetzung bildet für eine Festlegung der Konstituentengrenzen. Da eine Widerspruchsfreiheit in verschiedenen Silbentheorien nicht erreicht wird, soll hier die intuitive Verfügbarkeit des Silbenbegriffs zur vorläufigen Grundlage genommen werden, die aber nicht willkürlich ist, sondern ihre Begründung im perceptiven sowie im produktiven Zusammenhang erfährt.

Bank (ebd.) nur für die Koda der deutschen Silbe möglich sind (ebd.: 106). Vater (1992) bewertet "[a]llein diese Restriktion [als Argument] für die Ansetzung einer O-Konstituente" (Vater 1992: 106). Die Untersuchungen von Konstituentenverteilungen in Korpora in Kapitel 7 deuten darauf hin, dass die Konstituenten während des Wachstums der Silben und damit der Ausbildung des Silbensystems unterschiedliche Funktionen einnehmen. Deshalb werden in dem Silbenmodell dieser Arbeit verschiedene Funktionen für die Silbenkonstituenten angenommen: Der Silbenkern gilt als Signalträger, wogegen die Silbenränder Modifikationen bzw. Eigenschaften des Silbenkerns darstellen, die sich in ihrer Funktion unterscheiden. Die Rollenverteilung der Silbenkonstituenten wird in phonetischen Beobachtungen deutlich: Die signaltragende Funktion des Silbenkerns spiegelt sich in der Hörbarkeit der Phoneme, die den Silbenkern bevorzugt bilden: So hat Wolf (1871 in Lindner 2009: 118f.) die "Tonstärke" (ebd.) der Phoneme gemessen: Dabei äußerte eine Person ein bestimmtes Phonem, während sich die andere Person so weit entfernte, bis sie das Phonem gerade noch hören konnte. Die im Versuch für jedes Phonem festgestellte Entfernung wurde als die "Tonstärke" des Phonems definiert (ebd.), die für die Phoneme unterschiedlich ist (ebd.), wobei alle Vokale eine größere Tonstärke als die Konsonanten haben. Da Silbenkerne in den Sprachen bevorzugt²⁷ von Vokalen gebildet werden, scheint die Hörbarkeit der die Silbenkerne u.a. bildenden Vokale eine Funktion zu sein, die ihre Konzeption als Signalträger unterstützt²⁸ und den Kern von den Rändern unterscheidet, die durch Konsonanten gebildet werden. Pouplier et al. (2011: 244) sehen diese Vokal-Konsonant-Trennung der Phonemmenge einer Sprache, die die Unterteilung einer Silbe mit vokalischem Kern in Nukleus und Ränder ausmacht, auch durch die Art der artikulatorischen Koordination bedingt: So verweisen sie auf Studien (u.a. Sproat and Fujimura 1993; Krakow 1999 beide in Pouplier et al. 2011: 244), die gezeigt hätten, dass die Artikulation der Konsonanten in Bezug auf den vokalischen Nukleus erfolge und dass der vokalische Nukleus die Organisationsgrundlage für die Artikulation der Konsonanten darstelle (ebd.) und ziehen aus diesen Studien das Fazit, dass "vowels seem to serve a very specific function within the syllable and the general confound between a vowel and the nucleus position may be no accident." (Pouplier et al. 2011: 244). Die unterschied-

²⁷ neben silbischen Konsonanten

²⁸ Im Abschnitt 5.4.1. wird die Sonoritätshierarchie besprochen, die über phonologische Beobachtungen einer Sprache hergeleitet wird. Sie hat große Übereinstimmungen mit der von Wolf (1871 in Lindner 2009: 118f.) festgestellten Hörbarkeitshierarchie (Lindner 2009: 118f.) und zudem ebenfalls mit der von Ladefoged und Johnson (2011: 245f.) angegebenen Messung der akustischen Energie der einzelnen Phoneme. Auch in diesen nur akustischen und nicht auditiven Messungen wird die Zweiteilung von Konsonanten- und Vokalphonemen deutlich.

liche Funktion der Köpfe als Haupteigenschaft der Silbenkerne und der Kodas als Nebeneigenschaften scheint sich wie zuvor bei der Gegenüberstellung des Nukleus (in Form der Vokale) und der Ränder auch in ihrer Hörbarkeit auszudrücken: So untersuchte Benkí (2003) die Perzeption von C_1VC_2 -Silben durch junge Erwachsene (ebd.: 130), denen jeweils akustisch mit gleichzeitigem Rauschen einer von vier Intensitäten (-5 dB, -8 dB, -11 dB, -14 dB) C_1VC_2 Silben präsentiert wurden (ebd.: 130ff.). Die Teilnehmer mussten dann die wahrgenommenen C_1VC_2 -Silben notieren (ebd.: 130ff.). In der anschließenden Analyse wurde ein Vergleich der gehörten und notierten Silben mit den präsentierten CVC-Silben vorgenommen (ebd.: 132), wobei man zum Ergebnis kam, dass im Schnitt (für alle Teilnehmer und alle Rauschstufen) die Köpfe C_1 seltener mit anderen Phonemen vertauscht wurden als dies für die Kodas C_2 der Fall war (ebd.: 133, 146). Insbesondere²⁹ wurden Silbenkodas häufiger weggelassen als Silbenköpfe (ebd.). Daraus zieht Benkí (2003: 150) den Schluss, dass dieses Ergebnis der Untersuchung "a perceptual advantage for onsets over codas" (Benkí 2003: 150) zeige. Diese bessere Hörbarkeit von Silbenköpfen gegenüber -kodas, die sich in der geringeren Verwechslung von Köpfen gegenüber den Kodas (für die verschiedenen Rausch-Signalverhältnisse) zeigt, kann als Ausdruck einer primären Funktion der Köpfe zur eindeutigen Perzeption der Silben verstanden werden. Plötzliche Zunahmen in der Signalenergie können zu einer verstärkten Aktivität in den Nervenfasern des Hörsystems führen, die zu einer Verstärkung des entsprechenden Signalteils führen (Kiang et al., 1965; Smith, 1979 beide in Benkí 2003: 150). Diese verstärkte Aktivitätsrate der auditiven Nervenfasern sei bei der Präsentation von CV-Silben, aber nicht von VC-Silben beobachtet worden (Sinex and Geisler, 1983; Delgutte and Kiang, 1984a, b; Sinex, 1995 alle in Benkí 2003: 150). Content et al. (2001) gestehen den Silbenköpfen verglichen mit den Silbenkodas auch eine besondere Rolle bei der Sprachverarbeitung zu: Über sie erfolge die Segmentierung des gehörten Sprachflusses (Content et. al. 195) und für die Worterkennung seien sie "privileged alignment points in the signal" (ebd.), weil sie potentiell den Beginn neuer Wörter signalisierten (ebd.). Es kann aufbauend auf den Konstituentenbegriffen des Kopfes, Kerns und der Koda eine weitere Gruppierung von je zwei Konstituenten zu einem Paar vorgenommen werden. So erhält man statt der Auffassung der Silbe als einer linearen Folge der Konstituenten Kopf, Kern und Koda, die auch als *flache* Silbenstruktur bezeichnet wird (Bosch 2010: 785), eine weitere hierarchische Ebene in dem zugehörigen Baummodell der Silbe, weshalb die

²⁹ wenn man die Weglassung eines Phonems als die Vertauschung mit dem "Nichts" bzw. einem "leeren Phonem" ansieht

sich ergebende Silbenstruktur auch mit *hierarchisch* benannt wird (Bosch 2010: 787ff.). So wird das Paar aus Kopf und Kern als *Körper*, das aus Kern und Koda als *Reim* und das aus Kopf und Koda als *Schale* bezeichnet (Restle, Vennemann 2001: 1326). Entsprechend werden die sich zu diesen Gruppierungen ergebenden Baummodelle der Silbe auch als das *Körper-*, *Reim-* und *Schalenmodell* (Restle, Vennemann 2001: 1326) bezeichnet³⁰. Diese Gruppierungen werden wie die zuvor dargestellte Konstituentenbildung dadurch gerechtfertigt, dass sie die besonders passende Beschreibung von Phänomenen ermöglichen (Restle, Vennemann 2001: 1327f.). Für jedes dieser drei Modelle lässt sich, wenn auch für unterschiedliche Zwecke, ausreichend Evidenz anführen (Vennemann, Restle 2001: 1327), so dass jede der Gruppierungen, d.h. Körper, Reim und Schale, als Merkmal bei den Untersuchungen von Konstituentenverteilungen in Korpora im Kapitel 7 beachtet werden sollen, die hier auch zusammengefasst als Menge der *Binärrelationen* benannt werden.

5.4. Vorstellung und Merkmale verschiedener Silben(erwerbs)theorien

Das Phänomen der Silbe zeigt sich auf vielerlei Weise: So können schon kleine Kinder mit dem, was als Silbe bezeichnet wird, in Sprachspielen umgehen. Ältere Sprachbenutzer können die Anzahl der Silben von Wörtern bestimmen und auf gleiche Weise wie Kinder in Sprachspielen Operationen wie Vertauschungen, Ersetzungen etc. vornehmen. Dass sie auch eine physiologische Bedeutung hat, wird aus entsprechendem Sprachverhalten abgeleitet wie z.B. dem Silbenhäufigkeitseffekt, bei dem häufige Silben ein anderes Sprachverhalten zeigen als seltenere. Da ein Sprachverhalten, wie es im Silbenhäufigkeitseffekt auftritt, dem einzelnen Sprachbenutzer nicht bewusst ist, prägen solche bewussten Operationen mit Silben (z.B. die Nennung der Silbenanzahl) den für jeden (bis auf ganz kleine Kinder) verfügbaren intuitiven Begriff der Silbe.

Mit diesem intuitiven Silbenbegriff arbeiten auch die verschiedenen Ansätze zur Silbentheorie, indem ihre Beschreibungen der Silbe auf der Silbenbestimmung des Beschreibenden bzw. Forschenden oder einem Korpus mit bereits durch einen anderen Sprachbenutzer in Silben eingeteilten Wörtern basieren oder indem sie ihre Resultate aus entwickelten Silbenbestimmungsverfahren mit diesen verfügbaren Silben vergleichen. Die verschiedenen theoretischen Silbenbegriffe erweisen sich aber bisher als nicht geeignet,

³⁰ Ein weiteres Silbenmodell ist die Baumstruktur, die den Kern sowohl zu dem Kopf als auch der Koda gruppiert und damit sowohl eine Körper als auch eine Reimkomponente enthält und als "Überlappungsstruktur" (Vennemann, Restle 2001: 1326) der Silbe bezeichnet wird (ebd.).

widerspruchsfreie Einteilungen vorzunehmen, weil Silben oder Teile derselben (wie z.B. die extrasilbischen Segmente bei Wiese (1988)) als Ausnahmen bezeichnet, die Silbenbestimmungsprozedur eine Restmenge der intuitiv bestimmten Silbenmenge nicht erfasst, oder aber nicht alle intuitiv bestimmten Silben der entsprechenden Sprache erklärt werden können. Um die Problematik der theoretischen Behandlung der Silbe besser verstehen zu können, sollen einige Silbenbegriffe vorgestellt und voneinander abgegrenzt werden. Aus dieser Kritik sowie der von Theorien, die die Silbe explizit ablehnen, werden Merkmale entwickelt, die bei einer Definition der Silbe berücksichtigt werden sollten.

5.4.1. Silben(erwerbs)theorien

5.4.1.1. Das Koordinationsmodell (im Rahmen der Artikulatorischen Phonologie)

Das Koordinationsmodell (u.a. Nam, Goldstein, Saltzman 2009; Browman, Goldstein 2000, Pouplier 2011) der Silbe wird auf der Grundlage der Artikulatorischen Phonologie konstruiert, in welcher die Phoneme aus sog. Gesten zusammengesetzt sind (ebd.). Dies sind (Idealisierungen von) Verengungen im Mundraum wie z.B. dem Velum- oder dem Lippenverschluss für das Phonem /m/ (ebd.). Die Artikulatorische Phonologie fasst die Gesten als die Atome der Sprachverarbeitung (sowohl der Perzeption, Produktion als auch Planung) auf, die verbunden bzw. koordiniert werden müssen, um einen sprachlichen Ausdruck zu produzieren (ebd.). Dabei wird die zeitliche Abfolge des Auftretens einer Geste als relevant angenommen, weil die Verschiebung einer Geste zu einem völlig anderen sprachlichen Ausdruck führen kann, z.B. die der Velumgeste (d.h. Zeitpunkt der Velumöffnung) in den englischen Wörtern *ban* gegenüber *man* (ebd.). An beiden Wörtern sind identische Gesten beteiligt, nur das Velum öffnet sich zu verschiedenen Zeitpunkten: In *man* zu Beginn, in *ban* erst gegen Ende. Würde sich die Velumöffnung hingegen zeitlich nach hinten verschieben, so würde *ban* zu *man* (mit evtl. nasaliertem Vokal) (ebd.). Deshalb wird die zeitliche Anordnung der Gesten als wichtig betrachtet und ihre zeitlichen Abstände zueinander gemessen (ebd.): So u.a. die der Gesten des Kopfes einer Silbe zum Vokal, was (für einige Sprachen) zeigte, dass der Abstand (der Gesten) eines Konsonanten zum Vokal dem Mittelwert der Abstände der (Gesten des) Konsonanten eines komplexen Kopfes zum Vokal entspricht, wobei dieser Mittelwert oder der Abstand eines einzelnen Konsonanten vom Vokal als Konsonantenzentrum (*c-centrum*) bezeichnet wird und dieser zeitliche Effekt als Konsonantenzentrumseffekt (*c-centrums-effect*) (ebd.).

Er kann als zeitlich-artikulatorische Entsprechung des linguistischen Konstrukts des Silbenkopfes angesehen werden (ebd.). Für die (Gesten der) Konsonanten des Reims einer Silbe konnte ein solcher Mittelwert nicht festgestellt werden, sondern nur eine Abfolge der Phoneme mit gleichen zeitlichen Abständen zueinander (ebd.). Für die Erklärung dieser zeitlichen Verhältnisse wird eine paarweise Koordination der Konsonanten jeweils mit dem Vokal angenommen, von einer bestimmten Art für den Silbenkopf und einer davon verschiedenen für die aufeinanderfolgenden Paare des Reims, die mit der der Konsonanten des Silbenkopfes in Konkurrenz steht (ebd.).

Neben diesen Koordinationsmodi oder -arten wird für jede Geste ein Oszillator angenommen, eine Uhr, die schwingt und zusammen mit den Koordinations- oder Kopplungsmodi für das zeitliche Auftreten der Gesten sorgt (ebd.). Dieses Koordinationsmodell der Silbe wird durch mathematische Gleichungen beschrieben (ebd.). Die lediglich paarweise Kopplung der Gesten und die Annahme einer Uhr für jede einzelne Geste, gegenüber einer Uhr für alle Gesten bzw. einer zentralen Steuerung der Gesten, sorgt dafür, dass das globale Verhalten der Gesten aus diesem lokalen Verhalten entsteht, d.h. emergent ist (ebd.). Auf diese Weise konnte der beobachtete Konsonanzentrumseffekt modelliert werden (ebd.). Weil der Koordinationsmodus, der für die paarweise Koordination der einzelnen Konsonanten des Silbenkopfes jeweils zum Silbenkern angenommen wird, als besser verfügbar gilt als der für die Konsonanten der Silbenkoda, sehen die Anhänger dieses Modells Fakten als erklärt an, die den Silbenkopf in den meisten Sprachen als viel verbreiteter auszeichnen als die Silbenkoda (ebd.). Weil die zeitliche Anordnung von Gesten für das Hervorbringen von sprachlichen Signalen wichtig ist, wie im obigen *ban-man*-Beispiel demonstriert wurde, sehen die Vertreter der Artikulatorischen Phonetik das Erlernen der richtigen Anwendung der verfügbaren bzw. angeborenen zeitlichen Koordinationsmodi (*in-phase*, *anti-phase*) von Gesten entsprechend dem Zielsprachwort als eine (oder sogar die) wichtige Aufgabe im Spracherwerb an. Abweichungen in den Äußerungen der lautlichen Formen eines Zielsprachwortes durch das Kind führen sie auf die unzulängliche Trennung der Gesten in einem Wort mit den beiden verfügbaren Koordinationsmodi zurück und vermuten, man könne bei genauen zeitlichen artikulatorischen Messungen die für das zielsprachliche Wort notwendigen Gesten feststellen, die aber (nahezu) simultan geäußert würden. Spracherwerb bestünde somit im Trennen der Gesten eines zielsprachlichen Wortes durch die korrekte Anwendung der beiden Koordinationsmodi. Diese Auffassung wird auch in einer Lernsimulation bei Nam, Goldstein, Saltzman (2009: 310) deutlich:

The adult representation includes modes corresponding to CV (in-phase) and VC (anti-phase), where the relative strength of these modes can differ from language to language, corresponding to the relative frequency of CV and VC structures in that language. At the onset of learning, the child's representation does not include any CV or VC modes, so the child displays no preference for producing any relative phases over others; the relative phases produced are randomly distributed. As the result of the learning process, modes develop that correspond to the modes found in the adult speakers' representations and to their relative strength. (ebd.)

Danach besteht das Lernen im Koordinationsmodimodell der Artikulatorischen Phonologie aus der Anwendung des richtigen eines der beiden gegebenen Gestenkoordinationsmodi (*in-phase*, *anti-phase*). Der Gestenansatz wird auch für die historische Veränderung von Lautformen und die verschiedenen Koordinationsmodi herangezogen: Den historischen Übergang der Lautfolge /nd/ zu /d/ im Brasilianischen Portugiesisch erklärt Goldstein (2012) mit dem Verlust der zeitlichen Trennung der Zungenspitzenegeste, die in den Phonemen /n/ und /d/ gleich ist, von der Velumegeste, die /n/ und /d/ unterscheidet³¹ und deren Ersetzung durch eine synchrone Ausführung dieser beiden Gesten. Da das Koordinationsmodimodell sich lediglich auf Silbenpositionen unabhängig vom segmentalen Inhalt bezieht, ist es als Modell für den Spracherwerb problematisch, sofern bei diesem nicht nur segmentunabhängige CV-Schablonen (CV, VC, CCV, VCC, etc.) in ihrem zeitlichen Auftreten erklärt werden sollen, sondern die verschiedenen Silben mit verschiedenen Phonemen an verschiedenen Positionen³². Zudem werden die kindlichen Silbenäußerungen als Abweichung von dem für die Zielsprache als normal angesehenen zeitlichen Muster der *in-phase*, *anti-phase* betrachtet, was einer Fehlerklassifikation gleichkommt.

5.4.1.2. Stetsons (1951) Silbendefinition

Das Wesentliche an der artikulierten Sprache sieht Stetson (1951: 4ff.) in der Bewegung, die sie im Brustkorb (Rippenmuskeln) erzeugt. Über diese Bewegungen wird die gegenüber den Phonemen als grundlegend angenommene Silbe (Stetson 1951: 91) definiert. So sagt er (Stetson 1951: 91) hinsichtlich der Silbe, dass ihre "[...] delimitation is not due to a 'point of minimum sonority' but to the conditions which define a movement as one movement." (Stetson 1951: 91). Stetson (1951) geht davon aus, dass jeder Silbe während

³¹ zeitsequentielle Planung dieser beiden Gesten bzw. *in-phase locking* bzw. *in-phase*-Koordinationsmodus

³² Siehe auch die in Abschnitt 6.2.7. geäußerte Kritik, die über die hier genannte hinausgeht. Denn es ist überhaupt erst möglich, den Silbenerwerb als den Erwerb von Koordinationsmodi anzunehmen, weil die Gesten und die ihnen zugeordneten invarianten Mundraumorte, die "targets", angeboren sein sollen.

der Ausatmung³³ ein Luftimpuls entspricht, der durch die Aktivität einer Rippenmuskulatur ("internal intercostals") hervorgerufen wird (Stetson 1951: 30). Durch diese Rippenmuskulaturimpulse ist die Einheit der Silbe für Stetson (1951) gewährleistet und ihre primäre Rolle gegenüber den Phonemen ergibt sich dadurch, dass diese erst im Silbenkontext bedeutungsvoll seien, weshalb Stetson (1951: 94) eine Konzeption der Sprache als lineare Folgen von Phonemen ablehnt (Stetson 1951: 4):

It is impossible to teach the "separate sounds" of a language (the sounds cannot be separated), and then assembled into syllables (syllables are not assemblies), and the proper stresses, pauses, and "intonation" added (they are not additions but basic overall traits) (Stetson 1951: 4).

Bedeutung erlangen die Phoneme im Silbenkontext durch die grundlegende Konsonant-Vokal-Unterscheidung: Stetson (1951: 91) ordnet den Vokalen eine den Luftimpuls durch die unterschiedliche Gestalt der supraglottalen Räume (v.a. des Mund- aber auch des Rachenraumes) formende Funktion zu (Stetson 1951: 91) und den Konsonanten eine Rolle als Hilfs- oder Zusatzbewegungen (Stetson 1951: 50f.). Diese Hilfsfunktion der Konsonanten besteht dabei im Arretieren oder Freigeben (ebd.) des durch den Rippenmuskulaturimpuls freigesetzten Luftimpulses (Stetson 1951: 50f.) durch die Artikulatoren der Glottis, der Zunge, des Unterkiefers und der Lippen (Stetson 1951: 29, 44). Durch das konsonantische Arretieren werden mit Konsonanten besetzte Silbenkodes (d.h. -VC bzw. nicht-offen) erzeugt (Stetson 1951: 2, 50) und durch das Freigeben mit Konsonanten besetzte Silbenköpfe (d.h. CV- bzw. nicht-nackt) (Stetson 1951: 2, 50). Erfolgt das Arretieren und Freigeben hingegen nicht durch die glottalen und supraglottalen Artikulatoren, sondern durch die Rippenmuskeln selbst, so entstünden Silben alleine aus Vokalen (Stetson 1951: 50). Stetsons (1951) Beobachtungen der durch Rippenmuskeln initiierten Bewegungen in Verbindung mit Lautäußerungen können als Indizien für die Einheit der Silbe gesehen werden, obwohl die Annahme einer eindeutigen Entsprechung zwischen Silben und den von den Rippenmuskeln während der Ausatmung hervorgerufenen Bewegungen und den damit verbundenen Luftdruckimpulsen widerlegt³⁴ worden ist (Ladefoged et al.

³³ während der ein konstanter Druck durch die auf das Diaphragma einwirkenden Bauchmuskeln herrsche (Stetson 1951: 30)

³⁴ Ladefoged et al. (1958) üben zum einen Kritik an den von Stetson (1951) verwendeten Messmethoden, die keinen eindeutigen Rückschluss auf die behauptete Muskelaktivität zulassen würden (Ladefoged et al. 1951: 1f.). Denn durch zwei dieser Methoden werde die Bewegung bestimmter Körperbereiche gemessen, die den entsprechenden Muskeln benachbart seien, deren Bewegung aber nicht zwingend durch diese Muskeln hervorgebracht werde (ebd.). Ladefoged et al. (1958: 2ff.) verwenden stattdessen zur Überprüfung der von Stetson (1951) behaupteten Muskelaktivitäten die Elektromyographie (ebd.). Diese Messmethode er-

1958). Eine generelle Geltung von Annahmen, die durch die Verallgemeinerung isolierter Phänomene gebildet werden, ist auch nicht zu erwarten. Da Stetson (1951) sich auf die Bewegungen konzentriert, die mit Sprachäußerungen verbunden sind, und akustische Faktoren, die eher als zweitrangig oder nebensächlich erachtet (Stetson 1951: 6) werden, nicht in seine Silbendefinition einbezieht, kann diese auch keinen umfassenden Erklärungswert beanspruchen. Insgesamt ist die Definition der Silbe über die durch die Rippenmuskeln initiierten Luftdruckimpulse zu mechanisch konzipiert, weil ein konstanter physiologischer Mechanismus für das Wesen der Silbe angenommen wird. Stetson (1951) konstatiert zwar, dass das Kind einen für die jeweilige Sprache spezifischen Sprachmechanismus lernen müsse (Stetson 1951: 4) und dass die "habits of utterance of the language constitute an *organized system* [Hervorhebung durch den Verfasser] of skilled movements" (Stetson 1951: 6), aber der Gedanke eines Systems bezieht sich nur auf die Sprachphysiologie ohne Berücksichtigung anderer systematischer Faktoren. Stetsons

laubt die Erfassung der elektrischen Aktivität von einer oder mehrerer zu einer Einheit gruppiert Muskelfasern, die jeweils mit einer Nervenzelle verbunden sind (ebd.), die auch als "motor unit" (ebd.: 3) bezeichnet wird (ebd.). Zwar werde diese Messmethode auch in wenigen Fällen von Stetson (1951) verwendet, deren korrekte Anwendung bzw. Ausführung Ladefoged et al. (1958: 2) jedoch anzweifeln.

Neben dieser sich auf die methodischen Fragen beziehenden Kritik zweifeln Ladefoged et al. (1958) auch die Korrektheit der Messungen und damit ihre Brauchbarkeit zur Begründung von Stetsons (1951) Silbentheorie an: Ihre Überprüfung bestätigte keine Entsprechung von Silben und durch die internen Rippenmuskeln ("internal intercostals") hervorgebrachten Luftimpulse:

"Sometimes a single increase in tension spans a group of articulations including two vowels separated by a consonant closure (our records show that words such as *pity* and *around* may be spoken in this way); and sometimes there are two separate bursts of activity in what is normally regarded as a single syllable (e.g. in *sport*, *stay*, and other words beginning with a fricative followed by a plosive)." (Ladefoged et al. 1958: 6).

Beobachtungen von Ladefoged et al. (1958: 12) hinsichtlich der Aktivität des Diaphragma-Muskels, der bei der Ausatmung wesentlich beteiligt ist (Ladefoged et al. 1958: 12), scheinen die Feststellung zu bestätigen, dass die Bewegung eines Körperteils durch verschiedene Muskeln hervorgerufen werden kann (Ladefoged et al. 1958). Zudem deutet sie darauf hin, dass eventuell auch die physiologischen Funktionen und insbesondere die Muskelaktivitäten wie hier die Diaphragmaaktivität für Menschen individuell ausgeprägt bzw. strukturiert sind. Denn Ladefoged et al. (1958: 12) machen die Beobachtung, dass die Mehrheit der untersuchten Experimentellteilnehmer kaum eine Aktivität des Diaphragmas zeigt, während zwei Teilnehmer eine solche aufweisen. Diese charakterisieren Ladefoged et al. (1958: 12) wie folgt:

"The other two out of our eleven subjects used a very complex pattern of activity of the respiratory muscles. The diaphragm was active not only in inspiration, but also during speech. We were able to study only one of these subjects in detail. With this subject we often recorded activity of expiratory muscles such as the internal intercostals, at the same time as the diaphragm was exerting an inspiratory effort. Thus this subject was using some muscles to decrease the size of the thoracic cavity while he was simultaneously using other muscles to expand the cavity. He regulated the pressure of the air below the vocal cords by balancing the expiratory effort of one set of muscles against the inspiratory action of another. The other subject who maintained the diaphragm in action throughout most of speech also used a similar complex pattern of muscular activity." (Ladefoged et al. 1958: 12)

(1951) Theorie ist auch im Wesentlichen eine der abstrakten Silbenstrukturen, weil neben der Konsonant-Vokal-Unterscheidung und der Unterscheidung der Konsonanten in Silbenkopflaute durch ihre lösende Funktion und Silbenkodakonsonanten durch ihre arretierende Rolle, nur Silbenkonstituenten beachtet werden. Stetsons Beobachtungen haben dennoch einen großen Wert, weil sie auf physiologische Möglichkeiten und Begrenzungen der Sprachproduktion hinweisen und insbesondere Indizien für die Einheit der Silbe liefern, die im Rahmen einer umfassenderen Theorie geprüft werden müssen.

5.4.1.3 Die Frame-Content-Theorie (MacNeilage (1998, 2008))

MacNeilage (1998, 2008) definiert den Begriff der Silbe für die Wortproduktion von Erwachsenen³⁵ als eine abstrakte C...V...C-Silbenstruktur³⁶, auch als "Rahmen" bezeichnet, die die Funktion eines inhaltsleeren, mit "Steckplätzen" versehenen "Behälters" hat (MacNeilage 1998: 502). In diesen sollen während des Abrufs von Lexemen zur Sprachproduktion die getrennt von diesen "Behältern"³⁷ für einzelne Lexeme gespeicherten abstrakten Phoneme eingefügt werden, die auch mit "Inhalt" ("content") benannt werden (ebd.). Die Rahmen werden dabei als mentale Abstraktionen von Kieferoszillationen aufgefasst, die das Kind kurz nach Sprachbeginn durch Selbstorganisationsprozesse bildet (MacNeilage 2008: 188). Der "Inhalt" sind hingegen gleichzeitig zu der Kieferoszillation als dem grundlegenden zyklischen Rahmen stattfindende Bewegungen anderer Artikulatoren wie der Zunge, dem Velum oder der Lippen (MacNeilage 1998: 504f.). MacNeilage (1998: 506ff., 2008: 163ff.) geht zudem davon aus, dass Rahmen und Inhalt durch unterschiedliche neuronale Bereiche unabhängig voneinander gespeichert würden (ebd.).

Seine (MacNeilage 1998, 2008) Theorie, die auf Beobachtungen von Kindern während der Plapperphase und der Produktion der ersten Wörter beruht, begreift die ersten Wörter evolutionär als Adaption von Kieferoszillationen des Kauens bei der Nahrungsaufnahme (MacNeilage 1998: 503): So hat er zum einen auf Einzelsilben bezogen drei phonetisch bestimmte CV-Muster ausgemacht (MacNeilage 1998: 504f.): Vordere Konsonanten werden bevorzugt (d.h. häufiger als der Zufall es erwarten lassen würde) mit vorderen Vokalen, hintere C mit hinteren V und labiale C mit zentralen V produziert (ebd.). Auch die Mehrheit der zweisilbigen Lautformen können ausschließlich als durch Kieferbewe-

³⁵ wie in der Theorie des Lexikonzugriffs bei der Wortproduktion von Levelt et. al. (1999) oder dem Modell zur Erklärung von Versprechern von Shattuck-Hufnagel (1979)

³⁶ V soll die Platzhalter für die Phoneme des Silbenkerns bezeichnen und C die, für die Konsonanten vor und nach dem Silbenkern.

³⁷ vgl. den Begriff der "metrical templates" bei Levelt et. al. (1999)

gungen erzeugt angesehen werden, auch wenn dies nicht zwingend gelten muss (ebd.). Beobachtungen zeigen, dass Vokale und Konsonanten benachbarter Silben sich häufig nur in der Artikulationsart bzw. vertikalen Zungenlage, aber nicht im Artikulationsort bzw. der horizontalen Zungenlage unterscheiden (ebd.). Diese empirischen Beobachtungen verwendet MacNeilage (1998) (neben neuronalen Vergleichen zwischen Primaten und dem Menschen) für seine Argumentation, dass die Rahmen, d.h. die Kieferoszillationen, als grundlegendes zyklisches Konstrukt die Äußerungen des Kindes während der Plapperei und der ersten Wörter dominieren (ebd.: 104f.), während später die Hinzunahme anderer Artikulatoren für "Inhalt" in den geäußerten Schablonen Sorge (ebd.). Deshalb spricht MacNeilage (1998) auch von einer "Rahmen"-Phase und einer "Inhalts"-Phase sowohl für die Sprachevolution als auch den Spracherwerb (ebd.).

MacNeilages (1998) Beobachtungen können als Argument für eine evolutionär sich entwickelnde Lauterzeugung angesehen werden, die anfänglich stärker auf den Unterkiefer als auf die Zunge oder andere Artikulatoren zurückgreift, weil ihr Ursprung auf die natürlichen Kieferoszillationen beim Kauen und bei der Nahrungsaufnahme zurückzuführen ist. Die Schlussfolgerung, dass die beobachteten Silbenmuster allein durch den Unterkiefer erzeugt wurden, ist aber nicht zwingend, weil ähnliche akustische Eindrücke durch unterschiedliche Artikulatoren hervorgebracht werden können. Die frühe Abstraktion eines von den konkret geäußerten Silben unabhängigen mentalen Rahmens wirft die Frage auf, auf welche Weise diese Rahmen entstehen und wie Phoneme als "Inhalt" in diese abstrakten Rahmen eingefügt werden. Die angenommene gegenseitige Unabhängigkeit übersieht auch die in den Silben vorliegende phonetische und phonologische Einheit im Kontext des Silbensystems. Die Fokussierung auf Rahmen, die durch Kieferoszillation festgelegt werden, denen wiederum die Bewegung der anderen Artikulatoren als "Inhalt" gegenübergestellt wird, beachtet nicht die Entwicklung der einzelnen Phoneme, die stattdessen unter diesen beiden mechanistischen Systemannahmen subsumiert werden. Gegen die Annahme der Kieferoszillation als Basis der Silbenbildung, die zudem durch die Abstraktion zu einem mentalen Konstrukt erhoben wird, wurde die Kritik vorgebracht (Oller, Griebel 2008: 151), dass das Kind seine ersten nicht-sprachlichen Äußerungen nicht mit Kieferoszillationen beginnt, sondern zunächst mit reinen Stimmgebungen. Es bleibt auch unklar, warum und wie sich durch Selbstorganisation die abstrakten Rahmen aus den konkreten Kieferoszillationen bilden sollten und falls diese Annahme zutreffend sein sollte, warum dies zu einem so frühen Zeitpunkt stattfindet. Die zentrale Frage ist jedoch, warum überhaupt eine Trennung in abstrahierte Kieferoszillationsrahmen und die Bewe-

gung der anderen Artikulatoren stattfinden sollte. Zudem ist MacNeilages (1998, 2008) Silbentheorie sehr stark auf die Artikulatoren beschränkt, während kognitive oder systembedingte Faktoren außer der fragwürdigen Abstraktion der Rahmen nicht beachtet werden.

5.4.1.4. Die Sonoritätshierarchie

Eine wichtige und umfassende Regelmäßigkeit für die Kombination von Phonemen³⁸ innerhalb der Silbe ergibt sich beim Vergleich der Phonemkombinationen von Silbenkörpern mit denen von Silbenreimen: Während /t/ vor /r/ im Kopf des deutschen Wortes *Trog* steht, haben die beiden Phoneme in der Koda von *hart* die umgekehrte Reihenfolge (Meinschaefer 1998: 32), das Gleiche gilt für andere Phoneme (ebd.): So steht /p/ vor /l/ im Kopf von *Plan*, aber /l/ vor /p/ in der Koda von *Alp* (ebd.); /k/ vor /n/ im Kopf von *Knie*, hingegen /n/ vor /k/ in der Koda von *Schrank* (ebd.); ebenso /f/ links von /l/ im Kopf von *Fliege*, aber rechts von /l/ in der Koda von *Schilf* (ebd.) und genauso /ʃ/ vor /r/ innerhalb des Kopfes von *Schreck*, aber /ʃ/ nach /r/ in der Koda von *harsch* (ebd.). Die genannten Phonempaare /t,r/, /p,l/, /k,n/ und /f,l/ kommen also jeweils in den Körpern der aufgeführten deutschen Wörter³⁹ in der angegebenen Reihenfolge vor, während sich diese in den Silbenkoda der genannten Wörter umkehrt. Für die Mehrheit der Wörter des Standarddeutschen mit CC-Kopf oder CC-Koda lässt sich ebenfalls eine solche Umkehr der Reihenfolge der jeweiligen Phonempaare beobachten. Dies wird beim Vergleich einiger Tabellen (Wiese 1988: 59⁴⁰, 90) deutlich: Die folgende zeigt das (Nicht-)Vorkommen konkreter C₁C₂-Konsonantenpaare im CC-Silbenkopf in allen Wörtern des Standarddeutschen durch ein "+" ("-"),

³⁸ d.h. die Silbenphonotaktik

³⁹ in ihrer Form des Standarddeutschen

⁴⁰ die Tabelle der Silbenkopfkomplexe ist aus Wiese (1988: 59) entnommen, es sind jedoch die Affrikaten weggelassen worden. Von Klammern "()" eingeschlossene "+"-Zeichen bedeuten ein marginales Vorkommen der entsprechenden CC-Kopfkomplexe (ebd.).

Tabelle 2: C₁C₂-Kopf-Komplexe des Standarddeutschen

	l	ʀ	s	n	v	m
p	+	+	(+)	(+)	-	-
t	-	+	-	-	-	-
k	+	+	(+)	+	+	-
b	+	+	-	-	-	-
d	-	+	-	-	-	-
g	+	+	-	+	-	(+)
f	+	+	-	-	-	-
v	-	(+)	-	-		-
ʃ	+	+	-	+	+	+

(aus Wiese 1988: 59 in abgeänderter Form)

während die nächste Tabelle⁴¹ dies analog für die konkreten Folgen aus zwei Konsonanten der CC-Silbenkoda aller Wörter des Standarddeutschen darstellt:

Tabelle 3: C₁C₂-Koda-Komplexe des Standarddeutschen

	ʀ	l	m	n	f	s	ʃ	ç	p	t	k
ʀ	-	+	+	+	+	+	+	+	+	+	+
l	-	-	+	+	+	+	+	+	+	+	+
m	-	-	-	-	+	+	+	?	+	+	?
n	-	-	-	-	+	+	+	+	?	+	?
s	-	-	-	-	-	-	-	-	(+)	+	+
f	-	-	-	-	-	-	-	-	-	+	+
ç	-	-	-	-	-	-	-	-	-	+	-
ʃ	-	-	-	-	-	-	-	-	-	+	(+)

(entnommen aus Wiese 1988: 90)

Fasst man /p,t, k, b, d, g/ zu der phonetischen Klasse der Plosive zusammen, /f, ç, ʃ, s, v/ zu der der Frikative und die Phoneme /m, n/ zu der der Nasale, und nimmt man zusätzlich die Klasse der Vokale hinzu, so ergibt sich die folgende Rangordnung (aufsteigend in Pfeilrichtung), die die zuvor für die Phoneme der Silbenränder beobachtete Reihenfolge widerspiegelt, indem Phoneme einer ranghöheren Klasse im Silbenkörper⁴² rechts und im Silbenreim⁴³ links von Phonemen einer rangniedrigeren Klasse stehen (Wiese 1988: 91)⁴⁴.

⁴¹ entnommen aus (Wiese 1998: 90)

⁴² d.h. der Kombination aus Silbenkopf und dem Silbenkern

⁴³ d.h. der Kombination aus dem Silbenkern und der Silbenkoda

⁴⁴ Die angegebene Rangordnung ist aus (Wiese 1988: 91) entnommen, aber die Vokale wurden hier nicht weiter in hohe und nicht-hohe Vokale differenziert. Die auf der dargestellten Skala markierten Abstände haben keine Bedeutung. Es soll nur eine ordinale Skala, d.h. eine Reihenfolge, aber kein Abstand dargestellt werden.

Abbildung 2: Sonoritätshierarchie des Deutschen



(abgeänderte Version aus Wiese (1988: 91))

Die Plosive haben also den untersten Rang und die Vokale den höchsten. Man nennt den Rang einer Klasse bzw. der sie bildenden Phoneme ihre "Sonorität"⁴⁵ (Sievers 1893; Blewins 1995; Lindner 2009; Meinschaefer 1998; Vennemann 1988, Wiese 1988) und die Rangordnung selbst eine "Sonoritätshierarchie" (ebd.). Die zuvor angesprochene Entsprechung zwischen der Sonoritätshierarchie und der Reihenfolge von Phonemen im Körper oder Reim wird dann auch als "Sonoritätsabfolgeprinzip" (ebd.) bezeichnet. Das erste Sonoritätsabfolgeprinzip, d.h. dass Phoneme (aus Klassen) mit einer höheren Sonorität dichter am Silbenkern stehen oder denselben bilden (Clements 1990: 284), hat Sievers (1893: 185) formuliert (Clements 1990: 284). Sievers (1893) hat für den Begriff der "Sonorität" auch den der "Schallfülle" verwendet (ebd.). Sowohl der Begriff der Sonorität als auch der der Schallfülle beinhalten akustische Termini, die auf eine phonetische Fundierung des Sonoritätsbegriffs verweisen: Neben dem Deutschen lassen sich auch für andere Sprachen auf analoge Weise Sonoritätshierarchien aufstellen, die sich zwar hinsichtlich der Art der Einteilung des Phoneminventars in phonetische Klasse und die Zahl der Klassen unterscheiden (Vennemann, Restle 2001: 1312, Wiese 1988: 91f.), aber sehr starke Ähnlichkeiten aufweisen (Clements 1990: 290). Deshalb wird davon ausgegangen, dass für die Anordnung der Phoneme in Sonoritätshierarchien in den Sprachen inhärente phonetisch-physikalische Eigenschaften der Einzelphonem(klass)e(n) verantwortlich sind, die objektiv gemessen werden können und gemeinsam den Begriff der Sonorität als multidimensionale phonetisch-physikalische Eigenschaft von Einzelphonem(klass)en bilden (Clements 1990: 291; Lindner 2009: 117; Meinschaefer 1998: 53ff.). So konstatiert Meinschaefer (1998: 59) hinsichtlich der von Price (1980 in Meinschaefer 1998: 58ff.) ausgeführten Experimente zur Ermittlung phonetischer Korrelate von silbischem und nicht-silbischem Lateral /l/: "Für das Mehr an Sonorität, das ein silbisches von einem nichtsilbischem /l/ unterscheidet, hat PRICE die phonetischen Parameter Dauer, Lautstärke und Periodizität ermittelt." (Meinschaefer 1998: 60). Das Sonoritätsabfolgeprinzip ist

⁴⁵ Man sagt dann auch, dass ein Phonem x "sonorer als" ein Phonem y ist, wenn x einen höheren Rang als y hat. Aus dem Begriff der "Sonorität" ergibt sich der der "konsonantischen Stärke" (Vennemann 1988: 8f.), indem x eine "größere konsonantische Stärke als" y hat, wenn y sonorer als x ist.

ein grundlegendes und wichtiges Prinzip, weil es die Frage bejaht, ob sich die Phoneme einer Silbe tendenziell in Abhängigkeit von ihrer Silbenposition unterscheiden. Müsste diese Frage verneint werden, wäre dies ein Indiz für die Irrelevanz der Silbe und die alleinige Bedeutung von Phonemfolgen. Durch das "relationale[...] Konzept[...]" (Venne-
mann, Restle 2001: 1312) der Sonoritätshierarchie wird jedoch auf den größeren Silben-
zusammenhang verwiesen und damit die Bedeutung der Einheit der Silbe gegenüber ihrer
Zerteilung in Phonemketten hervorgehoben und betont. Insbesondere wird eine (teilwei-
se) Antwort auf die Frage nach dem Wesen von Silbenkernen gegenüber den Rändern
einer Silbe gegeben, u.a. durch die Festlegung des Kerns einer Silbe als dem Phonem mit
einem ausreichend großen lokalen Sonoritätsmaximum:

Die Fähigkeit, Sonant [d.h. Silbenkern] zu werden, hängt bei jedem Laute zunächst von seiner
natürlichen Schallfülle ab, so dass beim Zusammentreffen mehrerer Laute jedesmal derjenige als Sonant
fungieren muss, welcher an und für sich die grösste Schallfülle besitzt. Nur Laute, welche auf gleicher
oder nahezu gleicher Stufe der Schallfülle stehen, können abwechselnd Sonanten oder Consonanten
sein. In diesem Falle gibt die jeweilige Expirationsstärke statt der natürlichen Schallfülle den Aus-
schlag. (Sievers 1893: 185)

Das Sonoritätsabfolgeprinzip ist zwar ein starker Beweis für die Bedeutung und Einheit
der Silbe gegenüber Einzelphonemfolgen, jedoch sollten die umfassenden Verallgemei-
nerungen von Phonemabfolgebeobachtungen in der Silbe nicht zu einem ursächlichen
Erklärungsprinzip erhoben werden. Denn zum einen gibt es zu den für eine Einzelsprache
aufgestellten Sonoritätshierarchien Gegenbeispiele, die der Sonoritätshierarchie wider-
sprechen und zum anderen systematische Lücken (Ohala, Kawasaki-Fukumori 1997:
346), die zwar nach der Sonoritätshierarchie zugelassen wären, aber in der betreffenden
Sprache nicht vorkommen: So gibt es im Deutschen Wörter wie *Skat* (Wiese 1988: 95),
Skalpell (ebd.), *Sprache*⁴⁶ (ebd.), *Straße* (ebd.), *Sklave* (ebd.), *Skelett* (ebd.), *Skorpion*
(ebd.), *skandieren* (ebd.), *Skala* (ebd.), *skeptisch* (ebd.), die der oben präsentierten Sono-
ritätshierarchie für das Standarddeutsche widersprechen, weil die Silbenköpfe eine Frika-
tiv-Plosiv-Folge enthalten, jedoch Plosive weniger sonor sind als Frikative und deshalb
im Silbenkopf Frikative nicht links von Plosiven stehen dürfen. Statt als ursächliches

⁴⁶ Wiese (1988: 60) legt für die Silbe des Deutschen die Maximalanzahl der in den Rändern erlaubten Pho-
neme auf zwei fest. Demgemäß wären Wörter wie *Sprache* (ebd.: 91), *Sklave* (ebd.), *Straße* (ebd.) nicht nur
wegen der Sonoritätshierarchie ausgeschlossen, sondern auch wegen der Besetzung einer vom Silbenkern
ausgehenden dritten Position durch die jeweils hinsichtlich der Sonoritätshierarchie kritischen Phoneme /ʃ/
und /s/.

Prinzip für die Phonotaktik der Silbe sollte das Sonoritätsabfolgeprinzip wegen der Ausnahmen neben seinem Hinweis auf die Bedeutung und Einheit der Silbe vielmehr als Ausdruck der gemeinsamen physiologischen Grundlage des Menschen gesehen werden, die zu großen Schnittmengen im phonetisch-physikalischen Möglichkeitsraum führen. Jedoch werden das Sonoritätstabfolgeprinzip und die in ihm subsumierte Sonoritätshierarchie auf der Suche nach Sprachuniversalien zu allgemeingültigen Prinzipien erhoben und als angeboren in der "universal grammar" postuliert (Clements 1990: 291):

I will suggest that the sonority scale is built into phonological theory as part of universal grammar, and that its categories are definable in terms of the independently-motivated categories of feature theory [...] (ebd.).

Schwierigkeiten, eine einzige Sonoritätshierarchie für alle Sprachen zu finden, bzw. die Vielzahl der aufgestellten Sonoritätshierarchien (Vennemann, Restle 2001: 1312) deuten auf die Problematik dieses Unterfangens hin. Wie im Fall des Abweichens einer Einzelsprache von ihrer Sonoritätshierarchie gibt es dann für die Annahme einer universalen Sonoritätshierarchie verständlicherweise auch Abweichungen: Würde beispielsweise die Rangordnung, bei der Plosive den niedrigsten Rang haben, gefolgt von Frikativen, Nasalen, Liquiden, Gleitlauten und schließlich Vokalen mit dem höchsten Rang (Ohala, Kawasaki-Fukumori 1997: 343) als universale, d.h. für alle Sprachen geltende Sonoritätskala angenommen werden, so würden die häufig in den Sprachen der Welt vorzufindenden Konsonantenkomplexe /st-/, /xt-/, /vd-/ als Silbenköpfe oder die Silbenkods /-ts/, /-kj/, /-dz/ diese Sonoritätshierarchie verletzen (Ohala, Kawasaki-Fukumori 1997: 347).

Neben solchen Phonemfolgen in umgekehrter Reihenfolge zur angegebenen Sonoritätshierarchie gibt es wie zuvor für das Deutsche auch Lücken, d.h. Phonemfolgen, die generell nach den formulierten Sonoritätshierarchien möglich wären, aber seltener in Silben der Sprachen der Welt vorkommen, als es dem Zufall entsprechen würde (Ohala, Kawasaki-Fukumori 1997: 346), weshalb Ohala et al. (ebd.: 345) eine ebenfalls systematische Behandlung solcher Fälle zusätzlich zu einer Sonoritätshierarchie fordern (ebd.).

So wäre zwar die Folge von Gleitlaut + Vokal nach der oben angegebenen Sonoritätshierarchie erlaubt, aber der Gleitlaut /j/ kommt sehr selten vor vorderen Vokalen als CV-Kombinationen in Silben der Sprachen der Welt vor (ebd.: 346). Diese Lücke ist in den Sprachen "Ainu, Capanahua, Guahibo, Cavineia, Huichol, Korean, Ignaciano Moxo, Dan, Totonaco, Tenango Otomi, Sinhalese, Japanese, Trique, Bulgarian, Mixtec, Yao"

(ebd.) vorzufinden. Die Konstanz⁴⁷, mit der einzelne Sprachen von einer für sie vorgeschlagenen universalen Sonoritätshierarchie abweichen, zeigt⁴⁸, dass die Fokussierung auf Universalien und deren Nachweis die Erkenntnismöglichkeiten der Entwicklung einzelner Sprachen einschränkt, da sie den Blick auf die Individualität der Sprachsysteme verstellt, die jedoch auf gemeinsamer physiologischer Grundlage ihren individuellen Weg beschreiten.

5.4.1.5. Die Rolle der Silbe in Moskowitzs (1970) Phonemerwerbsansatz

Moskowitz (1970) entwirft ein Modell des Phonemerwerbs, in dem das Kind stufenweise eine Rangfolge linguistischer Einheiten unterschiedlichen Umfangs mittels des sog. "transfer process[es]" (Moskowitz 1970: 21) erwirbt: Beginnend mit Sätzen, für die es die Intonationskonturen noch während der vorsprachlichen Phase lerne (Moskowitz 1970: 14), verwende das Kind ab dem Sprachbeginn zunächst die Silbe als Einheit zur Bedeutungsunterscheidung ganzer Einheiten auf phonetischer Basis (Moskowitz 1970: 16ff.), da die Intonationskonturen der Sätze nicht für die Bedeutungsunterscheidung ausreichen. Es folge die Segmentierung der Silben in einer Art Kettenreaktion, nach der nahezu alle Silben schließlich in Segmentform vorlägen und die Rolle der grundlegenden phonologischen Einheit von der Silbe auf das Phonem und die es definierenden distinktiven Merkmale übergehe (Moskowitz 1970: 21, 21ff.). Der Übergang von einer ranghöheren linguistischen Einheit zu einer rangniedrigen wie der von den Intonationskonturen zu den Silben und der von den Silben zu den Phonemen bzw. distinktiven Einheiten erfolgt dabei durch sog. Transferprozesse (Moskowitz 1970: 21). Beim Übergang der Silben zu den Phonemen sollen die phonologischen Speicherungen der Silben durch Phonemfolgen ersetzt werden und "once the transfer is partially completed, all the information which remains encoded in syllable form is extra-systematic" (Moskowitz 1970: 21-2).

Problematisch an Moskowitzs (1970) Ansatz ist, dass die Transferprozesse, die den Übergang von einer ranghöheren zu einer rangniedrigeren linguistischen Einheit umsetzen, unmotiviert sind. Es bleibt unklar, wann und warum sie eintreten bzw. beginnen sollen. Qualitativ werden diese Übergänge als sukzessive von den höheren zu den niedrigeren Einheiten konzipiert, die automatisch aufeinander folgen. Entsprechend einem Lernmodell der sog. "black boxes" (ebd.) orientieren sich die Lautproduktionen des Kindes

⁴⁷ Siehe (Ohala, Kawasaki-Fukumori 1997: 345ff.) für weitere Beispiele des Verhältnisses der Beschaffenheit von Einzelsprachen und einer für sie vorgeschlagenen universalen Sonoritätshierarchie.

⁴⁸ wegen des Überbleibens der Ausnahmen als nicht erklärte Restmenge

quantitativ an der Zielsprache und landen in einem Sortiervorgang nur dann in einer nachfolgenden Einheit, wenn sie ausreichend häufig aufgetreten sind (Moskowitz 1970: 13). Diese Lernkonzeption kann aber nicht die zuvor gestellte Frage nach der Motivation der Transferprozesse beantworten, sondern verstärkt nur den Eindruck eines mechanischen Automatismus. Damit bleibt auch das Wesen der Silbe in der Moskowitzschen Theorie unmotiviert und willkürlich: Zunächst kommt der Silbe, die als Einheit dem die Intonationskontur tragenden Satz folgt, die Funktion zur Unterscheidung von Bedeutungen durch phonetische Unterschiede zu (Moskowitz 1970: 18), die schließlich im Silben-Phonem-Transferprozess an die Phoneme abgegeben wird. Dabei wird die Silbe durch die Phoneme ersetzt (Moskowitz 1970: 21ff.) und Lautformen, die in der Form von unanalytierten Silbenfolgen verbleiben, werden als "extrasystematic" (ebd.) aussortiert (Moskowitz 1970: 21-2). Moskowitz geht zudem von einer Trennung der konkreten Phonemfolge einer Silbe und ihrer abstrakten Silbenstruktur aus. Denn Moskowitz (1970: 17, 26ff.) nimmt an, dass das Kind die abstrakten Silbenstrukturen einer Silbe unabhängig vom Phoneminhalt erlernen kann und dass die ersten Silben des Kindes nur die abstrakte Silbenstruktur CV aufweisen (Moskowitz 1970: 17, 26ff.), aus der sich alle anderen abstrakten Silbenstrukturen (CVC, VC, V, ...) entwickeln (ebd.). Wie bereits an anderer Stelle erwähnt, ist es problematisch, Erwerbsreihenfolgen abstrakter Silbenstrukturen aufzustellen, da auch das häufige Auftreten der CV-Struktur verglichen mit anderen abstrakten Silbenklassen keine strenge Reihenfolge bedingt, da tatsächlich konkrete Silben unterschiedlicher abstrakter Struktur (CVC, VC, V, CV, ...etc.) vermischt im Zeitverlauf vorzufinden sind (Lindner 2009). Man sollte also die Entwicklung der abstrakten Silbenstrukturen nicht getrennt von ihrem konkreten Inhalt betrachten.

5.4.1.6. Die Silbe als Epiphänomen in Ohalas (2008) Kontrastansatz

Ohala (1997, 2008) stellt eine perzeptiv basierte Theorie vor, die die Existenz von Phonemen voraussetzt (Ohala 1997: 355), was mit der Formel "segment first, then syllable" (ebd.) umschrieben wird. Von dieser Annahme ausgehend, soll die Phonotaktik einer Sprache, d.h. das (Nicht-)Vorhandensein aller Phonemfolgen der Sprache erklärt werden (Ohala 1997: 343ff.), die auf das als grundlegend angenommene Prinzip ausreichend großer akustisch-auditiver Unterschiede zwischen zwei in einem Wort direkt benachbarten Phonemen zurückgeführt werden. Ohala (1997: 349ff.) macht zwei Vorschläge, wie diese Unterschiede gemessen werden sollen: Entsprechend dem einen Vorschlag soll jedes Phonem durch eine gleiche Menge akustischer Merkmale, der Amplitude, der Fundamen-

talfrequenz, ...etc. definiert werden (Ohala 1997: 349). Dadurch wird jedes Phonem in einem durch diese Merkmale festgelegten Merkmals-"Raum" positioniert und der Unterschied zwischen den beiden benachbarten Phonemen wird als Länge des Pfades zwischen den beiden Punkten gemessen (Ohala 1997: 350). Abfolgen von zwei Phonemen gelten als bevorzugter, die deshalb auch öfter in den Sprachen der Welt anzutreffen sein sollten, je länger dieser Pfad zwischen ihnen ist. Der zweite Vorschlag (Stevens 1980; Stevens und Keyser 1989 beide in Ohala 1997: 353f.) sieht vor, den Unterschied zwischen den beiden Phonemen durch Merkmale des Übergangs zu messen. Die Merkmale sind bevorzugter bzw. wird angenommen, dass sie in Sprachen häufiger vorkommen, je schneller sie auditiv erkannt werden können (ebd.: 354). Auch wenn durch das auf dem Begriff der Sonorität beruhende Sonoritätsabfolgeprinzip ein Großteil der Silben einer Sprache erfasst werden könnten (Ohala 1997: 343), so lehnt Ohala (ebd.: 344ff.) den Sonoritätsbegriff ab, weil dieser nicht auf eine messbare und damit empirisch überprüfbare Weise definiert sei (ebd.: 344), weshalb eine Definition der Silbe über das Sonoritätsabfolgeprinzip nicht möglich sei (Ohala 1997: 344ff.). Auch wenn man eine Definition der Silbe über das Konzept der Sonorität akzeptiere, blieben oft Ausnahmen wie Silbenränder mit Konsonanten identischer Sonorität z.B. /st/, /xt/ oder /vd/ im Silbenkopf (Ohala 1997: 347), die aber wegen der im Sonoritätsabfolgeprinzip geforderten Zunahme der Sonorität nicht erlaubt seien (ebd.). Abgesehen von einer möglichen Definition der Silbe über den Begriff der Sonorität misst Ohala (1997) dem Konzept der Silbe kaum Bedeutung bei: Der Silbenbegriff sei ebenfalls problematisch, weil er nicht hinreichend definiert sei (Ohala 1997: 344) und außerdem würde er sich in seiner (Ohala 1997, 2008) Theorie der syntagmatischen Phonemkontraste als Epiphänomen ergeben (Ohala 1997: 356): Entsprechend der Definition der Silbe über das Sonoritätsabfolgeprinzip haben die am äußersten Rand der Silbe stehenden Phoneme ein Sonoritätsminimum, während der Kern das Sonoritätsmaximum der Silbe aufweist (ebd.). Diese lokalen Maxima und Minima des Sprachflusses würden sich aber aus dem Prinzip der vorgestellten Theorie ergeben, nach der direkt benachbarte Phoneme paarweise verglichen und dabei Unterschiede bevorzugt werden, die zu möglichst langen Pfaden im zugehörigen auditiv-akustischen Raum bzw. möglichst schnell erkennbaren Merkmalen führen (ebd.). Gäbe es, ausgehend von einem Phonem, keine Erweiterung mehr in eine akustische Richtung⁴⁹, so müsste es eine Umkehr geben. Das habe ein lokales Maximum zur Folge (ebd.). Eine analoge Aussage gelte

⁴⁹ Ohala (ebd.) bringt hier nur das Beispiel des Grades der Vokaltraktöffnung, d.h. die Annahme, der Vokaltrakt ließe sich nicht weiter öffnen.

für die lokalen Minima, weshalb insgesamt die Silbe, definiert über den Sonoritätsbegriff, ein Epiphänomen sei (ebd.). Einige der Funktionen, die Ohala (1997) dennoch für die Silbe in Betracht zieht, scheinen dieser Rolle als Epiphänomen jedoch zu widersprechen: So hält er (Ohala 1997: 356f.) es für möglich, dass die Silbe als Einheit zum einen zur Synchronisation von Segmenten und Gesten dienen könne und zum anderen für eine effiziente Sprachproduktion und -steuerung erforderlich sei (ebd.). Da Ohala (1997) die Phoneme als die grundlegend gegebenen Einheiten annimmt (Ohala 1997: 355), setzt er voraus, dass die Silbe für diese Funktionen durch Zusammengruppierung der Phoneme erst gebildet wird (Ohala 1997: 355). Rein formal widerspricht dies der vorhergehenden Rolle der Silbe als Epiphänomen.

Ohala (1997) legt den Schwerpunkt bei der Definition des Sprachsignals auf die Perzeption und setzt dabei die Existenz ausreichend großer aufeinanderfolgender Phonemunterschiede voraus, ohne die Bedingungen der Sprachproduktion hinreichend zu berücksichtigen, wodurch der Silbe wiederum eine grundlegende Rolle eingeräumt werden müsste.

Als ein Beispiel, dass nicht von der Sonoritätshierarchie erfasst wird, führt Ohala (1997: 347, 357) die Silbe /blo/ an, in der von Sprachteilnehmern bei ausreichender Dehnung des Liquids /l/ die zweisilbige Form /bəlo/ gehört wird. Mit diesem Beispiel will Ohala (1997) zum einen belegen, dass die Sonorität als eine den einzelnen Phonemen inhärente Eigenschaft nicht plausibel ist, wenn eine weitere Silbe gehört würde, obwohl nur das /l/ gedehnt worden sei und diese Dehnung nichts mit einer inhärenten Qualität der Phoneme zu tun haben kann. Zum anderen zieht Ohala (1997: 357) auf Grund derartiger Phänomene die Schlussfolgerung, dass "syllabicity is a perceptual construct, i.e., created in the mind of the listener." (ebd.). Es ist Ohala (1997) nur möglich, diese Phänomene als sich aus seinem Prinzip der ausreichenden akustisch syntagmatischen Kontraste ergebend anzunehmen, weil er die Phoneme als gegeben annimmt (s.o.) und deshalb die Sprache als Folge von Phonemen ansehen und den folgenden Vergleich ziehen kann:

From the point of view of the requirements of the vocal-auditory communication system, syllables do not seem to be the first priority; modulations of the carrier signal are. The segmental stream could be transmitted without being chunked syllabically. Similarly, typing simply requires depressing one typewriter key after another; grouping a certain number of keystrokes together is not required. (Ohala 1997: 355)

Sind jedoch die Phoneme nicht von Anfang an gegeben bzw. angeboren, so stellt sich vielmehr die Frage, in welchem Zusammenhang sie sich entwickeln, und dieser Zusam-

menhang muss anders gedacht werden, als das aufeinanderfolgende Anschlagen von Schreibmaschinentasten.

5.4.2. Notwendige Merkmale einer Silbendefinition

Während in den vorangehenden Teilabschnitten exemplarisch kritische Punkte einiger Silbentheorien aufgezeigt wurden, soll in diesem Abschnitt zusammenfassend auf einige⁵⁰ als notwendig erachtete Merkmale der Silbendefinition hingewiesen werden.

Die Silbe sollte als eine grundlegende und nicht erst abgeleitete Einheit angesehen werden, was zum einen durch die Probleme der Ansätze von Ohala (1997) und des Koordinationsmodimodells der Artikulatorischen Phonologie mit seiner Annahme der Zielorte (*targets*) begründet erscheint und zum anderen durch die frühe ontogenetische Verfügbarkeit der Silbe gegenüber dem Entstehen von Phonemen in einem werdenden Sprachsystem. Die Silbe sollte nicht nach einem statischen Prinzip als für alle Sprachen universell konzipiert angesehen werden. So können universelle Sonoritätshierarchien die Silbenentwicklung nicht erklären, ebenso trifft dies auf statisch-physiologische Erzeugungsmechanismen wie dem Stetsonschen (1951) Rippenmuskelluftimpuls und die Unterkieferoszillation von MacNeilage (1998, 2008) zu. Als universell sollten die zu bestimmenden Prinzipien des dynamischen Silbenwachstums begriffen werden. Die universalen Silbenwachstumsprinzipien lassen sich aber nur erfassen, wenn man von der individuellen Sprache eines Sprachteilnehmers ausgeht und diese als System und damit auch die Silbendefinition als systembezogen begreift. Auch die Verwendungsweise einzelner physiologischer Bestandteile zur Spracherzeugung scheint individuell ausgestaltet zu sein. Darauf deutet die systematische Diaphragmaverwendung von zwei Teilnehmern der Experimente von Ladefoged et al. (1958) bei der Produktion von Sprache hin, während dies auf die Mehrheit der Teilnehmer nicht zutrifft⁵¹. Die Ausbildung ähnlich großer Schnittmengen verschiedener Sprachteilnehmer durch unterschiedliche physiologisch-phonetische Mittel deutet auch auf das Primat der Kognition gegenüber den physiologischen Grundlagen beim Silbenwachstum hin, das in rein physiologisch orientierten Ansätzen von Stetson (1951) und MacNeilage (1998, 2008) missachtet wird. Trotz dieses kognitiven Primats sollte aber die Silbenkonzeption nicht als physiologisch vollkommen unmotiviert Einheit wie bei Moskowitz (1970) definiert werden. Obwohl es als möglich

⁵⁰ Weitere Merkmale der Silbendefinition können dem Phonemerwerbsmodell dieser Arbeit entnommen werden.

⁵¹ siehe Abschnitt 5.4.1.2.

erachtet wird, die Prinzipien des Silbenwachstums auf der durch die Phoneme notierten Höreindrücke gegebenen Ebene herzuleiten, sollte die Silbendefinition im Zusammenhang mit den phonetisch-physiologischen Grundlagen formuliert werden.

6. Phonem(erwerbs)theorien und -aspekte

6.1. Weitere Unterscheidungsmerkmale der vorgestellten Theorien

In diesem Kapitel sollen einige an Diskursmodellen (siehe Kapitel 2) orientierte Phonem(erwerbs)theorien beispielhaft diskutiert werden, um bestimmte Aspekte zu problematisieren und einander gegenüberzustellen. Zunächst werden aber Unterscheidungsmerkmale benannt, die bei der Diskussion dieser Theorien eine Rolle spielen. Sie lassen sich u.a. entlang der folgenden Merkmale einteilen:

1.) Das Phonem als angeborene oder sich entwickelnde Einheit:

Grundlegend können Theorien zum Phonemerwerb, die die (indirekte) Annahme von angeborenen Bestandteilen enthalten, von Ansätzen unterschieden werden, die versuchen, die mit dem Erwerb verbundene Phänomenologie aus der Entwicklung heraus zu erklären.

2.) Das Verhältnis des Phonems zu anderen sprachlichen Einheiten:

Entweder wird das Phonem für sich betrachtet, definiert durch die distinktiven Merkmale, oder aber als Bestandteil einer größeren sprachlichen Einheit wie beispielsweise der Silbe, des Wortes oder des artikulatorischen Musters.

3.) Das Phonem als quantitative oder kategoriale Einheit:

Phonemerwerbstheorien sehen das Phonem als kategoriale Einheit ohne quantitative Eigenschaften (kategoriale Ansätze) oder verbinden mit jedem Phonem bestimmte Mengen wie beispielsweise seine Häufigkeit oder das Ausmaß von Schnittmengen (quantitative Theorien).

4.) Nähere Bestimmung des Phonems als Teil menschlicher Verarbeitungs- oder Verhaltensprozesse:

Phoneme können auf unterschiedliche Weise in das menschliche Sprachverhalten integriert werden. So beispielsweise als kognitive Kategorie im Ansatz der wahrscheinkeitsbasierten Phonologie von Pierrehumbert (2003), als neuronales Netzwerk im

Konnektionismus (Stemberger 1992) oder als hauptsächlich artikulatorische Routine wie in den Schablonen (Vihman, Croft 2007) oder den artikulatorischen Mustern (Piske 2001).

5.) Universalität und Individualität des Phonems als Entwicklungs- und Zielsprachobjekt:

Der Gegensatz von Individualität und Universalität des Phonembegriffs beinhaltet die Frage, in welchem Ausmaß sich Phoneme für verschiedene Sprecher in einer Sprache unterscheiden dürfen. Minimale Individualität ist gegeben, wenn das Phonem als eine Einheit angesehen wird, die allen Sprechern auf identische Weise angeboren ist.

6.) Annahme einer konkreten Bedeutung für die kindlichen Lautformen:

Die kindlichen Lautformen können nach jeweils für sie angenommenen Zielsprachwörtern gruppiert oder aber die Lautformenseite des Kindes bei Annahme der generellen Bedeutsamkeit der kindlichen Äußerungen alleine für sich betrachtet werden.

6.2. Vorstellung einiger Phonem(erwerbs)ansätze

6.2.1. Konnektionismus

In diesem Abschnitt soll der konnektionistische Ansatz am Beispiel des Vorschlags von Stemberger (1992) vorgestellt werden. Auch wenn die konnektionistischen Ansätze breiter gefächert sind und unterschiedliche zusätzliche Annahmen machen, so ist doch bei allen der Bezug auf ein künstliches neuronales Verarbeitungsnetzwerk vorhanden und die mit demselben verbundene Annahme einer Reihe von Verarbeitungsmechanismen, wie sie beispielsweise auch zur Mustererkennung (Marques de Sá 2001: 147ff.) verwendet werden. Das künstliche neuronale Netzwerk besteht aus sog. Knoten bzw. künstlichen Neuronen, die miteinander verbunden sind (Stemberger 1992: 168), wobei jeder Knoten eine Grundaktivierung aufweist, die durch die Interaktion mit den anderen Knoten zu- und abnehmen kann, und zudem den Verbindungen zwischen den Knoten eine Stärke zugewiesen wird, die die Stärke der Aktivierung zwischen den entsprechenden Knoten bestimmt (ebd.). Stemberger (1992) nimmt an, dass dem Kind die Phoneme als exakte Folgen vorliegen, wie sie von den Erwachsenen geäußert werden, und fasst alle Äußerungen des Kindes als Fehler auf (Stemberger 1992: 166), weil das Kind zwar versuche, die ihm exakt vorliegenden Phonemfolgen zu äußern, ihm dies aber nicht gelinge (ebd.). Somit wird der Phonemerwerb des Kindes von Stemberger (1992) als Überwindung eines

fehlerhaften Prozesses gesehen, in dem das Kind als Fehlerverursacher gilt und beispielsweise im Zusammenhang mit der scheinbaren Unterdrückung bestimmter in den Zielsprachformen vorkommender Phoneme als "deleter" bezeichnet wird (Stemberger 1992: 173). Ebenso wie im Ansatz von Smith (2010) wird wegen der Annahme, das Kind verfüge über die Kenntnis der Phoneme, diese Fähigkeit als angeboren vorausgesetzt. Die „Fehler“ werden mit inneren Zusammenhängen im Netzwerk erklärt:

Im Gegensatz zum Erwachsenen, dessen Versprecher Stemberger (1992: 172) in Analogie zu den Äußerungen des Kindes im Phonemerwerb ansieht, mache das Kind aber neben Fehlern, die aus dem fehlerhaften Zugriff auf Grund von zu niedriger Aktivierung eines Knoten resultierten, auch solche, die mit falscher Gewichtung der Verbindungen zwischen den Knoten verbunden seien (Stemberger 1992: 172), die das Kind erst richtig gewichten müsse (Stemberger 1992: 170). Dies geschieht nach Stembergers (1992) Auffassung durch den allmählichen Abgleich der geäußerten Form des Kindes mit dem exakten Vorbild. Das Wesen des Phonemerwerbs ist ebenso wie bei Smith (2010) eine reine Abgleichungsprozedur, die zum Ziel hat, das Netzwerk immer feiner zu justieren, um die Fehler des Kindes fortlaufend zu minimieren. Neben der fehlerhaften Gewichtung von Verbindungen oder der unzureichenden Aktivierung von Knoten bzw. künstlichen Neuronen des kindlichen Netzwerkes werden noch weitere sich aus dem Netzwerk ergebende Eigenheiten als Erklärungen für die spezifische Sprachproduktion des Kindes genannt:

So wird der fehlerhafte Zugriff auf Grund von zu niedriger Aktivierung eines Knotens auch mit der Häufigkeit eines Phonems oder Wortes verknüpft, indem für häufigere Wörter oder Phoneme eine stärkere Grundaktivierung angenommen wird (Stemberger 1992: 170). Weitere Elemente in den angeführten Erklärungen sind ein angenommenes Rauschen bzw. zufällig schwankende Aktivierung, die zum Zugriff auf die falschen Phoneme führe (Stemberger 1992: 171), die Ähnlichkeit von Wörtern (Stemberger 1992: 170) und die von unten nach oben in der Ebenenhierarchie des Netzwerkes fließende Aktivierung (*feedback*) (Stemberger 1992: 174), die zur gleichzeitigen Aktivierung unterschiedlicher Wörter und Phoneme führe und damit zu einer Konkurrenz darum, aus dem Lexikon abgerufen zu werden (Stemberger 1992: 172). Der konnektionistische Ansatz kann als Reaktion auf die generativen Ansätze gesehen werden, bei denen mittels angeborener Regeln die sprachlichen Äußerungen des Kindes aus den dem Kind exakt zur Verfügung stehenden Phonemsequenzen abgeleitet werden (Stemberger 1992: 165f.). Der Konnektionismus wird von seinen Vertretern als realistischer angesehen, weil er die oben beschriebene Theorie der Spracherlernung durch sukzessive Fehlertilgung operationalisiert

hat (Stemberger 1992: 167f.). Der Konnektionismus wurde v.a. deshalb als Erfolg betrachtet, weil mit einem künstlichen neuronalen Netz gezeigt wurde, wie das englische *past-tense* ohne die Annahme von expliziten Regeln wie in der generativen Phonologie von einem künstlichen neuronalen Netz gelernt werden kann (Gerdes 2008: 90).

Auch wenn die in den künstlichen neuronalen Netzen verwendeten Annahmen aus in biologischen Gehirnen beobachteten Phänomenen abgeleitet werden wie beispielsweise das oben genannte Aktivierungsverhalten eines Knotens aus dem für wirkliche Gehirne gemessenen Aktionspotential (White 2002: 1ff.) oder die ebenfalls für künstliche neuronale Netze veränderbaren Verbindungsgewichte aus der wirklich beobachteten Langzeitpotenzierung (Spitzer 2000: 47f.), ist doch zu fragen, ob die auf diese Weise gewonnenen Erklärungen tatsächlich die Realität des komplexen menschlichen Gehirns⁵² und damit generell das *scaling-up* Problem (Bates et al. 1997: 75, Dabrowska 2004: 155) angemessen berücksichtigen und ob sich überhaupt Ergebnisse, die durch die Anwendung künstlicher neuronaler Netze mit ihrer vergleichsweise geringen Komplexität gewonnen wurden, auf Prozesse im Gehirn übertragen lassen (ebd.). Die Lernfähigkeit neuronaler Netze ist auch schematisch und erfordert ein umfassendes Training. Beide Eigenschaften treffen offensichtlich nicht auf die Arbeit des Gehirns und den Spracherwerb zu. Das Kind baut zunächst nur sehr langsam ein Lexikon auf, beschleunigt aber dann die Produktion von Silben und Wörtern. Diese unterschiedlichen Produktionsgeschwindigkeiten deuten auf komplexe geistige und sprachliche Prozesse hin, die sich nicht mit einfachen Computersimulationen auf der Grundlage eines vagen, neuronal basierten Plausibilitätskonzeptes ergründen lassen. Es ist sicher zu erwarten, dass durch die Fortschritte beim Verständnis der Funktionsweisen des menschlichen Gehirns und in der Datentechnologie in Zukunft wertvolle Erkenntnisse für die Erklärung und das Verständnis des Spracherwerbs gewonnen werden können, aber hier soll für einen linguistischen Ansatz plädiert werden, da die "Spuren", die das Kind mit seinen sprachlichen Äußerungen hinterlässt, Hinweise zur Erklärung des Phonemerwerbs liefern, indem man die inneren Gesetzmäßigkeiten auf der Formenseite der vorliegenden, zeitlich geordneten Transkriptionen aufzeigt. Auch wenn der klassische Phonembegriff (Trubetzkoy 1962) im Phonemerwerb nicht in seiner intendierten Form als kleinste bedeutungsunterscheidende Einheit anwendbar ist, so gilt doch analog für die Trennung der vagen Netzwerklogik und der linguistischen Formenseite das

⁵² "Für die im Großhirn verlaufenden Nervenfasern wird eine Gesamtlänge von 500.000 km angenommen. Aufgrund der enormen Komplexität des sich daraus ergebenden Netzwerks von signalverarbeitenden Zellen lässt sich die Komplexität des Gesamtsystems Gehirn nur schwer begreifen." (Müller 2003: 58)

Gleiche, was Trubetzkoy (1962) zur Trennung von Psychologie und Linguistik gesagt hat, dass nämlich

[die] Heranziehung der Psychologie [...] bei der Definition des Phonems vermieden werden [muss], denn das Phonem ist ein sprachwissenschaftlicher und nicht ein psychologischer Begriff. Jede Bezugnahme auf das 'Sprachbewusstsein' muss bei der Definition des Phonems ausgeschaltet werden. Denn das 'Sprachbewusstsein' ist entweder eine metaphorische Bezeichnung des Sprachgebildes ("langue") oder ein ganz vager Begriff, der selbst noch definiert werden muss und vielleicht gar nicht definiert werden kann. (Trubetzkoy 1962: 37f.)

Dies bedeutet aber weder, dass neuronale oder andere verhaltensbasierte Konzepte als unwichtig angesehen werden, nur können dieselben in ihrer jetzigen Form nicht zur Erklärung der linguistischen Formenseite angesehen werden, noch bedeutet es, dass eine "mere curve fitting" betrieben werden soll, wie Lindblom (1999) sie anprangert.

6.2.2. Jakobsons Theorie der distinktiven Merkmale und deren Anwendung auf den Phonemerwerb

Jakobsons (1962) Ansatz zum Phonemerwerb soll hier vor allem deshalb vorgestellt werden, weil er, wie Moskowitz (1970: 1f.) zu Recht betont, gegen die Auffassung, die Sprache eines Kindes verkörpere rein Zufälliges, mit seiner systematischen und umfassenden Theorie eine klare Gegenposition bezogen hat. Von einer Systemperspektive aus gesehen ist jedoch jedes Element eines Systems unabhängig von seiner Quantität von Bedeutung. Jakobson hält dagegen das Häufige für grundlegend und schränkt dadurch willkürlich das Systemverständnis ein. Auch die von ihm vorgenommene Zuweisung von Bedeutungen zu Kindersprachformen muss willkürlich bleiben, da es zwischen kindlichen Sprachformen und ihren Bedeutungen und den entsprechenden Formen in der Sprache der Erwachsenen nur Schnittmengen geben kann, aber niemals eine vollständige Identität.

Jakobson (1962) nimmt eine Hierarchie von durch distinktive Merkmale festgelegten Phonemklassen an, wobei die Hierarchie durch die zeitliche Reihenfolge der jeweiligen Verwendung des ersten Phonems einer Klasse festgelegt ist. So nimmt Jakobson (1962) an, dass nach einer Anfangssilbe, die aus einem tiefen Vokal, meist /a/, zusammen mit einem labialen Plosiv wie /b/ oder /p/ gebildet wird, mit einem nasalen Konsonanten zu den bereits aufgetretenen oralen Konsonanten in Form der Labialen der erste konsonantische Gegensatz artikuliert wird. Darauf folgt der zweite konsonantische Kontrast von neu auftretenden Dentalen mit Labialen usw. Es wird jedoch von Kindern berichtet, die

erst einen Kontrast zwischen Dentalen und Velaren bilden, bevor überhaupt Labiale aufgetreten sind (Menn 1976 in Johnson, Reimers 2010: 54). Die Kontraste zwischen den Vokalen werden mit denen zwischen den Konsonanten in ein und derselben Hierarchie erfasst. So folgt entsprechend Jakobson (1961: 358) nach den zuvor genannten konsonantischen Kontrasten einer zwischen einem tiefen und einem hohen Vokal, beispielsweise der zwischen dem /a/ und dem /i/ (ebd.). Der nächste Kontrast soll dann auf eine von zwei Arten zwischen Vokalen etabliert werden: Als die eine Möglichkeit nimmt Jakobson (ebd.) die Etablierung eines weiteren Vokals an, der durch Verändern der Zungenhöhe von einem der bereits bestehenden Vokale /a/ zum /i/ erhalten wird, beispielsweise der Vokal /e/, so dass die dann drei Vokale den sog. "linearen Vokalismus" bilden. Als die andere Möglichkeit nimmt Jakobson (ebd.) hingegen einen Vokal an, der in Opposition zu dem bereits vorderen Vokal gesetzt wird, so dass sich z.B. das "Vokaldreieck" /a,i,u/ ergeben könnte. Aber auch hinsichtlich dieser beiden einzigen von Jakobson vorgesehenen Vokaldreiersysteme ist angemerkt worden, dass im Englischen beispielsweise zu /a,i/ statt eines hohen, hinteren Vokals ein tiefer, hinterer Vokal wie beispielsweise /ɔ/ kontrastierend gebildet wurde (Menn, Vihman 2011: 265). Jakobson (1961) nimmt an, dass die von ihm aufgestellte und hier zuvor nur bzgl. der ersten distinktiven Merkmale aufgeführte Hierarchie ein grundlegendes strukturelles Wesen der Sprache ist, was er durch seine Annahmen bestätigt sieht, dass Phoneme aus niedrigeren hierarchischen Klassen dieser Hierarchie in den Sprachen der Welt zahlreicher sind als solche aus höheren Klassen (Jakobson 1961) und dass Aphasiker erst alle Phoneme einer höheren Klasse verlieren, bevor ihnen solche aus der nächst tiefergelegenen Ebene nicht mehr zugänglich sind (Jakobson 1961). Jakobson (1961) geht davon aus, dass die von ihm beschriebene Hierarchie für jeden Menschen gilt, der eine beliebige Sprache lernt oder als Aphasiker verliert, d.h. universal ist. Dies begründet er im Wesentlichen damit, dass sie angeboren ist (1968: 136): So zitiert er Teuber (1976 in Jakobson 1986: 134), der gesagt haben soll, dass

die 'distinktiven Merkmale' mehr als ein universelles Schema für die Klassifizierung von Phonemen in ihrer Mannigfaltigkeit von Sprache zu Sprache sein [sollten]; die Merkmale sollten 'wirklich' sein in dem Sinne, dass sie universellen neurologischen Mechanismen für die Produktion und der Wahrnehmung von Sprachlauten entsprechen (Teuber 1976 in Jakobson 1986: 134).

Auf die Kritik von Ferguson und Farwell (1975 in Jakobson 1986: ebd.), dass er individuellen Unterschieden beim Phonemerwerb mit seinem Modell keine Beachtung schenke, antwortete Jakobson (1986: 173):

Diese Variationen heben die offenbare Invarianz nicht auf und verbergen sie auch nicht, und auch wenn die "Beschreibung und Erklärung von solchen Unterschieden" von großem Interesse für Linguisten sind und "Relevanz für Therapie und Pädagogik" haben, so können solche Varianten nicht völlig verstanden werden, ohne dass die hinter ihnen verborgenen Invarianten herausgeschält und erklärt werden. Mit Stanley S. Stevens (1951: 21) gesprochen ist es nicht der Fall, "dass die Variabilität jemals aus der empirischen Wissenschaft verbannt wird, aber die fortlaufende Ausbreitung des Bereiches der Invarianz kann die Herrschaft der Varianz auf überschaubarere Proportionen reduzieren. (ebd.).

Jakobsons Hierarchie ist aber auch vor dem Hintergrund zu sehen, dass die häufigen Lautzusammenhänge in den Sprachen dazu verleiten, das Häufige zum grundlegenden Status, zur wesentlichen Struktur oder zum Entwicklungsprinzip zu erheben und das Abweichende als Ausnahme, (krankheitsbedingte) Fehler oder Zufall anzusehen. Doch es gibt kaum lautliche Gegebenheiten, die absoluten Universalien entsprechen (Bybee 2004: 191), d.h. zu den meisten Aussagen zur Lautstruktur von Sprachen gibt es auch einzelne Sprachen, für die diese Aussagen nicht zutreffen. Deshalb sei es, so Bybee (ebd.), nicht sinnvoll diese Beinahe-Universalien als angeboren anzunehmen, sondern eher als Tendenzen, die sich aus der Art der Speicherung und Verarbeitung von Sprache beim Menschen ergäben (ebd.). Als Lösung des Problems schlägt Bybee (2004) vor, all die Mechanismen herauszufinden, die im Zusammenspiel bestimmte sprachliche Veränderungen herbeiführen und sieht in denselben bessere Universalien. Aber auch dieses Vorgehen ist zu schematisch und beinhaltet eine Logik der Rekurrenz von Teilprozessen, die mit einer gewissen Wahrscheinlichkeit in einer bestimmten Lautumgebung Anwendung finden. Die wirklichen Universalien der Sprache sind jedoch die, die auf der Ebene des Individuums wie auch von Sprachgemeinschaften die Entstehung von sprachlicher Struktur erklären können, ohne sich der Konstruktion von Ausnahmen und angeborenen Regeln zu bedienen. Jakobson (1961) hat dem Gedanken eines Systemzusammenhangs Gewicht verliehen, weil er entgegen Studien, die lediglich Einzelphoneme registriert haben, versucht hat, dieselben in einen größeren bedeutungsvollen Zusammenhang zu stellen, der gewissen Gesetzmäßigkeiten gehorcht und nicht zufällig ist. Zudem hat er durch die Hierarchie der distinktiven Merkmale zum einen eine zeitliche Dimension zur Erklärung des Phonemerwerbs hinzugenommen und durch den Gedanken, dass bestimmte Merkmale für ihr

erstes Auftreten die Existenz von anderen erfordern, darauf hingewiesen, dass es innere strukturelle Notwendigkeiten bei der Veränderung der Kindersprache und der Sprache als solcher geben muss. Außerdem sind die distinktiven Merkmale nicht unbedeutend, denn sie reflektieren auch ein grundlegendes Element der Silbenbildung, aber sie erklären dieselbe nicht alleine. Da Silben die lautlichen Grundroutinen eines geistigen Vorgangs sind und die Phoneme die gedachten oder empfundenen sequentiellen illusionären Einheiten, die zwar auf phonetischer Grundlage⁵³ gebildet und beeinflusst werden, aber nicht von ihr determiniert werden, können phonetische Gegensätze nicht ihr Urheber sein. Der von Jakobson hier vorgestellte Ansatz ist auch deshalb für das Verständnis des Phonemerwerbs problematisch, weil in ihm geschaut werden muss, wann ein Phonem einer Lautform des Kindes bedeutungsunterscheidend verwendet wird⁵⁴, um so die erstmalige Verwendung des zugehörigen distinktiven Merkmals konstatieren zu können.

6.2.3. Bybees Gebrauchsbasierte Phonologie (2001)

Bybee (2004) präsentiert einen sog. gebrauchsbasierten Ansatz zur Phonologie, den sie so nennt, weil sie annimmt, dass die Art, wie Sprache im Gebrauch verwendet werde, die Form derselben bedinge (Bybee 2004: 1, 57). Sie nimmt zudem an, dass es für die Sprache keine speziellen Speicherungen im Gehirn gebe, sondern nur allgemeine für alle Wahrnehmungsobjekte gleichermaßen geltend (Bybee 2001: 6). Für die Modellierung dieser Annahmen (Bybee 2004) verfolgt sie u.a. einen Exemplar-Ansatz (Wedel 2006: 252) für die Produktion und Perzeption von Wörtern, woraus folgt, dass alle Wörter, wie sie tatsächlich von einem Sprachteilnehmer während des Sprachgebrauchs bzw. der Sprachtätigkeit produziert oder wahrgenommen werden, auch im Lexikon als Speicherungen vorkommen (Bybee 2001). Als problematisch an diesem Ansatz wurde eine mögliche Überfüllung angesehen, wenn es im menschlichen Gehirn nicht genug Speicherplatz

⁵³ Auch Locke (1983) begreift methodisch Regelmäßigkeiten der kindlichen Sprache als unvollkommene Ersetzungen der zielsprachlichen Formen. Statt aber auf ein angeborenes Regelwerk zurückzugreifen, verweist er (ebd.) auf die artikulatorischen Möglichkeiten und Begrenzungen des Kindes (ebd.), die er schon in der Plapperphase für relevant hält (ebd.).

Dieser Ansatz kann allerdings ebenso wenig wie der von Jakobson (1941) die Dynamik der kindlichen Sprachentwicklung erklären, da eine Fokussierung auf phänomenale Aspekte der Zielsprache und der kindlichen Sprache stattfindet. Die Triebkräfte und Regeln der kindlichen Sprachentwicklung sind primär nicht als Transformation von zielsprachlichen Äußerungen in kindliche Lautäußerungen zu sehen und auch nicht in den physiologischen Grundlagen zu suchen (ebd.), allerdings verweisen die von Locke (1983) festgestellten Gemeinsamkeiten in den Sprachäußerungen des Kindes auf die gemeinsame physiologische Basis aller menschlichen Sprachen.

⁵⁴ Siehe auch den Abschnitt 4.7.

geben würde, was Johnson (1997: 146) als "head-filling-problem" (ebd.) bezeichnet und gleichzeitig aber darauf aufmerksam macht, dass der Mensch eine "apparent availability of unlimited memory in picture recognition" (Johnson 1997: 147) hat und somit vermutlich keine Platzprobleme vorhanden seien. Die Annahme einer vorkommensbasierten Speicherung von Wörtern bedeutet, dass von jedem Wort nicht nur eine phonetische Form gespeichert wird, sondern alle bisher angetroffenen Varianten desselben, wobei Bybee (2001: 52) (im Zusammenhang der phonetischen Kategorisierung) einschränkt, dass "tokens that are sufficiently similar are stored together, yielding a strengthened representation. At the same time, rarely used, nonsalient variants might fade from memory." (Bybee 2001: 52). Als grundlegendes Speicherungsformat der Wörter nimmt Bybee (2001: 31) die Gestenpartituren ("gestural score representation" (Bybee 2001: 31) der Artikulatorischen Phonologie zusammen mit akustisch-perzeptiven Speicherungen (Bybee 2001: 85) an. Die Gesten sind die für die Kodierung der eigenen Produktionen eines Sprechers von vornherein gegebenen sprachlichen Grundeinheiten, durch deren sequentielle und simultane zeitliche Verteilung nach der Auffassung der Artikulatorischen Phonologie alle größeren sprachlichen Einheiten wie beispielsweise Phoneme, Silben oder Wörter adäquat beschrieben werden können⁵⁵. Bybee (2001: 21, 26) führt den Begriff des Schemas ein, der ausdrücken soll, dass sich linguistische Regelmäßigkeiten aus einer netzwerkartigen Lexikonstruktur ergeben (Bybee 2001: 21), bei der "[s]imilar or identical properties of meaning and form are associated with one another across items" (Bybee 2001: 23). Deshalb ist der Schemabegriff als eine (teilweise nicht näher spezifizierte) in der Lexikonstruktur manifestierte Kategorisierung der gespeicherten Wortvarianten (Bybee 2001: 21, 27, 28) zu verstehen. Beispiele von Schemata seien alle Varianten eines bestimmten Wortes wie z.B. *send* (Bybee 2001: 32), alle Wortvarianten mit dem Reim *end* (ebd.), alle mit der Koda *Vend*, wobei V für einen beliebigen Vokal steht (ebd.) oder alle mit dem Merkmal *Vokal-Nasal-stimmhafter Plosiv* (ebd.). Die letztgenannten Beispiele sollen zudem auch verdeutlichen (ebd.), dass ein Schema sich auf unterschiedlich abstrakte Regelmäßigkeiten beziehen kann (ebd.), weil mit jedem der zuletzt aufgeführten Beispiele jeweils in der Reihenfolge ihrer Nennung potentiell zunehmend mehr Wortvarianten verbunden sind. Aus den mittels Gesten kodierten bzw. gespeicherten Wortvarianten ergeben sich nach Bybees (2001: 31, 85f.) Vorstellung andere Einheiten wie beispielsweise Phoneme, Silben und Morpheme durch Schemabildung. Silben sind

⁵⁵ Siehe auch Abschnitt 6.2.7.

somit nach ihrer (Bybee 2001: 23ff., 31, 85f.) Auffassung nicht grundlegende Einheiten, sondern Verallgemeinerungen:

Generalization to a notion of syllable, however, occurs through schema-formation. Similarities among stretches of speech in terms of the temporal coordination of gestures will lead to an abstract notion of possible syllables for a language. (Bybee 2001: 31).

Die Artikulatorische Phonologie (Browman, Goldstein 1986) hat aber keine Erklärung für das Zustandekommen der in ihrem Ansatz angenommenen Gesten mit den "targets" als festen Verengungsorten, während die Silben für jeden Sprachteilnehmer intuitiv verfügbar sind. Auf welche Weise leitet nun Bybees (2001) Ansatz die Veränderungen von sprachlichen Formen aus ihrem Gebrauch ab?

Bybee (2001) greift im Wesentlichen auf die folgenden vier Modellkonzepte zurück: Erstens ist ein zentraler Punkt bei der Besprechung sprachlicher Phänomene durch Bybee (2001) ihre Verwendung von Exemplar-Kategorien gepaart mit der Konzeption des Schemas: Dadurch, dass sie annimmt, dass jeder wahrgenommene oder produzierte Stimulus eines Wortes⁵⁶ sehr detailgetreu gespeichert werde, spielt die Häufigkeit einer Variante eines Wortes in ihrem Modell eine wichtige Rolle: Schemata, d.h. die Zusammengruppierung von Wörternvarianten im Lexikon hinsichtlich ähnlicher Teile, verändern sich mit den gespeicherten wahrgenommenen oder produzierten Wörternvarianten. Tritt ein Wort sehr häufig auf, so werden viele Varianten von ihm im Lexikon gespeichert oder bestehende verstärkt. Damit sind diese auch wiederum leichter abrufbar, werden aber auch durch ihre häufige Verwendung immer mehr automatisiert und reduziert. Selten auftretende Wörter hingegen verlieren zunehmend an Speicherungsstärke. Zweitens wird diese Häufigkeit von bestimmten Varianten mit Merkmalen von häufig ausgeführten und damit geübten Bewegungen (Bybee 2001: 58, 71) verbunden, indem angenommen wird, dass andere geübte Bewegungen wie Klavierspielen (Bybee 2001: 14) mit der beim Sprechen ausgeführten Bewegungen vergleichbar sind: So unterlägen häufig ausgeführte Bewegungen Veränderungen wie u.a. zunehmender Automatisierung und zunehmender Reduktion, was auch gleichermaßen bei häufig geäußerten Wortvarianten der Fall sei. Drittens greift sie (Bybee 2004) auf die Abhängigkeit des Schemabegriffs von den lexikalischen Einheiten zurück: Diese äußert sich beispielsweise in einer Prototypenorganisation

⁵⁶ Ein Wort definiert Bybee (2004) als "a unit of usage that is both phonologically and pragmatically appropriate in isolation."

der Kategorien: Die Zugehörigkeit von Wortvarianten zu einem Schema ist nicht absolut im Sinne von (nicht) zugehörig, sondern eine Wortvariante repräsentiert eine Kategorie umso besser, je näher sie dem Zentrum der Verteilung mit den meisten Wortvarianten und damit einem gedachten Prototypen ist. Deshalb bezeichnet Bybee (2001) Schemata auch als "gradient categories" (ebd.). Viertens versucht Bybee (2001) Veränderungen von Lautfolgen durch Einordnung in bestimmte räumliche und zeitliche Gestenkonstellationen plausibel erscheinen zu lassen, wie dies auch schon im Ansatz von Studdert-Kennedy (1992) getan wird. Bybee (2004: 15) nimmt für den Erwerb der Kinder auch v.a. ein wortbezogenes Lernen an, was nicht unproblematisch ist⁵⁷.

Die nach Bybees (2004) Ansatz möglichen Veränderungen sind vor allem solche, die sich aus dem Verhältnis einer Kategorie und der sie durch Ähnlichkeitsrelationen bildenden Wortvarianten ergeben: So verschiebt sich die Kategorie qualitativ mit ihren Exemplaren und nimmt mit der Häufigkeit derselben auch an Häufigkeit und Stärke bzw. Zugänglichkeit beim Abruf aus dem Lexikon zu. Alle Eigenschaften der Exemplare übertragen sich also auf die Kategorie. Zudem kann durch häufigen Gebrauch eine Lautform automatisiert und verkürzt werden. Dieser Ansatz hat den Vorteil, dass er durch diese Kategorien-Varianten-Verhältnisse die Verbindung zum tatsächlichen akustischen Sprachsignal nicht verliert. Doch fehlen vollkommen funktionale Aspekte, die eine von innen motivierte Entwicklung bedingen könnten. Außer der Kategorisierung von Varianten und Kategorien nach Ähnlichkeit hat der Ansatz keine weiteren inneren Komponenten und so spiegelt sich auch wie bei Pierrehumbert (2003) die äußere Häufigkeitsverteilung als angenommener Einflussfaktor der Entwicklung. Eine Relation zwischen Häufigkeit und Relevanz anzunehmen, kann naheliegen aber auch ebenso in die Irre führen, denn Naturprozesse wie die Sprachentwicklung bedienen sich unterschiedlicher Möglichkeiten, um ihre Funktion zu erfüllen. Das Konzept der Schemata ist aber auch willkürlich, und wird nicht aus einer inneren Systemlogik entwickelt. Insgesamt geht der Ansatz von Bybee (2004) vom Wirken eines Regelsystems aus, dessen Entstehung oder Veränderung nicht thematisiert wird und deshalb zu den Theorien gerechnet werden muss, die den Kernbereich der Sprachentwicklung für angeboren halten.

⁵⁷ Siehe hierzu auch den Abschnitt 4.7.

6.2.4. Vihmans und Crofts sog. "Radikale schablonenbasierte Phonologie"

Vihman und Croft (2007) definieren den in ihrer "Radikalen"⁵⁸ schablonenbasierten Phonologie" verwendeten Begriff der "Schablonen" als eine Menge von bevorzugten Wortbildungsmustern (Vihman 2007: 697), die zusammengenommen alle (Vihman 2007: 697) der anfänglich (Vihman 2007: 692) vom Kind geäußerten Lautformen subsumieren, wobei jede Schablone für sich ausreichend vielen Kinderformen als Muster dienen muss (Vihman, Croft 2007: 663). Die Definition der Schablonen beinhaltet zudem, dass die früh von einem Kind geäußerten Zielsprachwörter, auch "selected" genannt, besser durch eine der Schablonen angenähert werden als dies für die vom Kind später artikulierten Zielsprachwörter, auch mit "adapted" bezeichnet, der Fall ist (Vihman, Croft 2007: 696). Die genaue Artikulation der von Vihman und Croft (2007) für die kindlichen Lautformen angenommenen Zielsprachlautformen sehen sie (Vihman, Croft 2007) durch den oben bereits im Abschnitt 4.6.2. erwähnten sog. artikulatorischen Filter als erklärt an, durch den die zu den kindlichen Äußerungen aus der Plapperphase sehr ähnlichen Zielsprachlautformen besonders hervorgehoben würden (Vihman, DePalois 2009: 168ff.). Vihman und Croft (2007) bringen eine Reihe von Beispielen von geübten Produktionsmustern, u.a. auch für das Kind Annalena (Elsen 1991): Bei diesem die deutsche Sprache lernenden Kind beziehen sich die ausgemachten Produktionsmuster v.a. auf die Formen abstrakter Silbenstrukturen, in diesem Fall CV, CVCV, CVC, wobei für die CVCV - Formen zusätzlich die Identität der beiden Vokale (Vokalharmonie) und der Konsonanten (Konsonantenharmonie) angenommen wird. Neben diesen CV-Reduplikationen wird auch eine C₁VC₂V-Form mit labialem C₁ und alveolarem C₂ als Schablone angenommen. Zusätzlich zu diesen Beschränkungen (abstrakte Silbenstrukturen, Auftreten bestimmter Konsonantenkombinationen in CVCV-Formen hinsichtlich phonetischer Merkmale) der vom Kind produzierten Lautformen, werden auch vereinzelt ganz konkrete Phoneme angegeben, so der zweite Vokal als /ɪ/ oder /ʊ/ in Diphtongen VV. Darüber hinaus gibt es aber keine Betrachtung der konkreten Laute bzw. der vom Transkribierenden angegebene-

⁵⁸ Wie in anderen Ansätzen, die linguistische Struktur durch Konstruktionen beschreiben, wie z.B. die Syntax in der "Radical Construction Grammar" von Croft (2001), werden hier die positionsbezogenen Phoneme als Einheiten in einer größeren, sie umfassenden Struktur aufgefasst und definiert und deshalb von Vihman und Croft (2007: 712) als grundlegend ("basic" (ebd.), "primary unit" (Vihman, Croft 2007: 714)) bezeichnet, während sie die Phoneme aus diesen als sich durch Kategorisierungs- bzw. Verallgemeinerungsprozesse ableitend ("derivative" (ebd.)) betrachten (Vihman, Croft 2007: 711f., 718). In Anlehnung an die "Radical Construction Grammar" von Croft (2001) versehen sie (Vihman, Croft 2007: 711f., 718) ihren Phonemansatz deshalb mit dem Attribut "radical".

nen Phoneme. Da von Vihman und Croft (2007) nur häufige Merkmale (v.a. abstrakte Silbenstrukturen, vereinzelt konkrete Laute) in ganzen Lautformen betrachtet werden und so die detaillierten und graduellen quantitativen Veränderungen einzelner Laute in unterschiedlichen Kontexten verlorengelassen, kann auf Grund dieser Reduzierung von Lautformen ein innerer systematischer Zusammenhang nicht hergestellt werden. In der Schablonenphonologie von Vihman und Croft (2007) werden also einzelne konkrete Phoneme nur sporadisch betrachtet, und dienen, sofern sie häufig genug vorkommen, wie die anderen hochfrequenten Merkmale zu einer Zusammenfassung bzw. Einteilung von Silben bzw. Lautformenmengen im Zeitverlauf, wodurch die individuelle phonembezogene quantitative Entwicklung nicht beachtet, sondern deren quantitative Entwicklungskontinuität (durch die Zusammenfassung unterschiedlicher Silben mit diversem segmentalen Inhalt zu einer schablonenbezogenen Silbenmenge) zerrissen wird. Auch wenn es sich bei den Schablonen um einen Begriff handelt, der auf einem Häufigkeitskriterium basiert, so ist doch andererseits positiv zu sehen, dass die Schablonen sich deshalb für verschiedene Kinder unterscheiden können. Vihman und Croft (2007: 707) nehmen die Schablonen nicht als universale, angeborene Strukturen oder für alle Kinder einer Sprache gleichermaßen gegeben an, sondern als individuelle Ausprägungen, die sich für jedes Kind individuell durch die perzeptive Vertrautheit mit Phonemfolgen der Zielsprache, durch zunehmende motorische Übung mit einem Teil dieser Phonemfolgen u.a. auch während der Plapperphase und durch Vertrautheit mit seinem eigenen, bereits gebildeten Lexikon ergeben (ebd.). Jedoch bleibt unklar, wie es nach der Definition der für die ersten Lautformen des Kindes über das Häufigkeitskriterium definierten Schablonen mit der Entwicklung derselben weitergehen soll. Hier macht sich die Zusammenfassung von Lautformen zu Schablonen und damit die Missachtung des eigentlich vorhandenen, größeren Lautformenumfangs bemerkbar durch eine wiederum willkürliche Eingrenzung. Die von Vihman und Croft (2007: 689f.) aufgeführten Gründe für die Annahme der Schablonen statt des Phonems sollen im Folgenden aufgeführt werden, weil sie in diesen wortbezogenen Ansätzen zugleich auch als Argumente gegen die Silbe verstanden werden. Dabei wird aber missachtet, dass sich alle im Folgenden aufgeführten Gründe überhaupt erst aus der Gruppierung der kindlichen Lautformen zu für sie angenommenen Zielsprachformen ergeben⁵⁹: Als ersten Grund sehen sie an, dass zwischen den vom Kind geäußerten Lautformen oft eine größere Ähnlichkeit besteht als zwischen diesen und den für sie ange-

⁵⁹ Siehe dazu auch die Diskussion im obigen Abschnitt 4.7.

nommenen Zielsprachformen (Vihman, Croft 2007: 690, Vihman, DePaolis, Keren-Portnoy 2009: 173). Dies ist in jedem Fall ein Argument dafür, dass das Kind eine eigene Laut- und Sprachstruktur ausbildet, aber nicht unbedingt dafür, dass man durch Häufigkeitsverallgemeinerungen, wie sie die Schablonen darstellen, die feinen Details der einzelnen konkreten Phoneme von der Untersuchung ausschließt. Zweitens sei es oft schwer eine Relation zwischen einer Kindform und einer Zielsprachform segmentweise herzustellen (Vihman, Croft 2007: 690, Vihman, DePaolis, Keren-Portnoy 2009: 173). Auch dies ist lediglich ein Problem, welches sich ergibt, wenn man direkte Vergleiche zwischen konkreten Kinder- und Zielsprachformen vorzunehmen versucht, wodurch die unterschiedlichen Systemzusammenhänge ignoriert werden. Drittens werde ein und dasselbe Phonem der Zielsprache in verschiedenen Kinderlautformen durch unterschiedliche Phoneme wiedergegeben und einige Zielsprachlautformen würden in mehreren unterschiedlichen Lautformen ("variability of segment production" (Vihman, Croft 2007: 709)) vom Kind produziert (Vihman und Croft 2007: 269, Vihman, DePaolis, Keren-Portnoy 2009: 172). Dies zeige, dass das Kind zwar Wissen von bestimmten Wörtern gespeichert habe, aber noch keine einzelnen Phoneme besitze (ebd.). Dies ist wiederum ein Problem der Relation von Zielsprachformen zu den vom Kind produzierten Formen, und geht auch von einem statischen Phonembegriff aus, der der dynamischen Entwicklung des kindlichen Phonemsystems nicht gerecht wird. Viertens wird angenommen, dass die auf der Grundlage der Schablonen vom Kind geäußerten Lautformen die Reaktion auf die für es viel zu schwer zu produzierenden Zielsprachformen darstellen, bei denen sich die Kompliziertheit oft in dem Auftreten unterschiedlicher Vokale oder Konsonanten innerhalb einer Silbe oder eines Wortes über Silbengrenzen hinaus zeige (Vihman, DePaolis, Keren-Portnoy 2009: 173). Damit wird indirekt ein inhärentes (angeborenes) Regelsystem angenommen, mit dem das Kind auf der Grundlage eines (angeborenen) Selektionsmechanismus seine Produktion steuert und kontrolliert. Auch dieses Argument ergibt sich wie die vorherigen aus einem Phonemerwerbsansatz mit einem Schwerpunkt, der auf die Betrachtung des Lernens einzelner, konkreter Zielsprachwörter durch das Kind gerichtet ist. Vihman und Croft (2007: 714) begreifen zwar die Phoneme, über die schließlich der Erwachsene verfügt, als sich aus ihren Positionen in den Schablonen durch Kategorisierungs- und Verallgemeinerungsprozesse ergebend, aber bleiben eine Erklärung schuldig, wie dieser Prozess aussehen könnte. Außerdem bleibt unklar, welches die Eigenschaften der Phoneme sind, über die das sich entwickelnde Kind letztend-

lich verfügt, gerade weil ihre Entstehung aus den Schablonen trotz angeführter allgemeiner Begriffe wie Kategorisierungs- und Verallgemeinerungsprozessen nicht schlüssig ist.

6.2.5. Pierrehumberts Probabilistische Phonemtheorie

Pierrehumbert (2003) basiert ihre sog. Probabilistische Phonologie auf der Grundannahme der sog. Vorkommenstheorie (*exemplar theory*), dass die Kognition eines Menschen auf jedem einzelnen Perzept bzw. jedem durch diesen Menschen angetroffenen Stimulus basiert (Pierrehumbert 2003: 132, Johnson 1997). Dies ist nach ihrer Auffassung u.a. deshalb der Fall, weil jedes akustische Signal, das von einem Sprachteilnehmer gehört werde, entlang der für das Hören wichtigen Merkmalsdimensionen im Gehirn des Sprachteilnehmers in einem von Pierrehumbert (2003) phonetisch-parametrische Karte genannten Bereich gespeichert werde. Da Stimuli einer Sprache entlang dieser kontinuierlichen Merkmalsdimensionen veränderlich seien, gebe es eine Streuung der abgebildeten Stimuli entlang der Dimensionen auf dieser Karte (ebd.). Für die Speicherung der Stimuli im Lexikon sei eine phonetische Kodierung notwendig (Pierrehumbert 2003: 116), deren phonetische Kodierungskategorien sich jeweils als Häufigkeitsverteilungen auf dieser Karte ergäben (2003: 132), womit Pierrehumbert (ebd.) nach ihrer Aussage das Kategorienkonzept nach (Luce 1963, Luce und Galanter 1963, beide in Pierrehumbert 2003b: 119) verwendet (Pierrehumbert 2003b: 132). Die Stärke einer phonetischen Kodierungskategorie hängt somit direkt von der Anzahl der angetroffenen Stimuli ab, die Merkmalsausprägungen aufweisen, die in den Bereich der Häufigkeitsverteilungen dieser Kategorie fallen und die Zeit seit ihrem Antreffen berücksichtigen (Pierrehumbert 2003: 132). Neue Stimuli werden der Häufigkeitsverteilung und damit der Kategorie zugeordnet, die ihnen die größte Wahrscheinlichkeit zuweist (Pierrehumbert 2003: 132). Um diese Kategorienzuordnung neu hereinkommender Stimuli plausibel erscheinen zu lassen, vergleicht sie (Pierrehumbert 2003: 132) diese mit der neurolinguistischen Situation, dass jeder neue Stimulus bereits gespeicherte Abbilder von Stimuli in seiner Nachbarschaft auf der Karte aktiviert (ebd.). Sodann würde die Kategorie bzw. die Häufigkeitsverteilung ausgewählt, die durch stärkste Aktivierung die anderen ausreichend hemmen könnte (ebd.). Dieser Ansatz könne auch erklären, wie sich eine ins Leben gerufene Kategorien durch zunehmend mehr angetroffene Stimuli veränderten und Hörer somit eine zunehmend bessere Schätzung der Zentren phonetischer Kategorien sowie ihrer Randbereiche bekämen (ebd.), was zu Häufigkeitsverteilungen und Veränderlichkeit führen würde, die zunehmend mehr denen Erwachsener entsprächen (Pierrehumbert 2003: 132).

Pierrehumbert (2003: 132ff.) bezieht ihr Modell der phonetischen Merkmale auch auf die Sprachproduktion, wobei für die Produktion ein Exemplar der Verteilung einer Kategorie (zufällig oder auch durch soziokulturelle oder andere Faktoren beeinflusst) ausgewählt wird. Anschließend würden die Exemplare in seiner Nachbarschaft aktiviert und die durchschnittlichen Eigenschaften dieser Nachbarschaft schließlich produziert. Könne nun etwas auf diese Weise Produziertes durch den Hörenden (z.B. die Sprachgemeinschaft) nicht adäquat verstanden werden (ebd.: 133), so werde beim nächsten Mal ein anderes Exemplar gewählt, was schließlich auch zum Verschwinden einer Kategorie führen könne, weil diese hinsichtlich der Perzeption oder der Produktion unvorteilhaft sei (ebd.: 133). Diesen Mechanismus der Kopplung von Perzeption und Produktion nennt Pierrehumbert (2003: 133) auch "perception-production loop" bzw. "community feedback" (ebd.). Pierrehumbert (2003: 120ff.) sieht die detaillierten phonetischen Unterschiede zwischen den Produktionen von durch das IPA-Alphabet als identisch vorgegebenen Phonemen, die bei der objektiven instrumentellen akustischen Messung erkennbar werden, als Hauptgrund an, die Universalität der Phoneme abzulehnen (Pierrehumbert 2003b: 127). Sie lehnt aber nicht nur die Universalität des Phonems als solche ab, sondern auch seine über Kontexte und unterschiedliche Wortpositionen hinweg als gegeben angenommen Existenz (Pierrehumbert 2003b: 129), d.h. die Zusammenfassung von positionsbezogenen Varianten, die sie dadurch begründet, dass die Unterscheidung von zwei Phonemen wie z.B. /s/-/z/ hinsichtlich ihrer phonetisch-akustischen Merkmale wie z.B. der VOT (*voice onset times*) in Form einer Häufigkeitsverteilung zwar in konkreten Positionen möglich sei, aber nicht für mehrere Positionen gleichzeitig (Pierrehumbert 2003b: 129), weil dadurch zu stark überlappende Häufigkeitsverteilungen für die zur Trennung verwendeten phonetischen Merkmale entstehen würden, die damit nicht erlernbar seien (Pierrehumbert 2003: 129). Denn Pierrehumbert (2003b: 127ff.) nimmt ausdrücklich an, dass sich entlang einer oder mehrerer kontinuierlicher Dimensionen im parametrischen Raum, in den die Stimuli abgebildet werden, in ausreichendem Abstand die Stimuli an unterschiedlichen Stellen häufen, was zu bi- oder multimodalen Verteilungen führt, wobei die am häufigsten vorkommenden Wertebereiche auf die in der Zielsprache vorkommenden positionsbezogenen Phoneme hinweisen würden. Pierrehumbert (2003b: 116, 138ff.) nimmt neben der parametrischen Ebene und der sich aus dieser, wie zuvor beschrieben, durch Kategorien als Häufigkeitsverteilungen ergebenden phonetischen Kategorien, die lexikalische Ebene an, die mittels der so gebildeten phonetischen Kategorien kodiert bzw. gespeichert wird (Pierrehumbert 2003b: 116). Es gäbe auch Phänomene des

Sprachverhaltens, die auf eine Rückkoppelung des Lexikons auf die Bildung der durch sie kodierten phonetischen (Kodierungs-)Kategorien schließen ließen (Pierrehumbert 2003: 138), weshalb Pierrehumbert (ebd.) eine Beeinflussung der Bildung der phonetischen (Kodierungs-)Kategorien durch Regelmäßigkeiten im Lexikon annimmt. Diese Regelmäßigkeiten seien in der dem Lexikon höher gelagerten Ebene der sog. "phonologischen Grammatik" gespeichert (ebd.). Als derartige lexikonbedingte Phänomene sieht Pierrehumbert (2003b: 138ff.) beispielsweise die Nicht-Wahrnehmung von phonetischen Merkmalen durch Hörer an, die in deren Sprache nicht vorkommen (Pierrehumbert 2003: 138). Dies sei z.B. für die tiefen Werte des dritten Formanten im vokalisiertem /r/ des Englischen durch Japanische Sprecher bzw. Hörer (Pierrehumbert 2003: 138) der Fall. Ein anderes Beispiel sei die Wahrnehmung von nicht-nativen oder sehr selten vorkommenden C_1C_2 -Konsonantenkomplexen von Sprechern entweder als ihrer Sprache ähnlichen C_1C_2 -Konsonantenkomplexen oder aber als durch eine Wortgrenze getrennt, die aber in den präsentierten Komplexen nicht präsent sei (Pierrehumbert, Beckman in press in Pierrehumbert 2003b: 139). Positiv ist an Pierrehumberts (2003) Ansatz, dass sie das gesamte akustische Signal beachtet, wie es dem Hörer vorliegt. Ihre Erklärung der Perzeption könnte die Beobachtungen begründen, warum Kinder im ersten Jahr immer stärker für native Laute empfänglich sind, aber ihre Empfindlichkeit für nicht in ihrer Muttersprache vorkommende Laute nachlässt. Ihre Deutung könnte plausibel sein, weil fremdsprachliche Laute nicht im akustischen Input vorliegen und somit keine Häufigkeitsverteilungen und damit phonetische Kategorien bilden. Die Vorstellung der Kategorienbildung ist aber insgesamt zu einfach konzipiert, weil eine Veränderung von Kategorien nur durch neu wahrgenommene Signale oder die Rückmeldung über die Perzeptions-Produktionsschleife möglich ist, wobei die letztere eine einfache Abgleichungsprozedur darstellt, ähnlich dem Ansatz von deBoer (2001). Wie das Lexikon an der Bildung der phonetischen Kategorien beteiligt sein soll, bleibt zudem im Unklaren.

6.2.6. DeBoers Vokalevolutionssimulation

DeBoer (2001) möchte in einer Computersimulation zeigen, dass aus der lokalen Interaktion einer geschlossenen Gruppe von Sprachteilnehmern, die paarweise miteinander nach bestimmten Regeln interagieren, ein für diese Gruppe gemeinsames Vokalsystem entstehen kann und dass dasselbe universale Eigenschaften von Vokalsystemen aufweist (deBoer 2001: 55). Auf den letzten Punkt soll hier nicht weiter eingegangen werden, sondern lediglich die Annahmen der Computersimulation genauer beschrieben und diskutiert

werden: Für jeden Sprachteilnehmer der Gruppe wird ein Vokalraum für die Werte des ersten und zweiten Formanten dargestellt, der anfänglich unbesetzt ist, und zudem besitzt jeder Sprachteilnehmer einen Perzeptions- und Produktionsmechanismus, mit denen er jeweils die von anderen Teilnehmern synthetisierten Vokale wahrnehmen kann bzw. selbst synthetisieren kann (deBoer 2001). Die Computersimulation besteht nun darin, dass jeweils in einem Simulationsschritt von je zwei Agenten einer einen Vokal "artikuliert" bzw. synthetisiert und damit als Initiator bezeichnet wird, während der andere ihn imitiert und damit der Imitierende ist (deBoer 2001). Dieses Zusammentreffen zweier Agenten nennt deBoer (2001) auch Imitationsspiel. Das Hauptziel des Imitationsspiels formuliert deBoer (2001: 52) wie folgt:

The goal of the agents is to imitate each other as well as possible with a repertoire of sounds that is as large as possible. For this they need to develop repertoires of sounds that are similar to the ones of the other agents. (deBoer 2001: 52).

Es geht also um eine möglichst genaue Imitation des Initiators durch den Imitator und entsprechend sind auch die Regeln für die Interaktion der Sprachteilnehmer miteinander konzipiert: Der Initiator erzeugt einen seiner Vokale, sofern sein Vokalraum bereits einen Vokal enthält, andernfalls platziert er einen neuen. Der Imitator bestimmt den Vokal seines Repertoires, der diesem gehörten am nächsten liegt und produziert denselben (deBoer 2001). Der Initiator signalisiert sodann eine erfolgreiche Imitation, auch kurz *Erfolg* genannt, wenn dieser vom Imitator produzierte und nun von ihm gehörte Vokal seinem ursprünglich produzierten von all seinen Vokalen am nächsten liegt, andernfalls wird ein Misserfolg signalisiert (deBoer 2001). Auch wenn die Regeln dieses Imitationsspiels sich auf die lokale Interaktion durch die Imitation der Sprachteilnehmer beziehen, so entsprechen sie zusammengenommen doch einer gewissen Ab- bzw. Angleichungsprozedur für die Vokalräume der simulierten Sprachteilnehmer. Denn durch sie werden bestehende Vokale zweier Sprachteilnehmer einander angeglichen, neue und den vom Initiator gehörten ähnliche geschaffen, zu dicht beieinander liegende im Vokalraum zusammengefasst und zu wenig erfolgreiche gelöscht. Durch diesen Abgleich einzelner Vokale der Vokalinventare immer jeweils zweier Sprachteilnehmer einer Simulation konvergieren letztendlich die Vokalinventare aller Sprachteilnehmer hin zu einem gemeinsamen Vokalinventar. Der Hauptmechanismus der Evolution eines gemeinsamen Vokalinventars einer Sprechergemeinschaft wird von deBoer (2001: 52) also als die möglichst genaue Imi-

tation eines Sprechers durch einen Hörer angenommen. deBoer (2001) verweist zwar an einer Stelle darauf, dass der Lernmechanismus seiner Simulation unrealistisch sei (deBoer 2001), scheint dies aber nicht auf die Angleichung der Vokale von zwei Sprechern zu beziehen, denn er versucht dies mit Hinweis auf den Spracherwerb plausibel erscheinen zu lassen: "Changing pronunciation in order to match others more closely is also necessary for children learning a language." (deBoer 2001: 52). Die möglichst exakte Imitation für die Entwicklung von Lautsystemen anzunehmen, missachtet aber, dass jeder Sprecher während des Erwerbs über die Silbenschnittmengen sein eigenes Silben- und das damit verbundene Phonemsystem aufbaut, das durch den Druck des kausalen Motors erweitert wird. Bei diesem Vorgang steht nicht primär im Vordergrund, möglichst in Übereinstimmung mit dem Gehörten zu sprechen, sondern aus dem Gehörten das für das eigene System notwendige phonetische Lautmaterial zu entlehnen. Dies zeigt sich gerade beim Kind, das sprechen lernt, nicht zuletzt in den Beschreibungen und Klassifikationen der "Fehler", die sich als Abweichungen der Kinderlautformen zu den angenommenen Zielsprachlautformen ergeben. Das Kind muss anstatt in der Rolle des Imitators vielmehr als produktiver Sprachgestalter gesehen werden, welches sein eigenes System unbewusst bildet bzw. schafft, indem es über seinen kausalen Motor den Bedarf an neuen, die zunehmende Anzahl an Bedeutungen unterscheidenden Lautformen erfordert.

Alle Modelle, die den Spracherwerb durch Imitation zu erklären versuchen, können grundsätzlich die Sprachentstehung oder den Sprachwandel nicht verständlich machen, da sie sich im Rahmen eines Zirkelschlusses bewegen. Die Vorstellung einer (unbegrenzten) Weitergabe und möglichst exakten Aufnahme von Informationen zwischen Sprechern einer Sprache lässt nicht einmal eine Idee entstehen, wie in diese Modelle etwas Neues hineingelangen oder in ihnen entstehen sollte. Im Gegenteil müsste geprüft werden, ob durch ungenaue Übertragungen und Unschärfen ein sukzessiver Informationsverlust entstehen müsste, eine Entropie analog zum 2. Thermodynamischen Gesetz oder ein Verlust an Deutlichkeit wie im Kinderspiel „Stille Post“. Im Gegensatz zu der Realität in den Sprachen der Welt und der Sprache von Individuen, die ständig in Veränderung begriffen sind und in denen sich fortlaufend Neues entwickelt, müsste die konsequente Anwendung dieser Zirkelschlussmodelle in Simulationen zu einer tendenziellen Reduzierung der in ihnen übermittelten Daten gegen Null führen.

6.2.7. Artikulatorische Phonologie

In dem Ansatz der sog. Artikulatorischen Phonologie wird die Geste als die grundlegende Einheit der Sprachproduktion und -planung und auch der Phonologie⁶⁰ angesehen. Der Gestenansatz ist als Reaktion auf die Unzulänglichkeiten der Phonologie entworfen worden, die nur lineare, nacheinander aufgeführte Phonemfolgen betrachtet (Chomsky, Halle 1968) und auch wegen der Umständlichkeit von Beschreibungen in der autosegmentalen Phonologie⁶¹ (Goldsmith 1979) (Browman, Goldstein 1986: 219ff.). Der an der Bildung des Konsonanten /g/ beteiligte Zungenrücken hat in der Silbenfolge /igi/ einen anderen räumlich-zeitlichen Bewegungsverlauf als in der Silbenfolge /uqu/ (Gafos, Benos 2006: 915). Dies soll im Begriff der Geste verallgemeinert werden, indem dieselbe als eine durch eine mathematische Bewegungsgleichung⁶² erfasste unendliche Menge von zeitlich-räumlichen Bewegungsbahnen eines Artikulators durch einen als invariant gegebenen Punkt des Mundraums definiert wird (Browman, Goldstein 1986: 224), der auch als "target" (Gafos, Benos 2006: 916) bezeichnet wird. Beispielsweise könnte der Artikulator der Zungenrücken und das "target" der Ort der stärksten Verengung im Mundraum zwischen dem Zungenrücken und dem harten Gaumen sein. Eine Geste muss aber nicht nur die Bewegung eines einzigen Artikulators enthalten. In der Artikulatorischen Phonologie ist das Konzept der "coordinative structures" (Browman, Goldstein 1986: 223) vorgesehen, dass mehrere Artikulatoren für die Erreichung einer bestimmten Verengung im Mundraum koordiniert. Wesentlich ist aber, dass die Ortsangabe bzw. das "target" für eine Geste durch eine oder die Koordination mehrerer Gesten erreicht wird. Das "target" bzw. die invariante Ortsangabe einer Geste ist dabei sowohl durch einen der Verengungsgrade *geschlossen, kritisch, eng, mittel, weit* als auch einen der Verengungsorte *vorgestülpte Lippen, labial, dental, alveolar, postalveolar, palatal, velar, uvular* und *pharyngal* spezifiziert (Studdert-Kennedy 1992: 92). Im Wesentlichen entsprechen diese Werte den Orten und Arten von Konsonanten und der Zungenlage in der Beschreibung der Phoneme im IPA-Alphabet, wie Studdert-Kennedy (1992: 92) selbst anmerkt. Auf Grund der oben

⁶⁰ Die Äußerungen *cammer* und *camber* unterscheiden sich vom Gestenstandpunkt aus nur durch den Zeitpunkt des Verschluss des Velums in Bezug auf den Zeitpunkt der Öffnung des Lippenverschlusses (Browman, Goldstein 1992: 246). Auf Grund solcher Beispiele wird angenommen, dass die phonologischen Kontraste zwischen Wörtern auf die (zeitlichen Relationen von) Gesten zurückgeführt werden können (Browman, Goldstein 1992: 241ff.).

⁶¹ Obwohl die verschiedenen Ebenen der segmentalen Phonologie auch in dem gestischen Ansatz auf Grund der teilweisen Gleichzeitigkeit der Gesten enthalten sind.

⁶² Es handelt sich um eine gewöhnliche Differentialgleichung zweiter Ordnung (Gafos, Benos 2006: 915).

gegebenen Definition einer Geste stellt sie eine zeitlich-räumliche Bewegung dar und ihr ist neben den Werten für das "target" eine zeitliche Wirkungsdauer zugeordnet, die auch "Aktivierungszeitraum" genannt wird (Studdert-Kennedy 1992: 92). In der Artikulatorischen Phonologie sind die Gesten die diskreten Einheiten, die sich vom Phonembegriff aber dadurch unterscheiden, dass sie nicht sukzessive getrennt bzw. linear hintereinander auftreten müssen, sondern dass ihr Auftreten bzw. ihre Aktivierung auch gleichzeitig erfolgen kann. Dies ist im Begriff des Aktivierungsintervalls einer Geste, d.h. der zeitlichen Dauer einer Geste, und der zeitlichen Überschneidung (Fowler, Saltzman 1993: 184) von gestischen Aktivierungsintervallen enthalten, die wiederum als in einer sog. Gestenpartitur organisiert, modelliert werden. Dabei wird der Begriff der Partitur in Analogie zur Musik verwendet, wo sich die Dauer von Tönen auch zeitlich überschneiden kann. Es stellt sich die Frage nach der Herkunft der als grundlegend und angeboren angenommenen Geste (Studdert-Kennedy 1992: 102). Hinsichtlich des für eine Geste invariant angegebenen Verengungsortes und der ihn spezifizierenden Werte vermutet Studdert-Kennedy (1992: 102), dass sie sich evolutionär entwickelt haben. Dies würde aber bedeuten, dass sie bereits angeboren sind. Denn der im folgenden von Studdert-Kennedy (1992) skizzierte Phonemerwerbsansatz gibt darauf keine Antwort: Er ist zielsprachbezogen und betrachtet die Äußerungen des Kindes als Fehler bei der Ausführung der Gesten. Im Sinne eines Ansatzes, der die kindlichen Lautformen nach für sie angenommenen Zielsprachformen gruppiert, werden die vom Kind produzierten Äußerungen konsequent als "errors of gestural location and timing required to make the shift from [the child form] to [the adult target form]" (Studdert-Kennedy 1992: 98) analysiert. Zusätzlich zu den Fehlern des falschen Ortes von Gesten (Studdert-Kennedy 1992: 97) und ihrer falschen zeitlichen Anordnung und Ausführung werden auch die Auslassung von Gesten der angenommenen Zielsprachlautform und das durch die Gestenlogik bedingte Auftauchen von ungewollten Gesten als gestische Fehler genannt (Studdert-Kennedy 2001). Es handelt sich also um eine Fehlerklassifikation wie bei Lust (2006: 158) präsentiert, nur in Gestenform. Da im Ansatz der Artikulatorischen Phonologie die Gesten als autonome, unabhängige Einheiten angesehen werden, sind sie und mit ihnen die "target"-Werte für den Ort der stärksten Verengung einer Geste dem eigentlichen Spracherwerb vorgeordnet und damit als angeboren anzusehen. Der Phonemerwerb besteht sodann im Erlernen der richtigen Koordination dieser Gesten: Der Gestische Ansatz "views the task of learning to talk as quite largely one of learning to coordinate gestures that may differ in their articulatory compatibility" (Studdert-Kennedy 1992: 97 in Bezug auf Menn 1983). Und es wird

zudem angenommen, "that segments be defined, superordinate to the gesture, as emergent structures, comprising recurrent, spatiotemporally coordinated, gestural routines." (Studdert-Kennedy 1992: 101). Das Erlernen der richtigen Gestenkoordination für ein Zielsprachwort soll dabei innerhalb der für dieses Zielsprachwort gewählten sog. "artikulatorische(n) Routinen"(Studdert-Kennedy 1992: 98) erfolgen, die vom Kind aus der Plapperphase übernommen worden sein sollen und besonders geübten Bewegungsabläufen entsprechen würden⁶³ (ebd.). Damit wird auch in diesem Ansatz wie bei Piske (2001) ein Großteil der Veränderlichkeit der Kinderlautformen der Unfähigkeit zugerechnet, die Gesten des Zielsprachwortes in angemessener Weise zu koordinieren und evtl. überhaupt zu äußern. Denn ausgehend von der stereotypen Routine, die einem oder mehreren Zielsprachwörtern zugeordnet ist, versucht das Kind, diese durch eine zunehmend adäquater auf Koordination und Produktion der Gesten des Zielsprachwortes gerichtete zu durchbrechen:

While a gestural routine may afford a child initial access to difficult words of similar gestural pattern, it cannot solve all the problems of gestural selection and phasing with which the child must contend in moving toward an acceptable pronunciation. Variability within the constraints of the routine is an important part of this process. (Studdert-Kennedy 1992: 97).

Es geht beim Phonemerwerb aber nicht vorrangig um das besonders präzise Produzieren von zielsprachlichen Einzelwörtern, sondern um den Aufbau, die Unterhaltung und die Weiterentwicklung eines in sich schlüssigen, gewachsenen Silbensystems, dessen Erweiterung durch den Druck zur zunehmenden Unterscheidung von vom Kind entdeckten kausalen Zusammenhängen vorangetrieben wird. Würde man jedoch die Zielsprachbezogenheit und die damit verbundene Fehlerklassifikation wie in der hier vorgestellten Vorgehensweise von Studdert-Kennedy (1992) fallen lassen und sich stattdessen auf die Untersuchung der detailliert gemessenen Bewegungsverläufe und ihres gegenseitigen Verhältnisses im Zeitverlauf beschränken, so wäre es denkbar, dass auch das Systematische der kindlichen Äußerungen besser dargestellt werden könnte. Der angenommene grundlegende und angeborene Status der Gesten zusammen mit denen für sie angegebenen invarianten als "targets" bezeichneten Verengungsorten ist jedoch fragwürdig. Denn es

⁶³ Dabei werden für die analytische Extraktion solcher Routinen unterschiedliche Kinderlautformen durch als äquivalent aufgefasste Konsonanten an gleichem Ort (labial oder velar) von unterschiedlicher Art als durch eine solche Routine erzeugt gruppiert (Studdert-Kennedy 1992: 97). Hier gilt die gleiche Kritik, wie sie bei der Besprechung des Ansatzes von Piske (2001) in Bezug auf die Zusammenfassung unterschiedlicher Kinderlautformen durch seine artikulatorischen Muster geäußert wird.

wird keine Erklärung geboten, wie sich dieselben entwickelt haben sollen. Im Gegensatz dazu stellt die Silbe eine wirklich grundlegende Einheit dar⁶⁴. Im Phonemerwerb werden Silben als Ganzheiten produziert und ihre Erzeugung wird durch Abstände und Dichtezustände geregelt, die die Unterscheidbarkeit der Silben als Signalträger gewährleisten. Die Phoneme sind Größen des Silbensystems, die sich durch Schnittmengenbildung ergeben.

6.2.8. Der Erwerb artikulatorischer Muster

Piske (2001: 91ff.) untersucht die Entwicklung der Wiedergabe der Zielsprachwörter durch ein Kind mittels sog. artikulatorischer Muster, die einzelnen C₁V(C₂)-Silbenklassen oder Kombinationen derselben entsprechen. Dabei wird eine solche C₁V(C₂)-Silbenklasse über die Angabe mehrerer Einzelphonemfolgen oder die Spezifikation von solchen über phonetische Merkmale festgelegt. Zusätzlich wird für (Kombinationen von) C₁V(C₂)-Silbenklassen, die Piske (2001: 91) auch eine "phonetische Struktur" nennt, folgendes gefordert:

- a) Es ist nur dann angemessen, davon auszugehen, daß eine bestimmte phonetische Struktur ein artikulatorisches Muster repräsentiert, wenn die betreffende Struktur wenigstens zeitweise für mindestens zwei Wörter charakteristisch ist.
- b) Für jedes einzelne Wort gilt dabei, daß nur dann davon ausgegangen werden kann, daß seiner Produktion tatsächlich ein bestimmtes artikulatorisches Muster zugrunde liegt, wenn die Mehrzahl der Belege dieses Wortes entweder in einer Aufnahmesitzung oder über einen längeren Zeitraum hinweg die betreffenden phonetischen Charakteristika dieses Musters aufweist. (Piske 2001: 91).

Diese letzten beiden Bedingungen beziehen sich also allein darauf, wie die Lautformen zeitlich und mengenmäßig verteilt sein müssen, die ein bestimmtes Muster aufweisen und als Versuche gewertet werden, ein Zielsprachwort zu produzieren. Piske (2001) versteht unter dem Begriff des artikulatorischen Musters also eine hinsichtlich phonetischer, zeitlicher und zielsprachbezogener Kriterien festgelegte stabile Struktur. Verantwortlich für die Bildung eines artikulatorischen Musters sind seiner (Piske 2001: 217) Meinung nach vom Kind schon früh in der Entwicklung beherrschte artikulatorische Bewegungen (ebd.), auditiv hervorstechende akustische Merkmale der Zielsprachwörter, für die angenommen wird, dass sie auf der Grundlage des artikulatorischen Musters gebildet werden (ebd.) und schließlich als letzten Einflussfaktor für die Bildung solcher Muster visuell

⁶⁴ Siehe dazu auch die oben angeführte Diskussion zur Bedeutung der Silbe im Kapitel 5.

erlangte Erfahrungen mit dem Sprechen Erwachsener, was vor allem die durch die Lippen bzw. Zungenspitze oder Zungenkranz sichtbaren labialen und apikalen oder koronalen Laute begünstigen würde (ebd.). Wie Vihman und Croft (2007, s.o.) geht Piske auch davon aus, dass bestimmte Zielsprachwörter "als eine Art Auslöser für ein Muster [gelten können], da sie erstmals eine Lautstruktur [...] [aufweisen], die später für eine Reihe weiterer Wörter charakteristisch" wird⁶⁵ (Piske 2001: 94), die der Bildung dieses ersten Wortes "phonetisch sehr ähnlich" (Piske 2001: 95) sind und so also "in ein bereits etabliertes Muster integriert" (Piske 2001: 94) werden. Dabei gibt es aber den Unterschied zu dem "adapted"-Begriff von Vihman und Croft (2007) (s.o.), dass hier die Zielsprachwörter nicht größere Abweichungen zum artikulatorischen Muster aufweisen wie die "Auslöser" (ebd.)-Zielsprachwörter. Denn Piske (2001) führt das Konzept der Veränderung von Mustern ein (s.u.), wodurch die Abweichungen zwischen Zielsprachwort und Muster nicht zu groß werden können. Piske (2001: 91ff.) untersucht die Spracherwerbsäußerungen in Form artikulatorischer Muster für zwei Kinder, von denen die Untersuchungen der Sprache des Kindes ISA (Piske 2001: 130ff.) zur Erläuterung des Begriffs des artikulatorischen Musters kurz vorgestellt werden sollen: Piske (2001: 131ff.) macht für das Kind ISA insgesamt zwölf Anfangsmuster aus, die durch die Beschreibung des Silbenkopfes C_1 und des Silbenkerns V einer $C_1V(C_2)$ Silbe durch die phonetischen Merkmale v.a. des Artikulationsortes und der -art für die Konsonanten und der vertikalen und horizontalen Zungenlage für die Vokale erfolgt. Dabei wird angenommen, dass die für C_1 festgelegten Silbenkopfkonstituenten auch in C_2 vorkommen können, aber nicht andere. Für die meisten der angegebenen Muster, bis auf zwei Ausnahmen, ist die Menge der C_1 -Konsonanten eines einzelnen Musters recht begrenzt, während die Silbenkernmenge als zeitlich sich vergrößernd angegeben wird und insgesamt viel größer als die Menge der C_1 eines Musters ist. Deshalb spricht Piske (2001) bei solchen Mustern von den Phonemfolgen der Ränder als "stabile[n] Grundelementen" (Piske 2001: 183) einerseits und den Silbenkernen als "stärker varianten Elementen andererseits" (Piske 2001: 183), während er (Piske 2001) bei Mustern mit weniger unterschiedlichen Silbenkernen als Rändern diese Begriffe umgekehrt verwendet. Also heißt die Konstituente C oder V eines CVC -Musters mit mehr unter ihr subsumierten Phonemen *variabel*, während die andere als *stabil* bezeichnet wird. Die zwölf für das Kind ISA ausgemachten Muster sehen im Einzelnen wie folgt aus: So handelt es sich bei Muster 1 um einen alveolaren oder dentalen Plosiv (Piske

⁶⁵ vgl. den Begriff "selected" bei Vihman und Croft (2007)

2001: 133), bei Muster 2 um einen alveolaren und palatalen Nasal, der in der Silbenkoda velar sein kann (Piske 2001), bei Muster 3 um einen bilabialen Plosiv ("stimmlos, stimmlos affriziert, stimmhaft, stimmhaft entstimmt" (ebd.)) (Piske 2001), bei Muster 4 um den bilabialen Nasal /m/ (Piske 2001), bei Muster 5 um einen glottale Knacklaut (Piske 2001), bei Muster 6 um einen velaren Plosiv ("stimmhaft oder -los, manchmal auch als entsprechende Affrikate" (ebd.)) (Piske 2001), bei Muster 7 um den stimmhaften oder -losen glottalen Frikativ (Piske 2001), bei Muster 8 um einen Konsonanten, der als Verengung zwischen dem "Zungenkranz oder verschiedene[n] Teile[n] des Zungenrückens"(ebd.) und dem alveolaren Bereich des Munddachs gebildet wird (Piske 2001), bei Muster 9 um einen palatalen oder alveolaren lateralen Approximanten (/l, ʎ/) (Piske 2001), bei Muster 10 um eine Spezifizierung von sowohl C₁ als mit bevorzugt dem glottalen Plosiv oder dem entsprechenden stimmhaften oder -losen Frikativ als auch dem Silbenkern V als den Diphthongen /aʊ/ oder /aɔ/ (Piske 2001), beim Muster 11 um einen "dental(isiert)en oder alveolaren Frikativ /θ, ð, ʧ, s, z/ oder Affrikaten" (ebd.) aus diesen Frikativen und einem dentalen oder alveolaren stimmlosen Plosiv und schließlich beim Muster 12 um einen labiodentalen Frikativ ("stimmlos oder -haft, manchmal auch stattdessen die Affrikate /pf/, der stimmlose bilabiale Frikativ /ɸ/ oder der labio-velare Approximant /w/" (ebd.)). Durch diese Anfangsmuster sind eine Reihe von CV(C) Silben bestimmt, von denen die meisten hinsichtlich der Menge der C₁-Konsonanten, wie eben angegeben, recht eng begrenzt sind, während für die Silbenkerne bzw. Vokale V sehr viele möglich sind. Im Grunde genommen handelt es sich also bei diesen zwölf für den Anfang des Spracherwerbs bei ISA ausgemachten artikulatorischen Mustern um eine Klassifizierung der Silben in zwölf durch phonetische Kriterien des Kopfkonsonanten festgelegte Klassen, bei denen die (phonetische) Identität der Vokale nicht zur Einteilung verwendet wird. Die von Piske (2001) ausgemachten Anfangsmuster von ISA haben also alle eine einsilbige CV-Struktur. Anstatt wie Vihman und Croft (2007) nur zwischen einer an das Muster gut angepassten Zielsprachform ("selected") und einer schlecht angepassten ("adapted") zu unterscheiden, verfolgt Piske (2001) die Weiterentwicklung von artikulatorischen Mustern auf die folgende Weise: Die Lautformen des Kindes, die nicht in eine dieser als artikulatorische Anfangsmuster phonetisch-zeitlich-quantitativ bestimmten CV-Silbenklassen fallen, werden durch weitere Operationen der Muster beschrieben: Dies sind die Kombinationen dieser Anfangsmuster wie z.B. "Muster 1 + Muster 6" (Piske 2001: 173ff.), die Erweiterung der Menge der für ein CV-Muster zulässigen Sil-

benkerne V, auch "interne Erweiterung" (Piske 2001: 167) genannt, das Anfügen eines der Kopfkonzonanten der zwölf Anfangsmuster an eines dieser Anfangsmuster CV, also die Bildung einer C_1VC_2 -Silbenstruktur, auch als "Ergänzung von Mustern um einzelne Komponenten anderer Muster" (Piske 2001: 185) bezeichnet, und schließlich als letzte Konstruktionsart neuer Muster das Anhängen von neuen Konsonantenfolgen an solche der Anfangsmuster, auch "die Ergänzung von Mustern um neue Komponenten" (Piske 2001: 193) genannt.

Die Kritik dieses Ansatzes richtet sich auf folgende Punkte: Erstens ist es generell problematisch, Lautformen des Kindes konkreten Zielsprachlautformen zuzuordnen: Denn es gibt zwischen Bedeutungen verschiedener Sprecher nur Schnittmengen, insbesondere zwischen den Bedeutungen eines Kindes und eines erwachsenen Beobachters (vgl. Abschnitte 2.1 u. 4.7.). Zweitens engt die Zuordnung von kindlichen Lautformen zu einer Zielsprachlautform die Erklärung und Beschreibung des Phonemerwerbs auf die Frage ein, warum und wie das Kind bestimmte Lautformen *für* eine bestimmte Zielsprachlautform äußert (vgl. Abschnitt 2.1.). Da dies oft auf die Weise erfolgt, dass man Verallgemeinerungen über die einer Zielsprachlautform zugeordneten Kinderlautformen macht, stellen einige der Kinderlautformen Ausnahmen hinsichtlich dieser Verallgemeinerung dar. Bei Piske (2001) wird für diese Verallgemeinerungen, die artikulatorischen Muster, per Definition gefordert, dass "die Mehrzahl der Belege dieses Wortes entweder in einer Aufnahmesitzung oder über einen längeren Zeitraum hinweg die betreffenden phonetischen Charakteristika dieses Musters aufweist" (Piske 2001: 91). Die Ausnahmen erklärt Piske (2001: 108) durch "noch nicht voll ausgereifte neuromotorische Kontrollfähigkeiten" (2001: 108):

"Festzuhalten bleibt insgesamt, daß die Muster LEOs nicht vollkommen stereotyp sind, sondern daß sie ein begrenztes Ausmaß an Variabilität sowohl im Hinblick auf die Silbenstrukturen früher Wörter als auch im Hinblick auf die Realisierung einzelner konsonantischer und vokalischer Segmente zulassen. Mit einem gewissen Grad an phonetischer Variation muß [...] in den frühen Stadien des Laut- und Lexikerwerbs besonders aufgrund der noch nicht voll ausgereiften neuromotorischen Kontrollfähigkeiten kleiner Kinder ohnehin auch gerechnet werden." (ebd.).

Drittens gehen aber nicht nur Lautformen des Kindes bei der Untersuchung verloren, die ein und demselben Zielsprachwort zugeordnet sind, sondern auch solche, die Silben enthalten, die in keiner der als artikulatorische Muster bezeichneten Silbenklassen vorkommen, weil die phonetischen Merkmale derselben nur in den Lautformen für eine Ziel-

sprachlautform auftreten und somit wegen der obigen Musterdefinition kein artikulatorisches Muster sein können. Viertens geht nicht nur wegen der Gruppierung von Lautformen des Kindes zu einem Zielsprachwort Information verloren, sondern auch dadurch, dass die Konstituenten einer als Muster oder Teil desselben angenommenen C₁VC₂-Klasse mehr als ein Phonem enthalten können. Diese Gruppierung der Konstituenten erfolgt in den meisten Fällen nach phonetischen Merkmalen wie labial, dental/alveolar etc. und somit ließe sich dieselbe dadurch rechtfertigen, dass die konkreten, einzelnen zusammengefassten Phoneme sich phonetisch ähnlich sind, weil sie beispielsweise alle mit herabgelassenem Velum, alle mit dem Zungenkranz an der Rückseite der oberen Schneidezähne oder demselben am Zahndamm etc. gebildet werden. Aber es handelt sich dennoch um eine Zusammenfassung von vom Transkribierenden gehörten Einzelphonemen, womit ihre Identität innerhalb der so gebildeten Muster verloren geht. Fasst man wie in diesem Ansatz der artikulatorischen Muster einzelne Silbenköpfe und -kotas durch Aufzählung oder mittels phonetischer Merkmale zusammen, so ist die Frage, wie man die Entstehung der konkreten, einzelnen Konstituenten erklären bzw. beschreiben möchte. Zudem räumt Piske (2001: 98ff., 151) selbst ein, dass diese phonetischen Kategorisierungen von Köpfen und Kotas in gewissem Grade willkürlich sind:

Insgesamt haben beide Vorgehensweisen, d.h. Zuordnen von Wörtern, die sich durch auffällige Gemeinsamkeiten auszeichnen, zu unterschiedlichen Mustern oder ihre Zusammenfassung unter einem Muster, auf jeden Fall sowohl jeweils Vor- als auch Nachteile. Werden solche Wörter sofort unter einem Muster zusammengefaßt, werden bestimmte Unterschiede in ihrer Ausspracheentwicklung unter Umständen nicht deutlich. Werden sie dagegen unterschiedlichen Mustern zugeordnet kann der Eindruck erweckt werden, als seien gemeinsame Charakteristika unbedeutend. Diese Anmerkungen werfen natürlich insgesamt die Frage auf, wie sehr sich eine phonetische Struktur von einer anderen unterscheiden muß, so daß es berechtigt ist, davon auszugehen, daß sie beide unterschiedliche Muster und nicht ein gemeinsames Muster repräsentieren. Die in dieser Arbeit analysierten Daten deuten darauf hin, daß in bezug auf diese Frage nicht immer eine eindeutige Entscheidung getroffen werden kann. (Piske 2001: 98f.)

Piske (ebd.) spricht hiermit zwar nur das Problem der Zusammenfassung von Silbenköpfen (und -kotas) bei der Musterbildung an und macht auf den damit möglicherweise einhergehenden Verlust von Informationen hinsichtlich der einzelnen Konstituenten aufmerksam. Aber der gleiche Einwand gilt ebenso für die Silbenkerne, die bei diesem Vorgehen eher als große Gruppe (zwar mit zeitlicher Unterteilung) aufgefasst werden, die den Kopfkonzonantengruppen eines Musters zugeordnet wird. Wirklich gelöst wird die-

ses Problem nur, wenn man auf die Zusammenfassung einzelner konkreter Konstituenten zu Mustern verzichtet und ihre Entwicklung einzeln für sich, aber im Verhältnis zu den anderen betrachtet, wie es in dieser Arbeit getan wird. Fünftens sieht Piske (2001) den Hauptantrieb hinter der Bildung neuer Lautformen durch das Kind darin, sich immer mehr den zielsprachlichen Lautformen anzunähern und diese immer präziser wiederzugeben:

Allerdings konkurriert die phonetische Stabilität, die durch den Rückgriff auf etablierte Produktionsmuster offenbar gewährleistet wird, auch mit der Notwendigkeit, daß ein Wort irgendwann einmal auch konsistent zielgerecht wiedergegeben werden muß. Folglich müssen die frühen artikulatorischen Muster im Laufe der Zeit aufgebrochen bzw. weiterentwickelt werden, so daß allmählich auch strukturell komplexere Wörter zielgerecht wiedergegeben werden können. (Piske 2001: 109).

Die vom Kind in seinen Lautformen enthaltenen Silben konstituieren jedoch zusammen ein Silbensystem, das unter dem Druck, für zunehmend mehr durch das Kind gebildete Bedeutungen, Bezeichnungen zu bilden, erweitert bzw. vergrößert wird. Es mag zwar auch Situationen geben, in denen Lautformen durch Versuche des Kindes entstehen, sich einer gehörten zielsprachlichen Lautform anzunähern. Doch maßgeblich und hauptseitig entscheidend für die Entstehung bzw. Bildung neuer Silben ist der wachsende Bedarf des Kindes die von ihm in Bedeutungen extrahierten kausalen Zusammenhänge voneinander unterscheiden zu können. Sechstens untersucht bzw. beschreibt Piske (2001) zwar die Veränderung bzw. weitere Entwicklung der von ihm postulierten Muster, es ist aber nicht klar, welchen inneren Notwendigkeiten bzw. Regelmäßigkeiten diese Musterveränderung folgt. Damit verbleibt sie im phänomenologischen und spekulativen Bereich. Zusammenfassend ist also kritisch an diesem Vorgehen zu sehen, dass der Begriff des Musters an sich das Häufige beinhaltet, was dazu führt, dass es Lautformen der Kinder gibt, die als Ausnahmen behandelt werden müssen. Außerdem ist nicht klar, warum die einzelnen Weiterentwicklungen der Muster auftreten und warum es einen Wechsel von Mustern für ein Zielsprachwort gibt, nicht zuletzt wegen des direkten Bezugs der Musterdefinition auf jeweils einzelne Zielsprachwörter und damit eine Gruppierung der kindlichen Lautformen nach Bedeutungen (vgl. Abschnitt 2.1.). Damit wird eine falsche Akzentuierung der kindlichen Lautformen vorgenommen, die als Ausdruck des kindlichen Systems begriffen werden sollten.

7. Untersuchungen von Konstituentenverteilungen in Korpora

7.1. Anforderungen an die Beschaffenheit von Korpora zur Untersuchung des Phonemerwerbs

Ein Korpus für die Untersuchung des Phonemerwerbs muss eine Reihe von Anforderungen erfüllen, um dem dynamischen Wesen des wachsenden Silbensystems gerecht zu werden. In diesem Zusammenhang soll die Problematik der Phonemtranskriptionen nicht weiter diskutiert werden. Wie bereits oben erwähnt, werden diese als nützliche Hinweise auf die im Zeitverlauf vorliegenden Zustände des wachsenden Silbensystems angesehen. Gegenüber rein phonetischen Messungen haben sie den Vorteil, dass sie von einem Sprecher einer Sprache gehört worden sind und nicht etwa von Maschinen ermittelt wurden. Und da sie geistige Größen sind, die sich zwar auf phonetischer Grundlage, abhängig vom organischen Silbenwachstum entwickeln, kann auf sie auch nicht verzichtet werden. Sie sollten vielmehr durch gezielte phonetische Messungen ergänzt werden. Eine erste wichtige Anforderung an Phonemtranskriptionen der Äußerungen eines Kindes ist ihre Durchführung durch eine Person, da bei einer Aufnahme unterschiedlicher Phasen des Phonemerwerbs durch mehrere Transkribierende ihre Konsistenz nicht mehr gewährleistet wäre. Diese Annahme beruht auf dem oben präsentierten PAM-Modell von Best (1995) in dem das allgemeine Verhalten von Sprechern gekennzeichnet wird, Laute einer fremden Sprache entweder als nicht-sprachliche wahrzunehmen oder bezogen auf ihr eigenes Phonemsystem zu assimilieren. Die Inkonsistenz der Transkriptionen lässt sich nicht nur über Sprachgrenzen hinweg, sondern auch bei Sprechern einer Sprache nachweisen, wie Port und Leary (2005: 938, 941f.) betonen. Eine weitere Anforderung an ein Korpus im Hinblick auf den Phonemerwerb diskutieren Jakobson und Waugh (1986: 168f.):

Durch ihre reichhaltigen Daten über Sprache und Umgebung und durch ihre präzisen und expliziten Beweise übertreffen diese Studien [in Tagebuchform u.a. von Leopold (1947)] die immer häufiger werdenden 'Schnappschüsse' des Sprachverhaltens des Kindes und seines linguistischen Repertoires in einem gegebenen Augenblick. Solche 'Schnappschüsse' werden gemacht, indem man das Kind während einiger Monate in wöchentlichen oder längeren Abständen für eine halbe 'Stunde besucht und währenddessen Tonbandaufnahmen macht; manche Forscher beschränken sogar ihre Aufnahme des Kleinkindes auf ein einziges Treffen von etwa 100 Minuten. Diese seltenen und oberflächlichen Treffen zwischen dem Kind und einer ihm unbekanntem Person zerstückeln, mechanisieren und entstellen das dynamische Gesamtbild, da die Triebfedern der Veränderungen und ihre Reihenfolge verborgen bleiben und die in-

nersten dynamischen Gesetze des allmählichen Aufbaus nicht entdeckt werden können (Jakobson, Waugh 1986: 168f.)

Jakobson und Waugh (1986) beziehen sich auf Tagebuchstudien wie etwa die von Leopold (1947) und stellen diesen Studien andere Untersuchungen gegenüber, bei denen das Kind nicht zeitlich kontinuierlich beobachtet wird, sondern die Transkribierenden das Kind lediglich in zeitlichen Abständen mit dazwischenliegenden längeren Pausen für Aufzeichnungen aufsuchen. Dadurch gehen aber mögliche Äußerungen und damit Silben der Zeiträume verloren, in denen das Kind nicht beobachtet wird, und damit Informationen über das sich entwickelnde kindliche Silbensystem. Ein Phonemerwerkskorpus sollte also idealerweise so zeitlich detailliert sein, dass alle vom Kind geäußerten unterschiedlichen Silben aufgezeichnet werden. Jakobson und Waugh (1986) beziehen sich in ihren Überlegungen u.a. auf Tagebuchaufzeichnungen von Leopold (1947), der jedoch im Vergleich mit dem von Elsen (1991) nicht ein ähnliches Maß an Detailliertheit und Fülle der Information aufweist. Da im Ansatz dieser Arbeit der Phonemerwerb als vorwiegend autonomer Prozess einer kontinuierlichen Systementwicklung mit einer rezeptiv gesteuerten Orientierung auf die Zielsprache gesehen wird, sollten für die primäre Untersuchung dieses Prozesses keine Korpora verwendet werden, die die Daten mehrerer Kinder mischen wie es z. B. bei Querschnittsuntersuchungen der Fall ist. Diese können zwar Hinweise auf Gemeinsamkeiten geben, lassen aber keine Schlüsse auf die Individualität der Entwicklung zu. Die erwähnte zeitliche Detailliertheit der Transkriptionen der kindlichen Äußerungen sollte auch eine qualitativ detaillierte Korpusbildung einschließen: Es sollten keine Lautformen als unbedeutend oder störend aussortiert werden, soweit sie eine eigene Lautidentität aufweisen. Dies ist aber bei vielen Ansätzen der Fall, die die kindlichen Lautformen nach Zielsprachlautformen gruppieren. So sagt Lindner (2009:163) beispielsweise:

Im Gegensatz zu anderen ähnlich schlicht strukturierten Zielwörtern (z.B. *ja, ei, na*), zeichnet sich *da* bei allen drei Kindern dadurch aus, dass es überproportional häufig belegt ist: SIM 1926 mal, ISA 376 mal, TL 442 mal. Um aufgrund dieser hohen Zahlen eine evtl. Ergebnisverzerrung zu vermeiden, wurde diese Vokabel bei der Datenanalyse nicht berücksichtigt. Folglich wird *da* nicht in der Kardinaltabelle geführt, die bei allen drei Kindern mit der laufenden Wortnummer 2 beginnt. (Lindner 2009: 163)

Mit dieser Begründung wird eine Vielzahl von Formen des späteren Silbenkorpus ausgeschlossen, für das Korpus des Kindes "SIM"(ebd.) 1926 Formen, die für die weitere Un-

tersuchung nicht mehr zur Verfügung stehen. Die Reduzierung von Lautformenmaterial in den entsprechenden Zeitabschnitten engt die Sicht auf die Silbenentwicklung ein und führt zu quantitativen Ungenauigkeiten. Außer der Notwendigkeit, nach Möglichkeit alle qualitativ unterschiedlichen Äußerungen des Kindes zu notieren, ist auch nach Möglichkeit die absolute Ordnung der geäußerten Lautformen und damit der Silben zu notieren, weil dem Phonemerwerbsmodell die Annahme zu Grunde liegt, dass die zeitlich früher geäußerten Silben in einem inneren Zusammenhang mit den später artikulierten stehen und dass sie in gewisser Hinsicht deren Voraussetzung sind. Denn jede Silbe verändert im Prinzip das Gleichgewicht im gesamten Silbensystem. Oft liegen leider die Äußerungen nur in zeitlich ungeordneten Mengen vor, weshalb den Silben keine absolute zeitliche Ordnung gegeben werden kann.

7.2. Beschreibung der verwendeten Korpora und Aufbereitung der Daten

7.2.1. Das Annalena-Korpus (Elsen 1991)

Das in Elsen (1999: 135ff.) angegebene und für die maschinelle Verarbeitung digitalisierte Annalena-Korpus beginnt mit der Äußerung des ersten Wortes des Kindes Annalena im Alter von 0;8,23 und endet Anfang 2;5, weil zu diesem Zeitpunkt der Phonemerwerb als abgeschlossen angesehen wurde (ebd.: 58ff.). Das Korpus umfasst 12117 unterschiedliche von dem Kind Annalena geäußerte Lautformen, die an 635 Tagen, auch *Sitzungen* genannt, aufgezeichnet wurden. Die mehrsilbigen Lautformen werden nach folgenden Kriterien in Einzelsilben zerlegt: Ist zwischen zwei Phonemen ein Silbenschnittdiakritikum angeführt, so erfolgt an dieser Stelle die Trennung. Andernfalls werden Einzelkonsonanten beiden benachbarten Silbenkernen zugeordnet⁶⁶ und Komplexe hingegen für die kindlichen Lautformen einzeln geprüft und nach intuitivem Sprachgefühl zerlegt.

⁶⁶ Bei der Silbifizierung des Annalena-Korpus gilt: Die Zuordnung von einzelnen Konsonanten C zwischen zwei direkt von beiden Seiten anschließenden Nuklei K_1 , K_2 , d.h. $\dots K_1 C K_2 \dots$ als Teil von mehrsilbigen Lautformen zu beiden Nuklei soll hier nicht als Annahme verstanden werden, dass alle zwischen zwei Vokalen V_1 und V_2 stehenden einfachen Konsonanten C ambisilbischen Status haben, wie dies im Deutschen durch den vorhergehenden Vokal V_1 auf die folgende Weise determiniert ist:

Das Vorkommen ambisilbischer Konsonanten wird durch die lautliche Umgebung restringiert [...]: nach ungespannten Kurzvokalen (z.B. *offen*, *Roggen*) werden Konsonanten mit zwei Silben assoziiert, nach Langvokalen bilden sie den onset der folgenden Silbe (z.B. *Ofen*, *Rogen*). (Lindner 2009: 87).

Die für die Überprüfung des Vorliegens eines ambisilbischen Konsonanten notwendigen Kriterien (vgl. u.a. Lindner 2009: 86; Wiese 1988: 78) sind unter diesen Voraussetzungen irrelevant, da durch die Weglassung der Diakritika auch nicht die Vokallängen beachtet werden. Diese bestimmen aber gerade mit (s.o.), ob es sich bei C um einen ambisilbischen Konsonanten handelt oder nicht (ebd.). Bei der Zuordnung eines Konsonanten C, der sich als einziger zwischen zwei Silbenkernen K_1 , K_2 in einer mehrsilbigen Lautform befindet, die $\dots K_1 C K_2 \dots$ als Lautformenbestandteil enthält, stand vielmehr der Wunsch im Vordergrund

Bei der Extraktion der Silben wurden keine phonemmodifizierenden Diakritika beachtet, um die zu untersuchenden Laute in überschaubaren Grenzen zu halten und somit eine Vereinfachung des Modells zu erreichen. Da Diakritika Phoneme modifizieren, die aber über einen gemeinsamen Kernbereich von Eigenschaften verfügen, wurde der Verzicht auf die auf Diakritika zurückgreifende Transkription für vertretbar gehalten. Durch diese Silbifizierung ergeben sich unter Verzicht auf die Diakritikamodifikationen 7190 unterschiedliche Silben, in denen wiederum 319 Silbenköpfe, 327 Kodas, 206 Silbenkerne, 49 Konsonanten und 42 Vokale enthalten sind.

7.2.2. Das ISA-Korpus (Lindner 2009)

Das ISA-Korpus ist der in Lindner (2009: 455ff.) enthaltenen sog. "Kardinaltabelle" (ebd.) entnommen und digitalisiert worden. Er enthält Aufzeichnungen der Äußerungen eines Deutsch lernenden Kindes namens "ISA", das während des Beobachtungszeitraums zwischen 1;0;30 und 2;00;15 alt war (Lindner 2009: 157). Im Gegensatz zum Annalena-Korpus liegen dem Korpus keine täglich stattfindenden Tagebuchbeobachtungen zu Grunde, das Kind wurde nur einmal wöchentlich besucht (Lindner 2009: 157). Es fanden insgesamt 37 Besuche bzw. Sitzungen statt (ebd.). Das ISA-Korpus (Lindner 2009) umfasst 3004 unterschiedliche Lautformen. "Die Silbengrenzen [...] sind interpretativ auf Grund des intuitiven Verständnisses des Deutschen angesetzt." (Lindner 2009: 161). Die über diese Silbengrenzen festgelegten Silben wurden extrahiert und ergaben insgesamt

durch die Zuordnung von C zu beiden Kernen K_1 und K_2 auf jeden Fall nicht durch eine falsche Zuordnung Silben $\dots K_1 C$ oder $C K_2 \dots$ wegzulassen.

Mit dieser Entscheidung wurden systematische Abweichungen von der Zielsprache in Kauf genommen: So zeigen beispielsweise die Abbildungen 83 und 85 sehr viele stimmhafte Plosive in der Silbenkoda, die im Deutschen in der Koda einer Silbe auf Grund der sog. Auslautverhärtung nicht vorkommen können (u.a. Lindner 2009: 41), die als sog. phonologischer Prozess formuliert wird (ebd.), die auf sog. zugrundeliegenden Repräsentationen operiert und dabei stimmhafte Obstruenten und damit auch stimmhafte Plosive entstimmt (ebd.).

Zudem soll darauf aufmerksam gemacht werden, dass eine für die Lautformen des Kindes durchgeführte Silbeneinteilung nicht unbedingt auf Regelmäßigkeiten zurückgreifen kann, wie sie für die Silbe des Standarddeutschen gelten: So fand Lindner (2009) eine Vielzahl von Verstößen der Silbenbildung einzelner Kinder gegen grammatische Regelmäßigkeiten, denen die Silbe des Standarddeutschen unterliegt, insbesondere gegen die im Deutschen gültige Auslautverhärtung (ebd.: 294). Dies soll verständlich machen, dass bei der Silbenbestimmung für Lautformen eines Kindes auch bei Anwendung all der grammatischen Regularitäten des Deutschen systematische Fehler auftreten können.

Zudem wird durch die Zuordnung von internukleiden Einzelkonsonanten zu beiden Silbenkernen der Lautformenkontext widergespiegelt, in dem sich die beiden Silben befinden, der u.a. durch die aufeinanderfolgende Artikulation von K_1 zu C zu K_2 gegeben ist, wodurch C mit beiden Nuklei in Relation steht.

Dem Verfasser ist bewusst, dass die vorgenommene Silbeneinteilung angreifbar ist. Um die Frage zu lösen, muss der Prozess der Silbenbildung und der darin enthaltene Prozess der ständigen Umschichtung des Silbensystems im Spannungsfeld eines begrenzten phonetischen und phonologischen Raumes weiter untersucht werden, wobei idealtypische Silbenbildungen in ihrem Verhältnis zu historischen Sedimentationen im Mittelpunkt stehen müssten.

1464 unterschiedliche Silben nach Ausschluss der Diakritikamodifikation. Diese Silbenmenge umfasst 114 Silbenköpfe, 73 Silbenkerne und 82 Silbenkodas.

7.3. Verteilungsanalyse der Silbenkonstituenten in den Korpora

7.3.1. Ausgangslage, Fragestellung und methodisches Vorgehen

Die Arbeit konnte wegen ihrer theoretischen Ausrichtung hinsichtlich der Korpusuntersuchung auf keine vergleichbaren Untersuchungen zurückgreifen, weil alle Untersuchungen zum L1-Phonemerwerb den Diskurstheorien (s.o.) zuzuordnen sind. Dadurch erfolgt in diesen Ansätzen eine Gruppierung von kindlichen Lautformen nach für sie angenommenen Zielsprachbedeutungen entsprechend der in den Abschnitten 2.1. und 4.7. beschriebenen Weise, es werden zudem Lautformen ignoriert, die keiner Zielsprachbedeutung zuzuordnen sind, und nur in den seltensten Fällen wird den sprachlichen Formen des Kindes ein eigener Status mit einer eigenen Bedeutung in Form der sogenannten Protowörter zugesprochen. Während in den Diskurstheorien die Frage verfolgt wird, wie es zu den Abweichungen der kindlichen Lautformen von den für sie angenommenen Zielsprachformen kommen kann (s.o.), verfolgt diese Arbeit das Ziel, die Bildung der Lautformen des Kindes als eigene bedeutungsvolle Zeichen und gebildet als Kombinationen eines sich über die Prinzipien der Ökonomie und Differenzierung selbst steuernden Silbensystems zu verstehen. Angesichts dieser Ausgangslage und neuartigen Zielsetzungen steht bei dieser Korpusuntersuchung mehr die explorative Datenanalyse im Vordergrund, sowie die Herstellung gedanklicher Relationen zwischen als wichtig erachteten Bestandteilen des Silbenbildungsprozesses und einer damit verbundenen Ordnung (zeitlich-statisch, (nicht-)relational) des Korpus. Die Korpusuntersuchungen beinhalten nicht die Anwendung statistischer Methoden oder eines mathematischen Kalküls, sollen aber als Vorbereitung für eine weitere Bestimmung des gesuchten Abstandsmaßes der Silbenbildung verstanden werden. Die Auszählung von Mengen und die Herstellung von Relationen zwischen ihnen ist eine Vorbedingung für die Entwicklung eines quantitativen kausalen (mathematischen) Modells. Diesen Zusammenhang betont auch Zellig Harris (1968):

[...] to show how one can arrive at an abstract system which characterizes precisely natural language. This is done by taking the data of language and finding within the data such relations as can be organized into a suitable model. The problem here was not to find a broad mathematical system in which the structure of language could be included, but to find what relations, or rather relations among relations, were necessary and sufficient for language structure

Bevor ein mathematisches Kalkül anzuwenden ist, müssen durch die Herstellung von Relationen oder Relationen von Relationen entsprechend der Aussage Zellig Harris' als wichtig erachtete Strukturbestandteile isoliert und zueinander in Beziehung gesetzt werden, um so Anhaltspunkte für die vermutete Abstandsregelung zu finden, wobei Zellig Harris jedoch eine synchrone Sprachbeschreibung anstrebt und in dieser Arbeit eine diachrone Untersuchung im Mittelpunkt steht. So wurden beispielsweise zunächst Silbennengen des Korpus auf die aufeinanderfolgenden Sitzungen bezogen, im Laufe der Untersuchungen stellte sich aber heraus, dass die Verwendung des "Silbenzeitstrahls", d.h. einer für die Silben angenommenen absoluten Auftretensreihenfolge, viel besser als Bezugszeitstrahl für andere Mengen des Silbenbildungsprozesses dient. Die zeitliche Folge aller unterschiedlichen Silben des gesamten Korpus wird als *Silbenzeitstrahl* bezeichnet, obwohl die absolute Reihenfolge der Silben nur näherungsweise angenommen werden kann, da sie sich nicht weiter über die pauschale Reihenfolge der Sitzungen hinaus aufschlüsseln lässt. Diese Entscheidung musste auch wegen eines besonders schwer wiegenden Mangels der Korpora getroffen werden, der in einer nicht klar durchgehaltenen Zuordnung von Zeiteinheiten zu den produzierten Lautmengen besteht. Somit entsprechen Teilmengen von Silben mit direkt aufeinanderfolgender, ununterbrochener Reihenfolge Zeiträumen. Entsprechend wird auch bei dem einer Silbe zugeordneten Zeitpunkt vom *Silbenzeitpunkt* gesprochen. Die Existenz von vergleichbaren Korpora ohne diesen Mangel würde eine präzisere und weitergehende Analyse aller hier behandelten Aspekte ermöglichen.

Die Analyse der Korpora beschäftigt sich mit der Verteilung der Silbenkonstituenten, um Indizien für die angenommene Aufteilung der phonetischen und phonologischen Räume zu finden, die sich in „Abstandsregelungen“ manifestiert. Es wird angenommen, dass Übereinstimmungen bzw. fehlende Übereinstimmungen bestimmter Mengen als Anzeichen für Abstandsmaße angesehen werden können. So unterscheiden sich z.B. die Mengen der auftretenden Köpfe und Kodas bei allen Silbenkernen quantitativ und in ihrer Zusammensetzung. Um die innere Verteilungsstruktur zu verstehen, werden einfache Verteilungen von Mengen festgestellt wie z.B. die statische (auf die Gesamtkorpora bezogen) Rangordnung der Silbenkerne oder der Köpfe und Kodas. Sie ermöglichen unter anderem einen Vergleich von Korpora. Die dynamische Verteilung von Mengen ermöglicht dagegen eine Zuordnung des zeitlichen Aspekts zu bestimmten Mengen und ist eine Vorbedingung für die Beurteilung von zeitlichen und evt. kausalen Abhängigkeiten. Da die Mengen der Silbenkonstituenten tatsächlich in gebundener Form in der Silbe vor-

kommen, lässt sich über die Isolierung von statischen und dynamischen binären Relationen weiterer Aufschluss über die Silbenentwicklung gewinnen. So wurde z.B. bisher die quantitative Gemeinsamkeit von Silbenkernpaaren hinsichtlich ihrer Schalenmengen noch nie untersucht und auch nicht das Verhältnis von einmalig vorkommenden Konstituenten im Zeitverlauf zu mehrfach vorkommenden. Beide Relationen und weitere in den folgenden Korpusuntersuchungen präsentierte Relationen und Prozessgrößen werden unter dem Gesichtspunkt der Abstandsregelung diskutiert.

Die Verteilungsanalyse der Kopora konzentriert sich v.a. auf eine globale Betrachtung von Silbenbestandteilen und nicht auf die Erklärung des Verlaufs von einzelnen konkreten Silbenbestandteilen. Letzteres ist erst zu leisten, wenn ausgehend von einer globalen Analyse, d.h. einer Gruppierung und Inbeziehungsetzung von Gruppen und Einzelbestandteilen der durch die angenommene Abstandsregelung angenommene funktionale Zusammenhang näher eingegrenzt wurde. Dies beinhaltet die Feststellung von (globalen) Verteilungsregularitäten von (Teil-)Mengen des jeweils untersuchten Korpus. Denn die gesuchte Abstandsregelung kann primär nur durch eine globale Analyse verstanden werden. Deshalb wird auch von Überlegungen abgesehen, wie sie Jakobson (1941) und Locke (1983) hinsichtlich einzelner konkreter Phoneme oder phonetisch gruppierter Phonemengen vornehmen. Phonetische Gruppierungen oder die Betrachtung einzelner konkreter Phoneme spielen bei der Analyse nur insofern eine Rolle, dass sie als die Realisierung einzelner konkreter Silbenbestandteile innerhalb einer globalen Verteilungslösung aufgefasst werden.

7.3.2. Darstellungsform der Ergebnisse und ihre Grundquantitätsarten

Die als wichtig erachteten Relationen sind als Graphiken dargestellt, und somit ist das Ergebnis der Korpusuntersuchungen primär visuell-darstellend. Die visuelle Darstellung beschränkt sich im Wesentlichen auf die folgenden sechs quantitativ-temporalen Relationsarten: Mit dem Begriff des *Merkmals* soll im Folgenden eine Menge von Elementen bezeichnet werden, beispielsweise eine Menge von Werten, eine Menge von Silbenbestandteilen oder eine Menge von Mengen von Silbenbestandteilen. Die Elemente dieser betrachteten Menge sollen wiederum als *Werte* dieses Merkmals benannt werden.

1.) Häufigkeitsverteilung eines Merkmals im Gesamtzeitraum:

Bei einer sog. *Häufigkeitsverteilung* eines Merkmals wird die Gesamthäufigkeit jedes seiner Werte (z.B. die Menge aller Silbenkerne) in einer festgelegten Bezugsmenge (z.B.

hier: die Menge aller unterschiedlichen Silben eines Korpus) bestimmt, in der jedes Element (z.B. hier: eine Silbe) dieses Merkmal aufweist.

2.) Kumulative Berechnung bestimmter Werte im Zeitverlauf:

Es liegen in diesem Fall zeitlich geordnete quantitative Werte $(f_i)_{i \geq 0}$ vor. Aus dieser Folge $F := (f_i)_{i \geq 0}$ zeitlich geordneter Werte wird eine neue zeitliche Folge $F' := (f'_i)_{i \geq 0}$ von Werten berechnet, indem der zeitlich erste Wert f'_0 der neuen Folge F' dem ersten der ursprünglich gegebenen Folge F entspricht, d.h. $f'_0 = f_0$ und sukzessive jeder dem ersten Wert der neuen Folge nachgeordnete Wert f'_j ($j \geq 1$) als Summe aus dem direkt vorhergehenden Wert f'_{j-1} der neuen Folge F' und dem Wert f_j der Ursprungsfolge F mit dem gleichen zeitlichen Rang j berechnet wird, d.h. $f'_j = f'_{j-1} + f_j$.

3.) Anteilswerte von Merkmalswerten an zeitlich variablen Silbenzeiträumen:

Zusätzlich zu der Gesamtsilbenmenge eines Korpus wird die zeitliche Dimension der Silben in Form der Reihenfolge ihres Auftretens beachtet. Dadurch können Silbenmengen als Silbenzeiträume begriffen werden, für die Anteilswerte berechnet werden.

4.) Mengenmäßiger Vergleich unterschiedlicher, gleich großer Silbenzeiträume:

Der Silbenzeitstrahl wird in gleich große, überschneidungsfreie und direkt aufeinanderfolgende Silbenteilzeiträume oder -mengen eingeteilt, die dann auf Grund ihrer gleichen Größe hinsichtlich weiterer quantitativer Werte verglichen werden können.

5.) Streudiagramme:

Bei einem sog. *Streudiagramm* werden für jedes Element einer zu untersuchenden Population, z.B. der Menge aller unterschiedlichen Silben des jeweils betrachteten Korpus, die Werte zweier Merkmale (z.B. Zeitpunkt des ersten Auftretens und Gesamthäufigkeit eines Silbenbestandteils) gemessen und jeweils an der dem jeweiligen Merkmal zugeordneten Achse eines zweidimensionalen kartesischen Koordinatensystems abgetragen (Ott 2010: 105ff.).

6.) Phaseneinteilungen:

Bei einer *Phaseneinteilung* wird der Silbenzeitstrahl nach bestimmten a priori festgelegten objektiven Kriterien in Zeiträume, sog. *Phasen* eingeteilt. Eine solche Phaseneinteilung wird in Abschnitt 7.3.9.7. vorgenommen.

Neben der reinen Form dieser Grundquantitätsarten werden in den Abbildungen auch Relationen bzw. Verhältnisse von ihnen dargestellt, so beispielsweise das Verhältnis zweier kumulativer Werte (im Zeitverlauf) oder das Verhältnis aus einem kumulativen und einem Anteilswerte im Zeitverlauf.

7.3.3. Aufführung und Wiederholung einiger häufiger Begriffe

Folgende Begriffe werden während der Untersuchung häufig erwähnt und sollen wegen ihres nicht unbedingt gängigen Gebrauchs vorab aufgeführt werden:

- *Silbenzeitstrahl*: Alle Silben eines Korpus haben einen Zeitbezug, indem sie nacheinander in unterschiedlichen Lautformen des Kindes auftreten und Silben in mehrsilbigen Lautformen ebenfalls eine zeitliche Folgerelation aufweisen. Durch diesen Zeitbezug lassen sich die Silben, zumindest rein theoretisch, in eine absolute Reihenfolge bringen. Die untersuchten Korpora beziehen das Auftreten von Lautformen auf einzelne Tage oder andere zeitliche Gruppierungen, und deshalb liegen die Silben in keiner absoluten Reihenfolge vor. Dennoch wird die sich aus der Reihenfolge der digitalisierten Lautformen für die aus ihnen bestimmten Silben ergebende absolute Reihenfolge angenommen und als *Silbenzeitstrahl* bezeichnet. Zukünftige Korpora zum Sprach- und insbesondere zum Phonemerwerb müssen die absolute Reihenfolge aller auftretenden Lautformen festhalten.
- *Silbenzeitraum*: Ausgehend von dem Begriff des *Silbenzeitstrahls* (s.o.) wird eine Menge von unterschiedlichen Silben eines Korpus als *Silbenzeitraum* bezeichnet, wenn sie eine durchgehende Folge von Silben des Silbenzeitstrahls bildet. D.h. eine Menge von Silben eines Korpus wird mit *Silbenzeitraum* benannt, wenn zu je zwei ihrer Silben auch alle die Silben des Silbenzeitstrahls in ihr enthalten sind, die zwischen diesen beiden Silben im Silbenzeitstrahl positioniert sind.
- *Silbenzeitpunkt*: Der Begriff des *Silbenzeitpunktes* bezeichnet die Position einer Silbe innerhalb der durch den Silbenzeitstrahl gegebenen absoluten Reihenfolge. So ist beispielsweise die 17-te bzw. 35-te Silbe innerhalb eines Silbenzeitstrahls der 17-te bzw. 35-te Silbenzeitpunkt.
- *Silbengesamthäufigkeit* eines Merkmals:
Mit *Silbengesamthäufigkeit* soll die Anzahl eines Silbenmerkmals wie beispiels-

weise eines bestimmten Kopfes in der Menge aller Silben des Gesamtkorpus bezeichnet werde.

- *Binärrelationen*: Der Begriff der *Binärrelation* fasst hierbei die des Körpers (d.h. der Kombination aus Silbenkopf und -kern), des Silbenreims (d.h. die Kombination aus Silbenkern und -koda) und der Schale (d.h. die Kombination aus Silbenkopf und -koda) zusammen.

7.3.4. Zur Gliederung des Vorgehens bei der Analyse

Im Folgenden werden die Ergebnisse der Untersuchung der beiden die deutsche Sprache lernenden Kinder Annalena und Isa vorgestellt. Für die Untersuchung wurde eine systematische Einteilung des Korpus in den Gesamtzeitraum bzw. das Gesamtkorpus betreffende Analysefragen, auch *statische Abfragen* genannt, und den Zeitverlauf betreffende, d.h. den Bezug der Analyse von Werten auf Teilzeiträume, auch *dynamische Analyse* genannt, vorgenommen. Sowohl in der statischen als auch der dynamischen Rubrik wurde weiter danach differenziert, ob sich die Analyseabfragen auf Silben oder -bestandteile an sich (d.h. Einzelkonstituenten, -silben oder -binärrelationen) beziehen oder darüber hinaus eine Relation zwischen verschiedenen Silbenbestandteilen hergestellt wird. Folglich werden die Analysen der Verteilungen in den beiden Korpora in den fünf Rubriken "Statische Verteilung (nicht-relational)" (6.3.5.), "Statische Verteilung (relational)" (6.3.6.), "Statisch - dynamische Relationen" (6.3.7.), "Dynamische Verteilungen (nicht-relational)" (6.3.8.) und "Dynamische Verteilungen (relational)" (6.3.9.) vorgestellt und anschließend eine Zusammenfassung und Interpretation der Ergebnisse (6.3.10.) gegeben.

Um Klarheit über die quantitativen Verhältnisse zu erhalten, wurden Verteilungen des Korpus gemessen bzw. durch Auszählung und Bildung von Relationen hergeleitet. Zuerst sollen die statischen Abfragen vorgestellt werden, weil sie quantitative Verhältnisse im Korpus darstellen und so einen Gesamtüberblick geben, während sich die dynamischen Abfragen auf zeitlich variable Silbenteilmengen des Korpus beziehen und nicht das gesamte Korpus umfassen.

7.3.5. Statische Verteilungen (nicht-relational)

Innerhalb der Rubrik der statischen Verteilungen betrachten die nicht-relationalen Abfragen Silbenbestandteile an sich, ohne sie zueinander in Relation zu setzen.

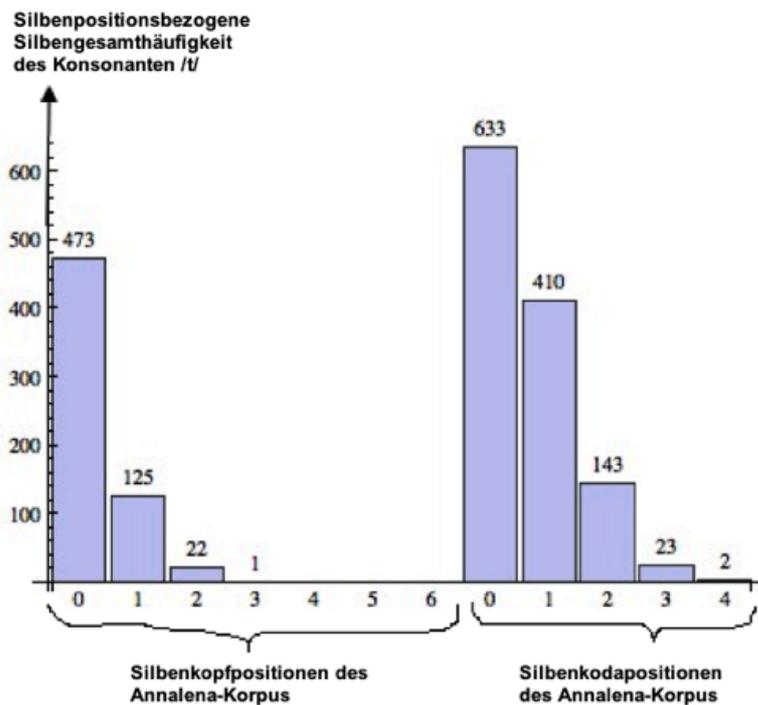
7.3.5.1. Silbenpositionsbezogene Gesamthäufigkeiten von Einzelphonemen

Im Verlauf dieser Arbeit verlagerte sich der Schwerpunkt der Analyse von der Betrachtung einzelner Phoneme ohne Bezug auf Positionen innerhalb der Silbe zu einer positionsbezogenen Untersuchung, da Eigenschaften von Phonemen, zu denen hier auch das quantitative Auftreten gezählt wird, sich als positionsgebunden erweisen. So haben die Phoneme beispielsweise eine Silbengesamthäufigkeit, die von ihrer Silbenposition abhängig ist. Der Begriff der *Silbenposition* wird dabei auf die Konstituenten aller Silben eines bestimmten Korpus bezogen. Im folgenden Schaubild sind alle möglichen Positionen für die Kopf- und Kodakonstituente aller Silben des Annalena-Korpus gezeigt. Die maximale Anzahl der Phoneme in den Köpfen aller Silben des Annalena-Korpus ist 7, weil die Silbe /bObObOgɪd/ mit dem Kopf /bObObOg/ als einer Folge aus 7 Konsonanten im Korpus vorkommt. Analog beträgt die maximale Anzahl von Silbenkodapositionen 5, weil maximal 5 Phoneme in den Kodas aller Silben des Annalena-Korpus vorkommen können. Beispielsweise tritt die Silbe /pɪkOθOt/ mit der Koda /kOθOt/ auf, das aus einer Folge von 5 Konsonanten (/k, O, θ, O, t/) besteht. Das nachfolgende Schaubild zeigt die Anzahl der Silben, die den Konsonanten /t/ an einer bestimmten Position des Silbenkopfes (Silbenkopfposition) bzw. der Silbenkoda (Silbenkodaposition) aufweisen, wobei die erste Position des Silbenkopfes bzw. der Silbenkoda mit der Ziffer "0", die zweite mit der "1", usw. bezeichnet ist. Die Ziffernbezeichnung der Positionen ist also so gewählt, dass sie für ein Phonem x in der Kopf- oder Kodakonstituente an der Position mit der Ziffer i ($i \geq 0$) angibt, wie viele Phoneme dem Phonem x in dieser Konstituente vorausgehen. So steht vor dem Phonem x an der Position "0" kein Phonem, an der Position "1" eins, etc.

Im Folgenden ist sowohl für den Kopf als auch für die Koda jeweils für jede Position, an der der Konsonant /t/ im Korpus vorkommen kann, eine Silbe als Beispiel angeführt, wobei die Silben in der Reihenfolge der aufsteigenden Positionsgröße angegeben sind, zunächst für den Silbenkopf: /ta/, /stɪf/, /sOtam/, /dʁttɪ/ und die analoge Silbenfolge für die Silbenkoda /bɪt/, /mənt/, /gɪnst/, / bukOθt/, /pɪkOθOt/.

Die Abb. 3 zeigt die Silbengesamthäufigkeit des Konsonanten /t/:

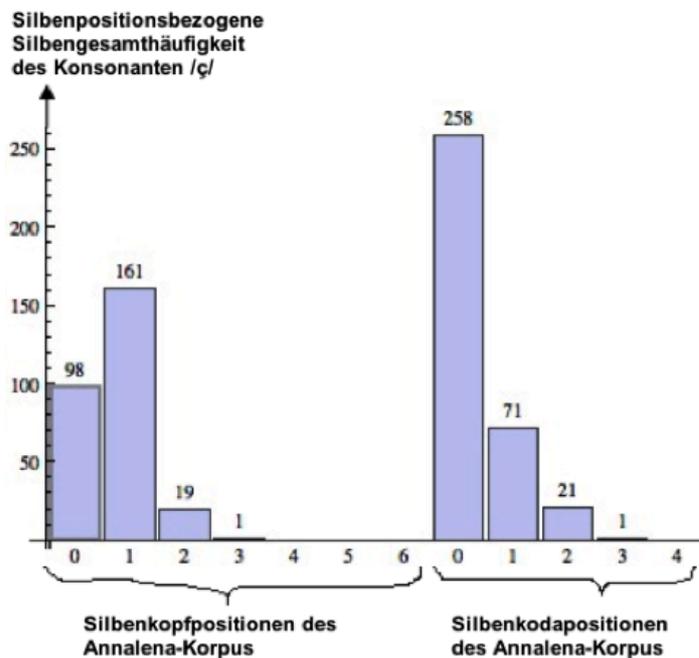
Abbildung 3: Silbenpositionsbezogene Silbengesamthäufigkeit von /t/



Für jede der im Annalena-Korpus möglichen sieben⁶⁷ Positionen des Kopfes und jede der fünf in der Kodaposition sind die Mengen der unterschiedlichen Silben des Gesamtkorpus als Säulenhöhe und entsprechender Zahl über der Säule abgetragen, in denen der Konsonant /t/ an entsprechender Position vorkommt. Es wird deutlich, dass die Silbengesamthäufigkeit des Konsonanten /t/ auf einer Position mit zunehmender Entfernung vom Silbenkopf bzw. von der Silbenkoda abnimmt und somit insbesondere von der Silbenposition abhängt. Das Häufigkeitsmaximum muss dabei aber nicht mit der ersten Position der Silbenkonstituenten identisch sein: So befindet es sich beispielsweise beim palatalen Frikativ /ç/ auf der zweiten Position des Silbenkopfes (Abb. 4):

⁶⁷ Diese ungewöhnlich hohe Anzahl an Kopfpositionen wird durch eine Silbe hervorgerufen, die als Silbenkopf die entsprechende Zahl von Konsonanten vor dem Silbenkern aufweist und nicht als silbisch gekennzeichnet sind (s.o.).

Abbildung 4: Silbenpositionsbezogene Silbengesamthäufigkeit von /ç/



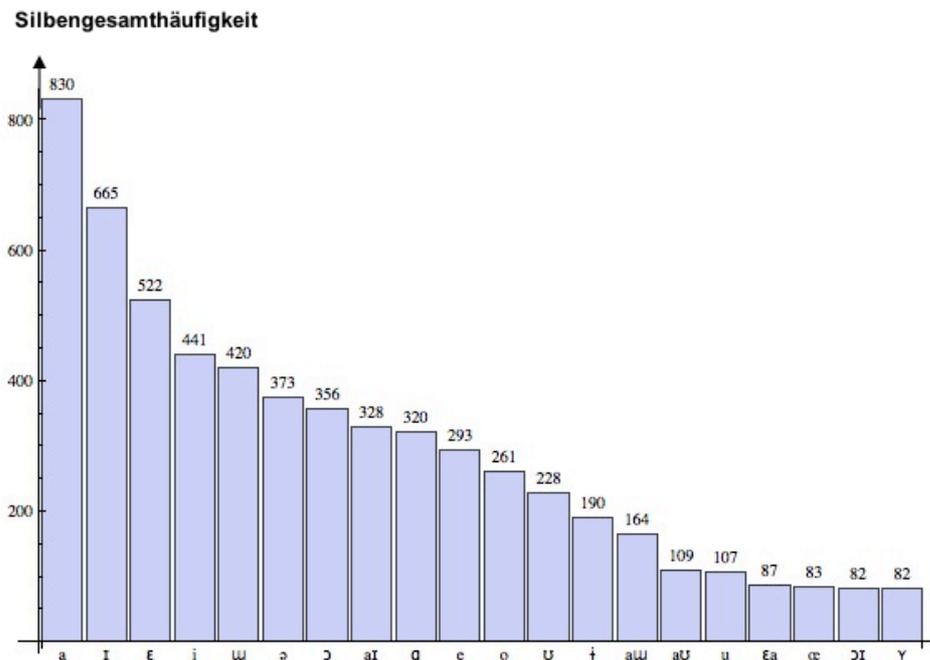
Der Positionsbezug der Phoneme wird jedoch durch die Überlegung eingeschränkt, die Silbenpositionen nicht einzeln für sich zu betrachten, sondern sie zu sog. (Silben)-Konstituenten zusammenzufassen⁶⁸. Deshalb werden in dieser Arbeit statt positionsloser Einzelphoneme (positionsbezogene) Silbenkonstituenten in Form der Köpfe, Kerne und Kodas untersucht.

7.3.5.2. Silbengesamthäufigkeitsverteilung der häufigsten Silbenkerne

Die folgenden Abbildungen zeigen die Silbengesamthäufigkeiten der zwanzig häufigsten Silbenkerne für das Annalena- und das Isa-Korpus:

⁶⁸ siehe Abschnitt 5.3..

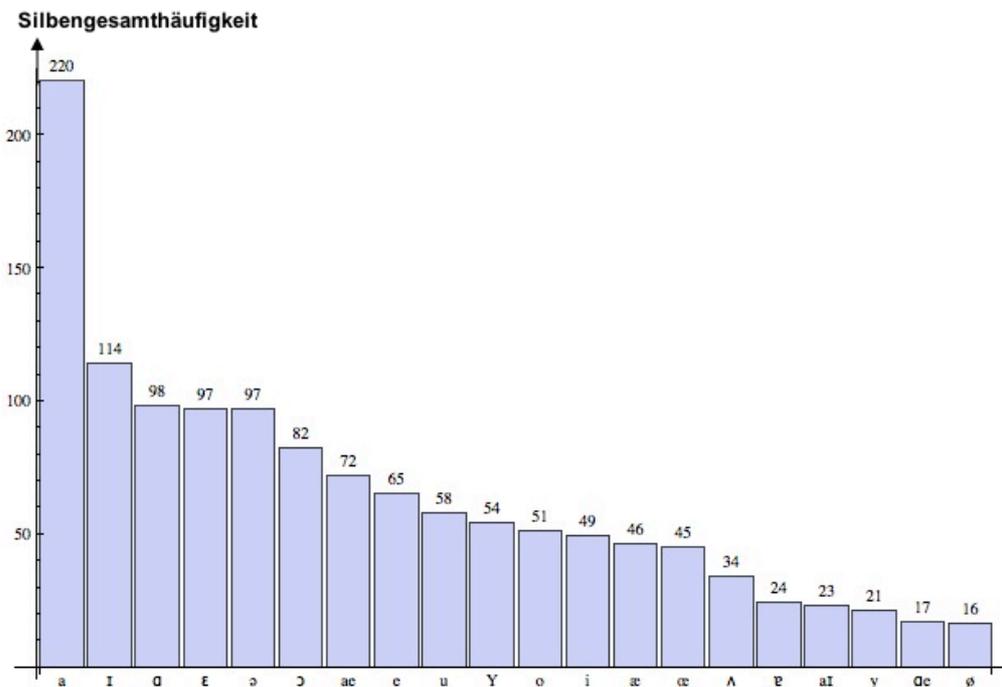
Abbildung 5: Silbengesamthäufigkeiten der 20 häufigsten Nuklei (Annalena-Korpus)



Die 20 häufigsten Nuklei des Annalena-Korpus.

Z.B. hat der Nukleus /a/ die Silbengesamthäufigkeit 830, d.h. im gesamten Korpus kommt der Silbenkern /a/ in 830 unterschiedlichen Silben vor. Der Silbenkern /ɔ/ kommt hingegen nur in 356 unterschiedlichen Silben im gesamten Annalena-Korpus vor, d.h. hat eine Silbengesamthäufigkeit von 356.

Abbildung 6: Silbengesamthäufigkeiten der 20 häufigsten Nuklei (Isa-Korpus)



Die 20 häufigsten Nuklei des Isa-Korpus.

Z.B. hat der Nukleus /a/ die Silbengesamthäufigkeit 220, d.h. im gesamten Korpus kommt der Silbenkern /a/ in 220 unterschiedlichen Silben vor. Der Silbenkern /ø/ kommt hingegen nur in 16 unterschiedlichen Silben im gesamten Isa-Korpus vor, d.h. hat eine Silbengesamthäufigkeit von 16.

Bei beiden Kindern fällt auf, dass sich die Mehrheit der Silben auf eine relativ geringe Anzahl von Silbenkernen konzentriert, während die Mehrheit der Nuklei jeweils in nur sehr wenigen Silben vorkommt: Das Annalena-Korpus enthält insgesamt 7190 unterschiedliche Silben, die 20 häufigsten Silbenkerne kommen insgesamt in 5941 unterschiedlichen Silben vor und haben damit einen Anteil an der Gesamtsilbenbildung von 82,6 %. Für das Isa-Korpus gilt ein analoges Verhältnis: Es enthält insgesamt 1464 unterschiedliche Silben, die 20 häufigsten Silbenkerne kommen insgesamt in 1283 unterschiedlichen Silben vor und haben damit einen Anteil an der Gesamtsilbenbildung von 87,6%.

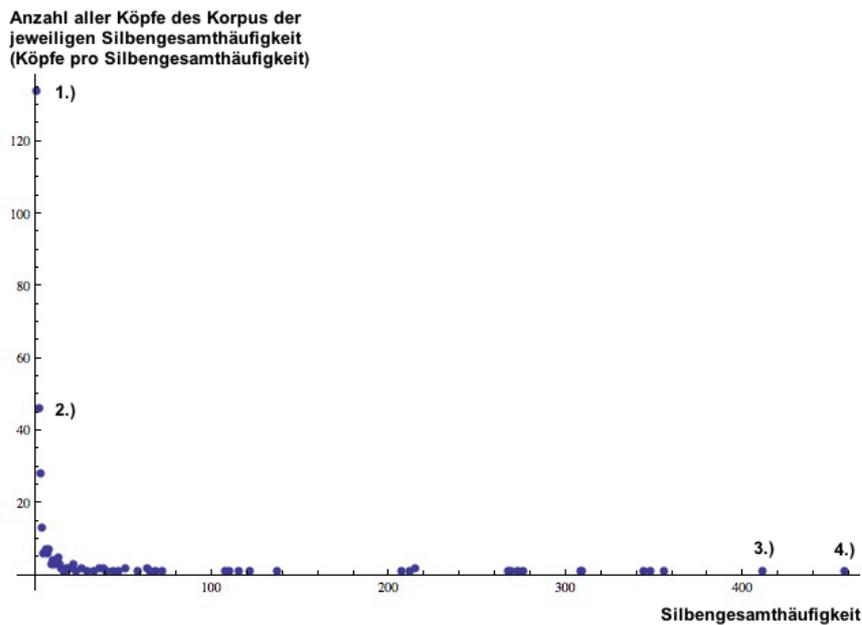
Trotz dieser Gemeinsamkeiten bei beiden Kindern gibt es Nuklei, die in der betrachteten Häufigkeitsverteilung unterschiedliche Ränge besetzen: So hat der Silbenkern /a/ im Isa-Korpus den Gesamthäufigkeitsrang 3 inne, d.h. er ist der dritthäufigste Nukleus aller Nuklei des Isa-Korpus, während dieser Silbenkern im Annalena-Korpus den Rang 9 besetzt. Hingegen hat der Silbenkern /ε/ im Annalena-Korpus den Gesamthäufigkeitsrang 3 inne, d.h. er ist der dritthäufigste Nukleus aller Nuklei des Annalena-Korpus, während dieser Silbenkern für das Isa-Korpus den Rang 4 besetzt.

7.3.5.3. Häufigkeitsverteilungen der Silbengesamthäufigkeiten von Silbenbestandteilen einer bestimmten Art

Die Häufigkeitsverteilung der Köpfe zeigt eine ungleiche Verteilung, von denen einige sehr häufig, dagegen sehr viele nur einmal vorkommen.

Es gibt also sehr viele Köpfe mit einer sehr niedrigen Silbengesamthäufigkeit und sehr wenige mit einer hohen Silbengesamthäufigkeit.

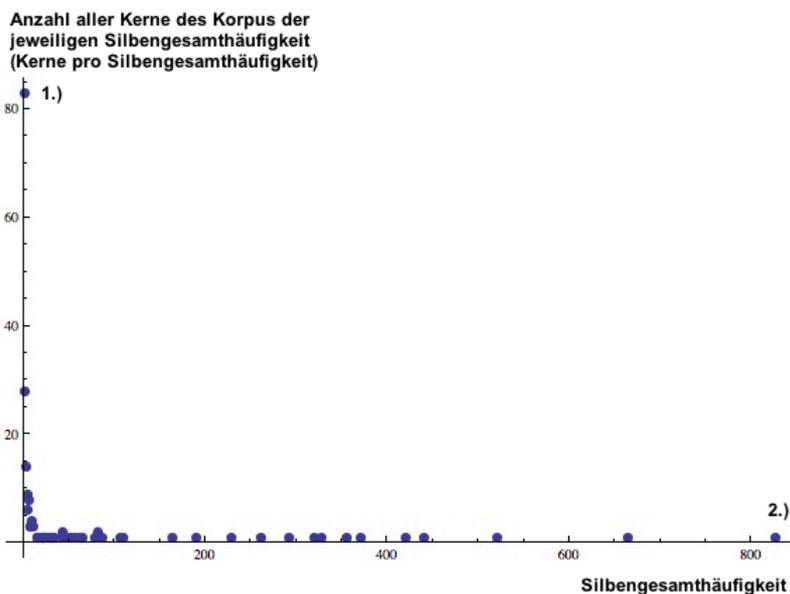
Abbildung 7: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Köpfe



In der Abb. 7 sind vier Punkte mit den Zifferen 1.) bis 4.) gekennzeichnet, die jeweils die folgende Bedeutung haben: Im Korpus sind insgesamt 134 unterschiedliche Köpfe vorhanden, die jeweils nur in einer einzigen Silbe enthalten sind (Punkt "1."). Hingegen gibt es im gesamten Korpus 46 unterschiedliche Köpfe, die jeweils in insgesamt 2 unterschiedlichen Silben enthalten sind (Punkt "2.") und jeweils nur einen Kopf, der in insgesamt 411 (Punkt "3.") oder 457 (Punkt "4.") unterschiedlichen Silben vorkommt.

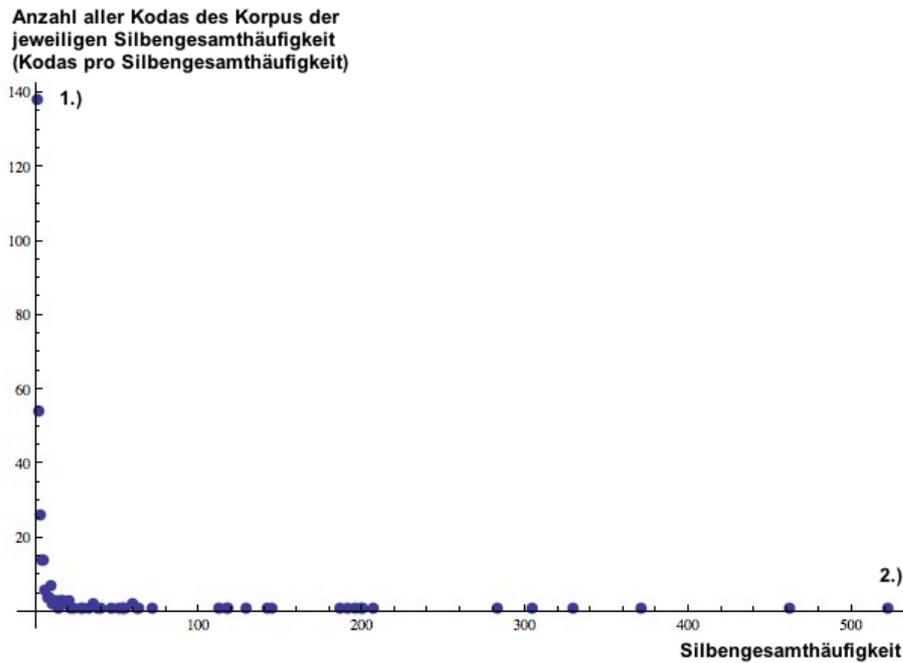
Eine ebenso ungleiche Verteilung ist auch für die Kerne, Koda und Schalen zu beobachten: Es gibt also sehr viele Nuklei mit einer sehr niedrigen Silbengesamthäufigkeit und sehr wenige mit einer hohen Silbengesamthäufigkeit.

Abbildung 8: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Nuklei



Der mit "1.)" gekennzeichnete Punkt in Abb. 8 besagt, dass im Gesamtkorpus 83 Kerne vorhanden sind, die jeweils nur in einer einzigen Silbe vorkommen und aber demgegenüber 1 Kern, der insgesamt 826 unterschiedliche Silben mit bildet (Punkt "2.")). Auch für die Kodas gilt, dass es sehr viele mit einer sehr niedrigen Silbengesamthäufigkeit und sehr wenige mit einer hohen Silbengesamthäufigkeit gibt.

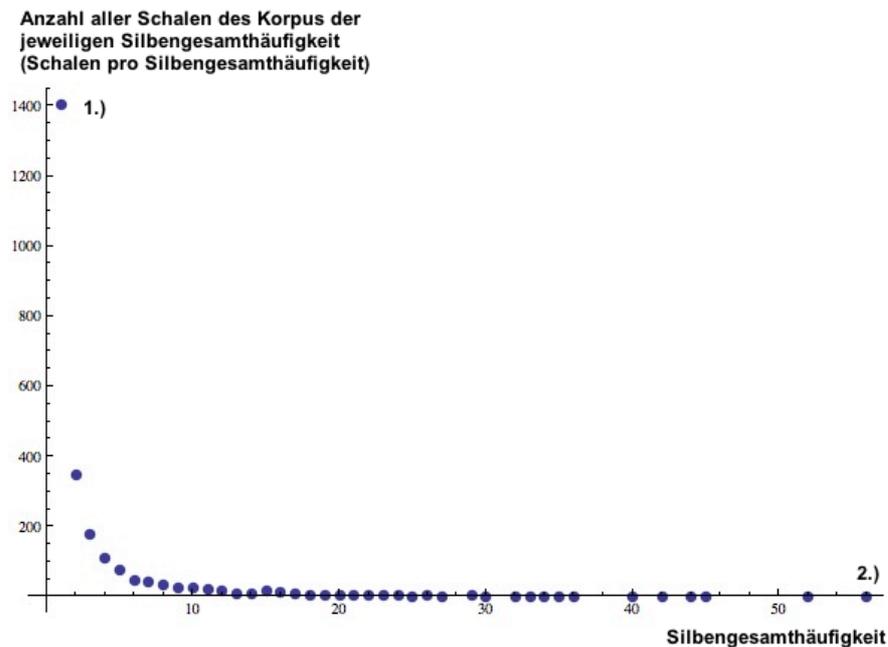
Abbildung 9: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Kodas



Der mit "1.)" gekennzeichnete Punkt in Abb. 9 besagt, dass im Gesamtkorpus 138 Kodas vorhanden sind, die jeweils nur in einer einzigen Silbe vorkommen, demgegenüber aber 1 Koda, die insgesamt 462 unterschiedliche Silben mit bildet (Punkt "2.)).

Wie zuvor für die Köpfe, Kerne und Kodas gilt auch für die Schalen eine vergleichbare Aussage: Es gibt sehr viele Schalen mit einer sehr niedrigen Silbengesamthäufigkeit und sehr wenige mit einer hohen Silbengesamthäufigkeit.

Abbildung 10: Häufigkeitsverteilung der Silbengesamthäufigkeiten aller Schalen



Der mit "1.)" gekennzeichnete Punkt in Abb. 10 besagt, dass im Gesamtkorpus 1402 Schalen vorhanden sind, die jeweils nur in einer einzigen Silbe vorkommen und aber demgegenüber eine einzige Schale, die insgesamt 56 unterschiedliche Silben mit bildet (Punkt "2.")).

7.3.5.4. Verteilungen der Binärrelationsarten im Gesamtkorpus

Die Anzahl unterschiedlicher Binärrelationen jeder Art sind in Tab. 3 für das Annalena- und das ISA-Korpus aufgeführt:

Tabelle 4: Anzahl unterschiedlicher Schalen, Körper und Reime in der Silbengesamtmenge

	Anzahl unterschiedlicher Binärrelationen der jeweiligen Binärrelationsart			Summe
	Schalen	Körper	Reime	
Annalena	2154	2103	1636	5893
ISA	417	722	337	1476

Bei beiden Kindern sind die Reime (d.h. die Kern-Koda-Kombinationen) am wenigsten zahlreich ausgeprägt. Hingegen sind bei Annalena die Schalen etwas häufiger als die Körper, während bei ISA hingegen die Körper erheblich zahlreicher sind⁶⁹.

⁶⁹ Evtl. ist dies auch ein Unterschied, der der nicht übereinstimmenden Detailliertheit der Korpora geschuldet ist.

7.3.5.5. Gesamthäufigkeiten von Silben in den Lautformen

Um eine Brücke von der Silbe zum Wort zu schlagen, soll hier kurz auf die Kombination von Silben zu mehrsilbigen Lautformen und damit auf ihre mögliche Wiederholung im Lautformenkontext eingegangen werden. In dieser Arbeit wurde primär die Menge aller Silben untersucht, die sich aus den Lautformen einsilbiger Wörter oder den aus mehrsilbigen Lautformen extrahierten Silben ergibt. Diese Fokussierung der Untersuchung auf die Betrachtung von Silben ergibt sich aus der Annahme des grundlegenden Status der Silbe auf der lautlichen Ebene der Sprache. Ein grundlegender Status wird der Silbe auch in Vennemann (1978) eingeräumt, in dem Sinne, dass u.a. die Phoneme, der Akzent und weitere Merkmale der Silbe als hierarchisch höhergestellter Einheit als Eigenschaften in einem Baummodell zugeordnet werden (ebd.: 189). In dieser Arbeit wird das Entstehen dieser Eigenschaften aus den holistischen Silbeneinheiten durch Schnittmengenbildung innerhalb eines funktional-dynamischen Silbensystems erklärt. Fasst man die Relationen innerhalb des in Vennemann (1978) gegebenen Baummodells (ebd.: 1978) als durch Schnittmengenbildung gegeben auf, so spiegelt es nicht die Auffassung des Silben-Wort-Verhältnisses dieser Arbeit wieder. Denn wie bereits eingangs erwähnt, wird davon ausgegangen, dass das Wort ein wirkliches Silbenkombinationsprodukt ist und nicht Ausgangseinheit für eine Schnittmengenbildung mit den Silben als Endprodukten. Die Silbe ist zwar die produktive Grundeinheit der Lautproduktion, die reine Abbildung ihres Wirkens wird jedoch durch sedimentative „Ablagerungen“ gestört, die sich auf der Silben- und Wortebene durch die ständigen Verschiebungen innerhalb des begrenzten artikulatorischen und phonologischen Raumes ergeben, wodurch sich die Gestalt des Silbensystems ändert, was wiederum zu Verschiebungen im Phonemsystem führt. Dem Verfasser ist bewusst, dass diese Zusammenhänge so lange spekulativ klingen müssen, bis sie durch weitere Untersuchungen erhärtet werden können. Die Daten der hier untersuchten kindlichen Korpora sind allerdings nicht umfassend genug, um eine hinreichende Wiederholung von Silben in unterschiedlichen Lautformen zu gewährleisten.

So zeigt die folgende Tabelle die Anzahl unterschiedlicher (diakritikafreier) Lautformen und die in ihnen vorkommende Anzahl unterschiedlicher Silben jeweils für das Annalena- und das ISA-Korpus. Das Verhältnis aus der Anzahl unterschiedlicher Lautformen zu der Anzahl unterschiedlicher Silben beträgt für das Annalena-Korpus 1,412 und für das ISA-Korpus 1,888. Hingegen ist das Verhältnis für die Konstituenten in den Silben erheblich größer, so beispielsweise das Verhältnis der Anzahl aller Silben zu der aller Köpfe des Annalena-Korpus 22,54 und für das ISA-Korpus 12,84. Die Korpora

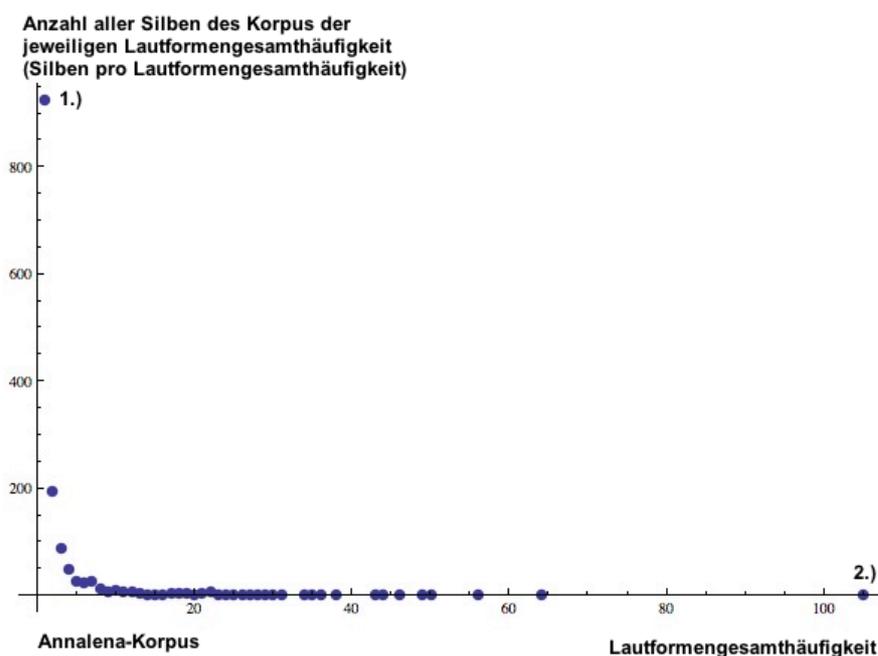
müssten hinsichtlich des Lautformenumfangs erheblich größer sein, um eine vergleichbare Untersuchung durchzuführen, wie sie für die Konstituenten in den Silben ausgeführt wurde, jedoch mit dem Unterschied, dass die Konstituenten und Phoneme als sich durch Schnittmengenbildungen aus den Silben ergebend aufgefasst werden,- auch wenn von "Kombinationen" von Konstituenten zu Silben gesprochen wird-, während die Lautformen als wirkliche Silbenkombinationsprodukte betrachtet werden (s.o.).

Tabelle 5: Anzahl aller unterschiedlichen Silben und Lautformen und Summe aller Silbenhäufigkeiten im Gesamtkorpus

	Anzahl der unterschiedlichen diakritikafreien Silben	Anzahl der unterschiedlichen diakritikafreien Lautformen (Wörter)	Summe aller Silben in ein- und mehrsilbigen Lautformen (inklusive der Wiederholungen)
Annalena	7182	10131	17818
ISA	1447	2732	4531

Die Abb. 11 stellt für das Annalena-Korpus die Häufigkeitsverteilung aller Silben in der Menge aller Lautformen dar. Es ist zu sehen, dass einige wenige Silben in sehr vielen unterschiedlichen Lautformen vorkommen, während sehr viele in nur wenigen Lautformen vertreten sind.

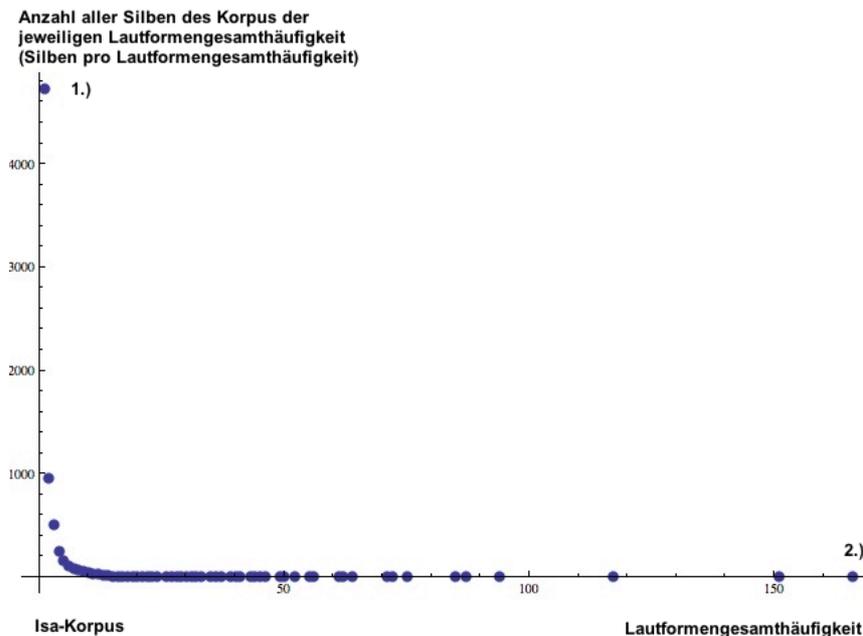
Abbildung 11: Häufigkeitsverteilung der Lautformengesamthäufigkeiten aller Silben (Annalena-Korpus)



Der mit "1.)" gekennzeichnete Punkt in Abb. 11 besagt, dass im gesamten Annalena-Korpus 4724 Silben vorhanden sind, die jeweils nur in einer einzigen Lautform vorkommen, demgegenüber aber eine einzige Silbe, die insgesamt 166 unterschiedliche Lautformen mit bildet (Punkt "2.>").

Abb. 12. zeigt die Abb. 11 entsprechende Abbildung für das ISA-Korpus:

Abbildung 12: Häufigkeitsverteilung der Lautformengesamthäufigken aller Silben (Isa-Korpus)



Der mit "1.)" gekennzeichnete Punkt in Abb. 12 besagt, dass im gesamten Isa-Korpus 925 Silben vorhanden sind, die jeweils nur in einer einzigen Lautform vorkommen und aber demgegenüber eine einzige Silbe, die insgesamt 105 unterschiedliche Lautformen mit bildet (Punkt "2.>").

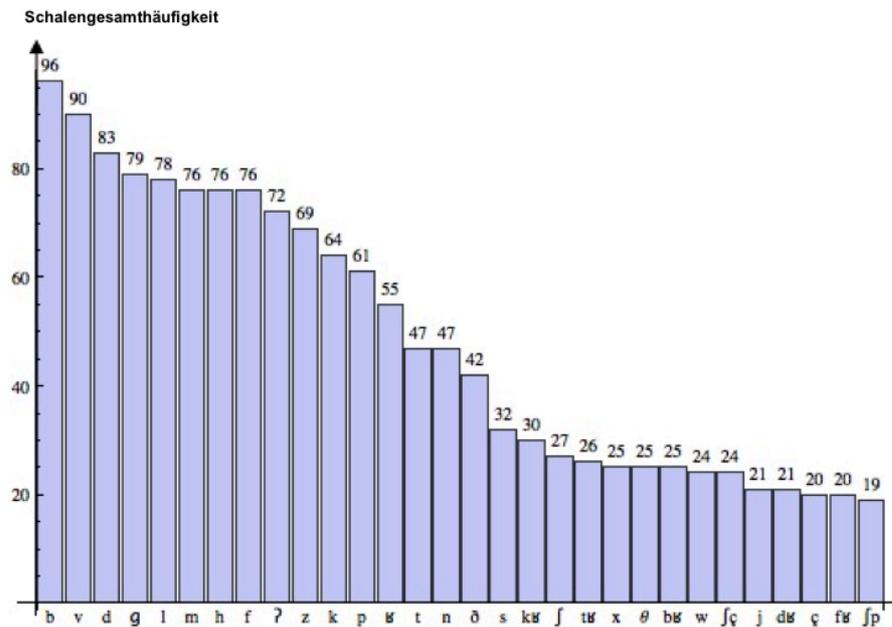
7.3.6. Statische Verteilungen (relational)

Neben einer Analyse der auf das Gesamtkorpus bezogenen Häufigkeiten von einzelnen Silbelementen zur Aufklärung ihres quantitativen Gesamtpotentials und der quantitativen Gesamtkonzentration des Prozesses wurden auch Relationen zwischen mehreren Silbenbestandteilen hergestellt.

7.3.6.1. Verteilungen der häufigsten Köpfe und Kodos in der Gesamtschalenmenge

Abb. 13 und Abb. 14 zeigen die Mengen der dreißig jeweils häufigsten Köpfe bzw. Kodos aller Schalen des Annalena-Korpus für den gesamten Zeitraum.

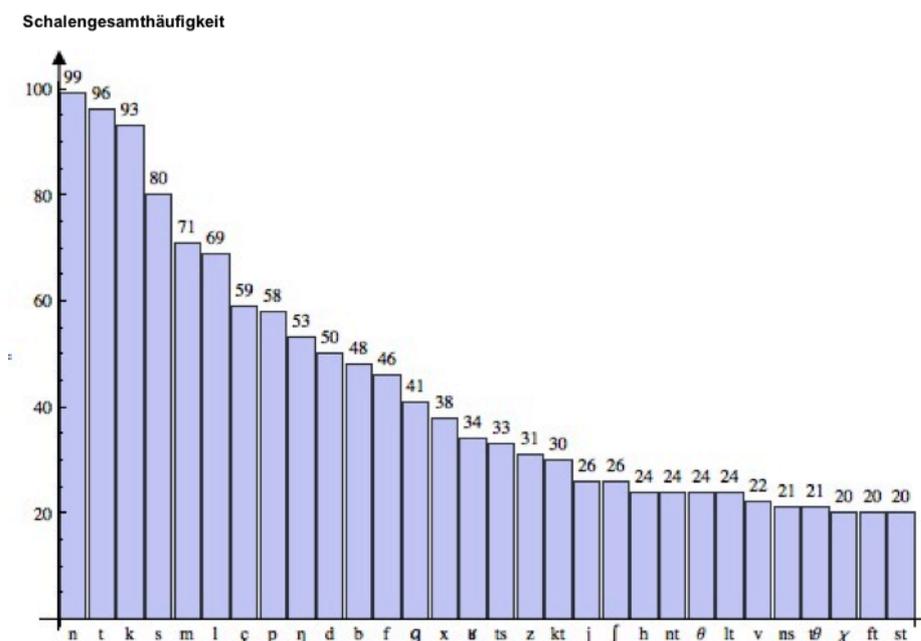
Abbildung 13: Gesamthäufigkeit der 30 häufigsten Köpfe in der Schalengesamtmenge



Die Rangfolge der 30 häufigsten Köpfe in der Menge aller Schalen.

Das vorhergehende Schaubild zeigt die Anzahl der Kodas, mit denen sich ein bestimmter Kopf verbindet: Beispielsweise "kombiniert" sich der Kopf /b/ mit 96 unterschiedlichen Kodas zu eben so vielen Schalen, während dies für den Kopf /m/ nur für 76 unterschiedliche Kodas der Fall ist.

Abbildung 14: Gesamthäufigkeit der 30 häufigsten Kodas in der Schalengesamtmenge



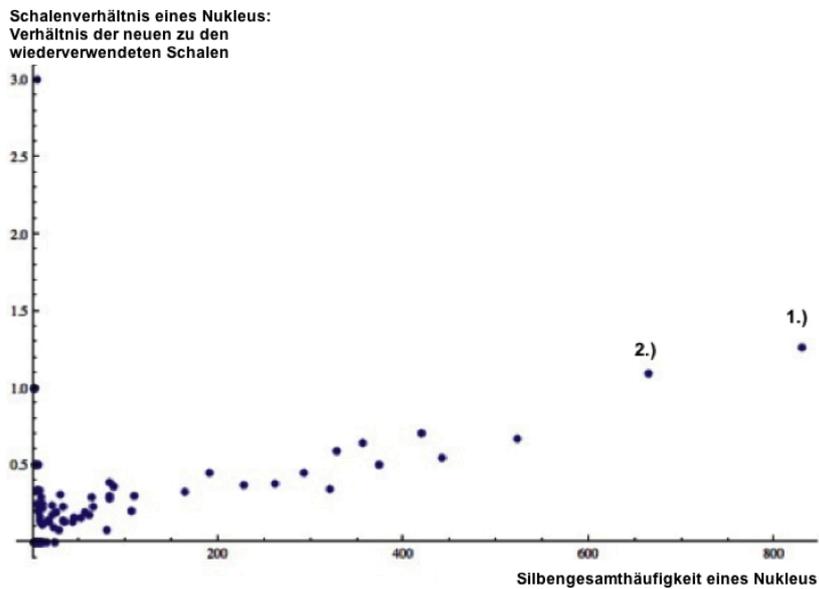
Die Rangfolge der 30 häufigsten Kodas in der Menge aller Schalen.

Die Abb. 14 zeigt die Anzahl der Köpfe, mit denen sich eine bestimmte Koda verbindet: Beispielsweise "kombiniert" sich die Koda /n/ mit 99 unterschiedlichen Köpfen zu eben so vielen Schalen, während dies für die Koda /ɤ/ für nur 34 unterschiedliche Köpfe der Fall ist. Abb. 13 und Abb. 14 zeigen, dass die Mehrheit aller Schalen des Annalena-Korpus von einer kleinen Mengen von Köpfen oder Kodas gebildet wird: Die dreißig häufigsten Köpfe in der Menge aller Schalen des Gesamtkorpus haben mit einer Gesamthäufigkeit von 1450 einen Anteil an der Schalengesamthäufigkeit 2154 von 67,38 %: Die dreißig häufigsten Kodas in der Menge aller Schalen des Gesamtkorpus haben mit einer Gesamthäufigkeit von 1301 an der Gesamthäufigkeit von 2152 aller Schalen einen Anteil von 60,46%.

7.3.6.2. Relation aller Silbenkerne in ihrem Verhältnis von neuen zu bereits verwendeten Schalen

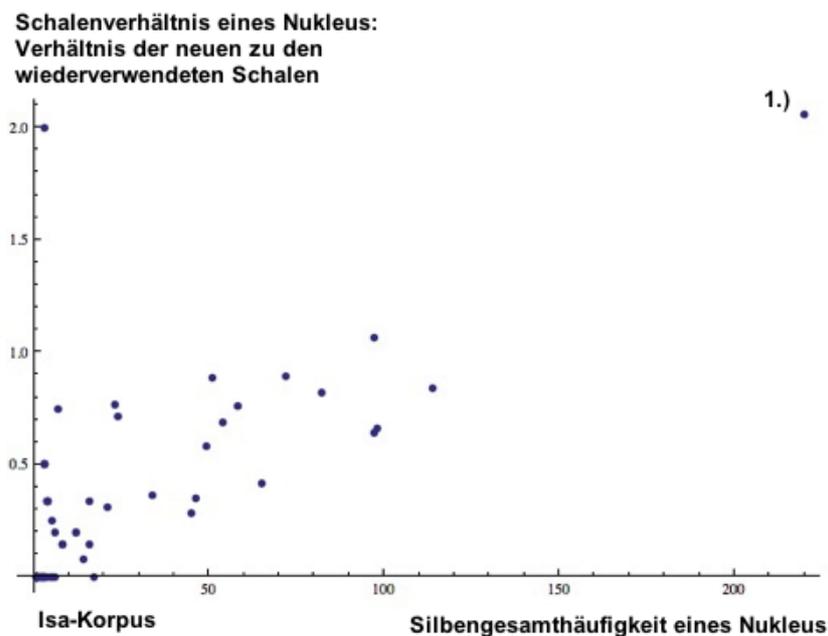
Für jeden einzelnen Silbenkern soll das Verhältnis aus der Anzahl aller seiner von anderen Silbenkernen bereits früher verwendeten zu der Anzahl aller seiner neuen Schalen für das Gesamtkorpus betrachtet werden, welches mit dem Begriff des *Schalenverhältnisses* eines Silbenkerns benannt werden soll. Der häufigste Silbenkern, das /a/ (Gesamthäufigkeit 830 Silben), hat ein Schalenverhältnis von 1,262, das durch den mit "1.)" gekennzeichneten Punkt der Abb. 15 dargestellt wird. Der zweithäufigste Silbenkern, das /ɪ/ (Gesamthäufigkeit 665 Silben), hat ein Schalenverhältnis von 1.098, dargestellt durch den mit "2.)" gekennzeichneten Punkt der Abb. 15.

Abbildung 15: Streudiagramm des Schalenverhältnisses und der Silbengesamthäufigkeit aller Nuklei (Annalena-Korpus)



Auch für das Isa-Korpus hat der häufigste Silbenkern, das /a/ (Silbengesamthäufigkeit 220), das größte⁷⁰ Schalenverhältnis mit einem Wert von 2.055 (mit "1.)" gekennzeichneten Punkt in Abb. 16).

Abbildung 16: Streudiagramm des Schalenverhältnisses und der Silbengesamthäufigkeit aller Nuklei (Isa-Korpus)



⁷⁰ abgesehen von einigen sehr seltenen

Zusammenfassend ist also festzuhalten, dass der jeweils häufigste Silbenkern in beiden Korpora das größte⁷¹ Schalenverhältnis hat.

7.3.7. Statisch-dynamische Verteilungen

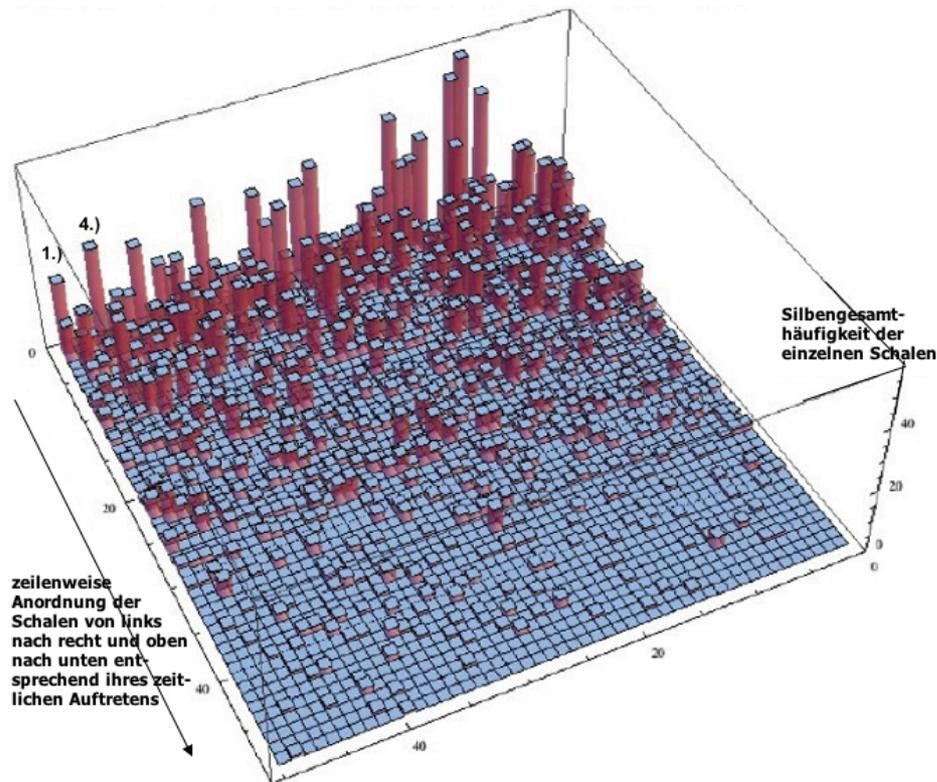
Neben statischen und dynamischen Abfragen gibt es auch die Kombination der beiden, wenn sowohl in der Analyse die Anzahl aller Silben des Korpus mit einer bestimmten Eigenschaft als auch Silbenteilzeiträume enthalten sind.

7.3.7.1. Quantitative und zeitliche Relationen in der Silbenentwicklung

Aus der Silbengesamthäufigkeit einer Schale kann auf den Zeitpunkt ihres ersten Auftretens im Silbenbildungsprozess zurückgeschlossen werden: Je häufiger eine Schale in der Silbenmenge ist, umso früher muss sie tendenziell im Verlauf des Gesamtprozesses der Silbenbildung zum ersten Mal auftreten. Dies zeigt Abb. 17 für die Schalen, die entsprechend ihres zeitlichen Auftretens zeilenweise von links nach rechts und von oben nach unten angeordnet sind, und für die jeweils durch die Höhe der Säule ihre Gesamthäufigkeit im Gesamtzeitraum dargestellt wird. Die Höhen der mit "1.)" bzw. "4.)" gekennzeichneten Säulen stellen die Silbengesamthäufigkeiten der zuerst bzw. als viertes im Zeitverlauf auftretenden unterschiedlichen Schalen dar. Die Säulen dieser ersten und vierten Schale befinden sich in der ersten Zeile des Grundtableaus, auf dem alle unterschiedlichen Schalen des Gesamtkorpus, wie zuvor beschrieben, entsprechend ihres zeitlichen Auftretens angeordnet sind. In weiter unten gelegenen Zeilen findet man keine Säulen mehr von der Höhe der Säulen "1.)" und "2.)". Dies bestätigt die obige Feststellung, dass eine Schale tendenziell umso früher im Verlauf des Gesamtprozesses der Silbenbildung zum ersten Mal auftreten muss, je öfter sie in der Menge aller Silben des Gesamtkorpus vorkommt.

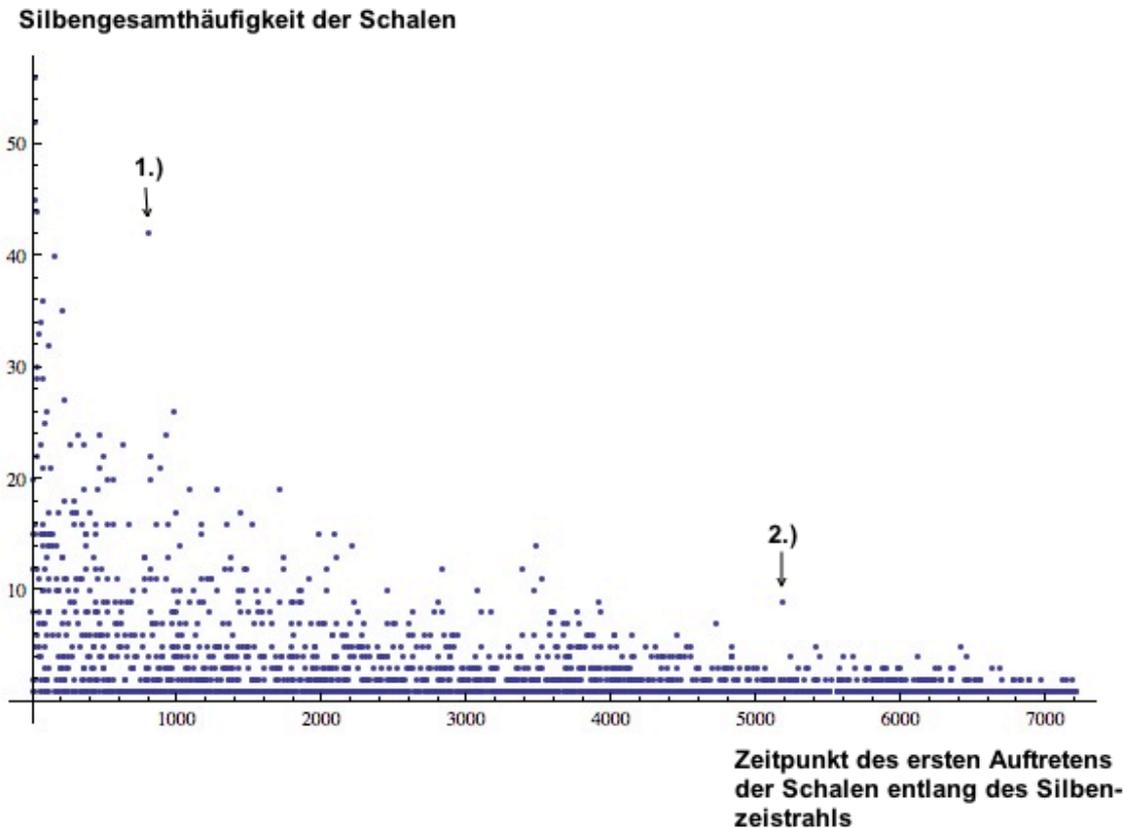
⁷¹ abgesehen von einigen sehr seltenen

Abbildung 17: Zeitliche Anordnung aller Schalen nach dem Zeitpunkt ihres ersten Auftretens und Darstellung ihrer Silbengesamthäufigkeiten



Auf andere Weise wird dieser Sachverhalt in der Abb. 18 verdeutlicht: Mit fortschreitender Entfernung vom Zeitpunkt des ersten Auftretens sinkt auch die Häufigkeit der Verwendung einer Schale im Gesamtprozess. So steht der mit "1.)" und einem hinweisenden Pfeil markierte Punkt in Abb. 18 für eine Schale x, die im gesamten Silbenbildungsprozess zum ersten Mal in der 801-ten Silbe als deren Schale vorkommt. Der Zeitpunkt des ersten Auftretens der Schale x ist die 801-te Silbe und ist an der waagerechten Achse der Abb. 18 gekennzeichnet. Diese Schale x hat zudem eine Silbengesamthäufigkeit von 42 unterschiedlichen Silben, d.h. sie kommt im gesamten Annalena-Korpus in 42 unterschiedlichen Silben vor. Die Silbengesamthäufigkeit von Schalen wird in Abb. 18 durch die senkrechte Achse des Koordinatensystems angezeigt. Der mit "2.)" markierte Punkt der folgenden Abbildung steht hingegen für eine Schale y, die gegenüber der Schale "1.)" erst in der 5177-ten Silbe zum ersten Mal im Silbenbildungsprozess in einer Silbe vorkommt. Entsprechend der obigen Aussage zur Zeit-Silben-Quantitäts-Relation ist ihre Silbengesamthäufigkeit (9) auch erheblich kleiner als die durch "1.)" repräsentierte. Mit fortschreitender Entfernung vom Zeitpunkt des ersten Auftretens einer Schale sinkt tendenziell auch die Häufigkeit ihrer Verwendung im Gesamtprozess.

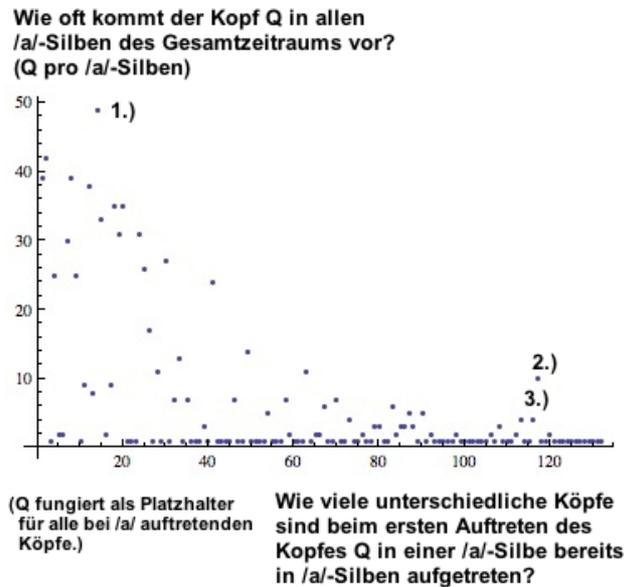
Abbildung 18: Streudiagramm des Zeitpunktes des ersten Auftretens und der Silbengesamthäufigkeit von Schalen



Diese Zeit-Mengen-Relation trifft auch für das Wachstum der einzelnen Köpfe und Kudas bei den einzelnen Kernen zu: Die Abb. 19 zeigt für jeden unterschiedlichen Kopf, der in Silben mit dem Vokal /a/ als Silbenkern vorkommt (kurz: /a/-Silben) diese Menge seiner /a/-Silben im Gesamtzeitraum und setzt sie zu der Anzahl der Köpfe in Beziehung, die zum Zeitpunkt seines ersten Auftretens in einer /a/-Silbe bereits in allen /a/-Silben aufgetreten sind. Auch hier zeigt sich wiederum die Beziehung zwischen der Häufigkeit der Köpfe und dem Zeitpunkt ihres ersten Auftretens: Je später sie in /a/-Silben erscheinen, umso niedriger ist tendenziell ihre Gesamthäufigkeit. So tritt der in Abb. 19 mit "1.)" gekennzeichnete Kopf als 14.-ter Kopf in allen /a/-Silben des Korpus auf, die als im Zeitverlauf zeitlich entsprechend der Reihenfolge ihres Auftretens angeordnet gedacht werden, während der mit "2.)" gekennzeichnete Kopf erst als 117-ter Kopf zum ersten Mal in allen /a/-Silben vorkommt und mit einer /a/-Silbengesamthäufigkeit von 10 erheblich seltener ist als der mit "1.)" gekennzeichnete Kopf (/a/-Silbengesamthäufigkeit von 49). Der mit "3.)" gekennzeichnete Kopf tritt bereits als 116-ter Kopf in /a/-Silben auf, hat aber mit 4 eine geringere /a/-Silbengesamthäufigkeit als der mit "2.)" gekennzeichnete Kopf.

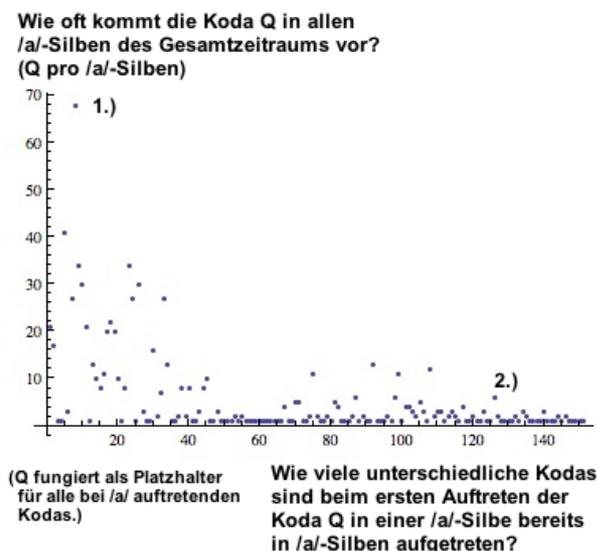
Dies zeigt zudem, dass die formulierte Zeit-Mengen-Relation nicht absolut, sondern nur tendenziell gilt, weil sie vereinzelt durchbrochen werden kann.

Abbildung 19: Streudiagramm der /a/-Silbengesamthäufigkeit eines Kopfes und des Zeitpunkts des ersten Auftretens beim Nukleus /a/



Auch für die Kodas nimmt die Gesamthäufigkeit in /a/-Silben tendenziell ab, je später sie zum ersten Mal in einer /a/-Silbe auftreten. So entspricht dem mit "1.)" gekennzeichneten Punkt in Abb. 20 die Koda, die in allen /a/-Silben als 8-te auftritt und eine /a/-Silbengesamthäufigkeit von 68 hat, während durch den mit "2.)" gekennzeichneten Punkt die Koda repräsentiert wird, die in allen /a/-Silben als 126-te auftritt und eine /a/-Silbengesamthäufigkeit von 6 hat.

Abbildung 20: Streudiagramm der /a/-Silbengesamthäufigkeit einer Koda und des Zeitpunktes ihres ersten Auftretens beim Nukleus /a/



Diese beispielhaft an dem Silbenkern /a/ demonstrierte Zeit-Mengen-Relation gilt ebenso für andere Silbenkerne.

7.3.8. Dynamische Verteilungen (nicht-relational)

Bei der Untersuchung der dynamischen Korpusstrukturen werden im Gegensatz zur ausschließlichen Betrachtung der Gesamtkorpusmenge bei der statischen Analyse Silbenteilmengen als Teilzeiträume des Korpusgesamtzeitraums betrachtet. Zusätzlich zu den Mengen der statischen Abfragen, kommt bei den dynamischen Abfragen also die zeitliche Dimension hinzu, die vielfältige neue Abfragen ermöglicht:

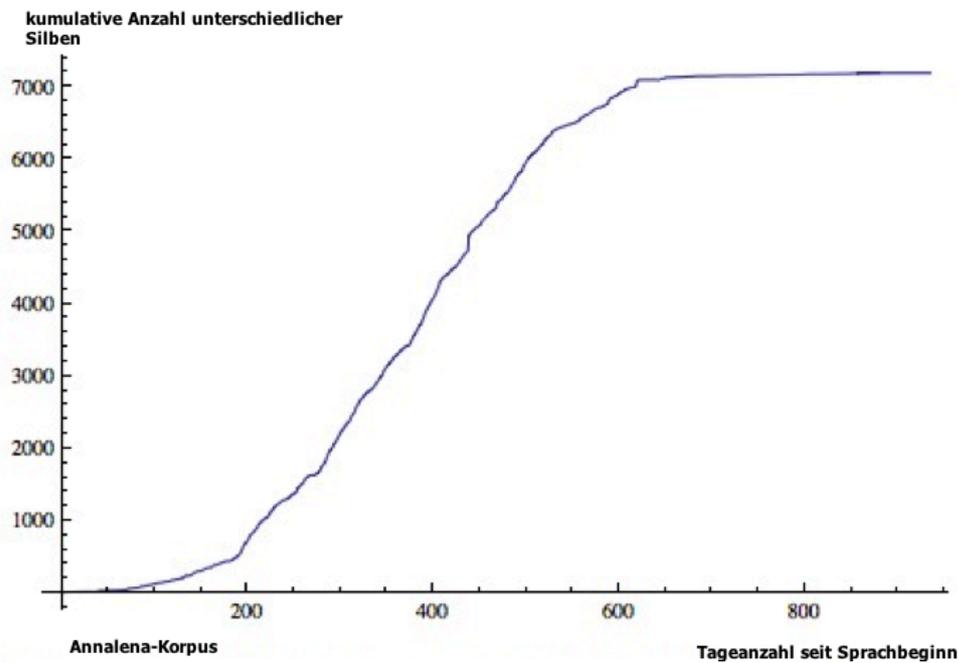
7.3.8.1. Silbenproduktivitätsänderungen im Zeitverlauf

Zu Beginn des Spracherwerbs werden neue Silben in geringerem Umfang gebildet als in späteren Phasen. So äußert Annalena (Elsen 1991) in den ersten 150 Tagen (Zeitraum 1. Tag - 150. - Tag) ihres Spracherwerbs 310 unterschiedliche Silben, während sie in den nächsten 150 Tagen (Zeitraum 151. Tag - 300. Tag) bereits weitere 1914 noch nicht produzierte Silben artikuliert und damit die Menge der vorher produzierten Silben mehr als versechsfacht. Ähnliches gilt für das Kind ISA (Lindner 2009), das während der ersten 151 Tage 242 Silben äußert und in den folgenden 153 Tagen 710 weitere, bisher nicht artikulierten Silben, wodurch die vorher produzierte Silbenmenge fast verdreifacht wird.

Die *Silbenzuwachsgeschwindigkeit*, d.h. die Anzahl neuer Silben pro einer gegebenen Zeiteinheit, nimmt für beide Korpora am Anfang der Entwicklung zu. Die beiden folgenden Schaubilder verdeutlichen diese Vergrößerung der Silbenzuwachsgeschwindigkeit als kumulative Silbenanzahl im Zeitverlauf: In der Abb. 21 sind auf der waagerechten Achse die Anzahl der Tage seit Sprachbeginn abgetragen, d.h. der Wert "1" dieser waagerechten Achse steht für den ersten Tag, an dem das Kind Annalena sein erstes Wort geäußert hat und z.B. die Zahl 96 für den 96-ten Tag einschließlich des ersten Tages seit Sprachbeginn. Nicht an allen Tagen hat das Kind neue Silben (evtl. innerhalb mehrsilbiger Lautformen) geäußert, so dass nicht jedem Tag seit Sprachbeginn auch eine positive Anzahl neuer Silben zugeordnet ist. So hat Annalena am ersten Tag 9 neue Silben geäußert, am zweiten Tag keine und am dritten Tag 15 weitere neue Silben, vom vierten bis zum neunten keine und erst am zehnten Tag wiederum 18 neue Silben, usw. Abb. 21 zeigt die kumulative Häufigkeit neuer Silben im Zeitverlauf. Die jedem Tag zugeordnete bereits erreichte Anzahl unterschiedlicher Silben wird dabei durch die senkrechte Achse in Abb.

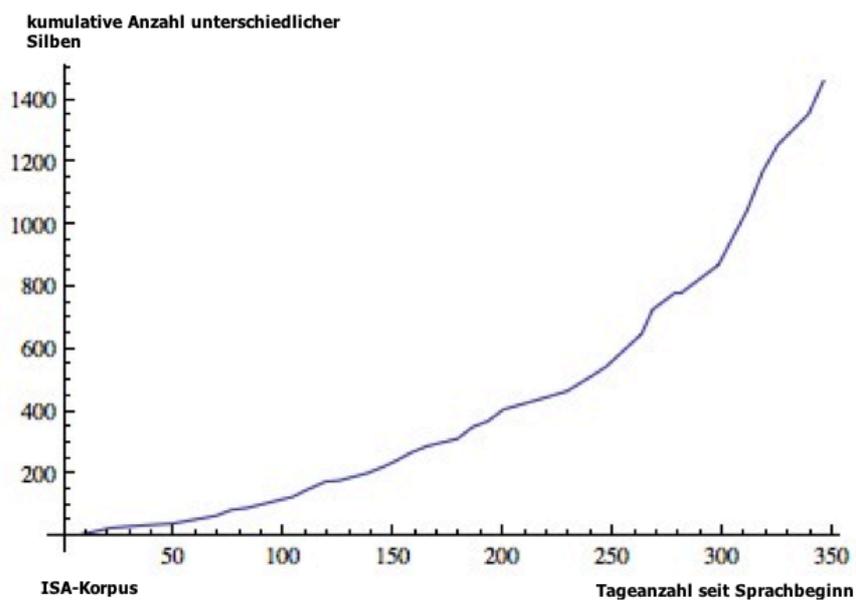
21 dargestellt. Diese einzelnen Werte werden anschließend durch Linien miteinander verbunden.

Abbildung 21: Kumulative Häufigkeit unterschiedlicher Silben (Annalena-Korpus)



Die analoge Verteilung für das Isa-Korpus zeigt die Abb. 22:

Abbildung 22: Kumulative Anzahl unterschiedlicher Silben (Isa-Korpus)



Die oben beschriebene Zunahme des Silbenwachstums wird in Abb. 21 und Abb. 22 an der zunehmenden Steigung des dargestellten Graphen deutlich. Während diese Vergrößerung der Silbenzuwachsgeschwindigkeit für beide Korpora zu beobachten ist, zeigt nur das Annalena-Korpus eine der Zunahme folgende Abnahme, die in der Abb. 21 für das Annalena-Korpus einer "Abflachung" bzw. Verringerung der Steigung des dargestellten Graphen entspricht⁷². Die Tab. 4 zeigt die durchschnittliche Anzahl neuer Silben pro Tag in Zeiträumen von ca. 50 Tagen⁷³ für beide Korpora und bestätigt diese Beobachtung einer der Vergrößerung der Wachstumsgeschwindigkeit neuer Silben folgenden Abnahme des Auftretens neuer Silben pro Zeiteinheit: Während für das Isa-Korpus nur eine Zunahme der Wachstumsgeschwindigkeit von anfänglich 0,96 neuen Silben pro Tag bis auf schließlich 7,37 neuen Silben pro Tag für den Zeitraum der ersten 304 Tage seit Sprachbeginn zu beobachten ist, erreicht die Wachstumsgeschwindigkeit neuer Silben für das Annalena-Korpus im Zeitraum des 400-ten bis 450-ten Tages seit Sprachbeginn mit 20,16 Silben pro Tag ihr Maximum und fällt anschließend wieder bis auf 0,12 neue Silben pro Tag im Zeitraum des 856-ten bis 934-ten Tages seit Sprachbeginn ab.

⁷² Das Isa-Korpus deckt nur eine Zeitspanne von 350 Tagen ab, die seit Sprachbeginn (d.h. dem Zeitpunkt der Äußerung des ersten Wortes) vergangen sind, während das Annalena-Korpus weit mehr als 800 Tage seit Sprachbeginn abdeckt. Dies ist vermutlich die Ursache dafür, dass beim Isa-Korpus noch kein Maximum mit einer folgenden Abnahme der Silbenproduktivität beobachtet werden kann. Zudem erfolgte die Beobachtung des Kindes Isa nicht kontinuierlich gegenüber der beim Kind Annalena, auch wenn dies nicht die Unterschiede der beiden Graphen der Abb. 21 und 22 erklärt, sondern der Umfang des Beobachtungszeitraums, der für das Kind Annalena mehr als doppelt so lang ist (s.o.) und damit eine viel weitergehende Entwicklung des Silbensystems zulässt als der für das Kind Isa abgebildete Zeitraum. (vgl. die Diskussion unter dem Punkt 13.) in Abschnitt 7.3.10.)

⁷³ Wegen der fehlenden Übereinstimmung dieses gewählten Zeitraums von 50 Tagen mit den tatsächlichen Angaben der Sitzungen werden zum Teil nur Annäherungen an die 50-Tage-Zeiträume erreicht.

Tabelle 6: Wachstumsratenänderungen der Anzahl neuer Silben

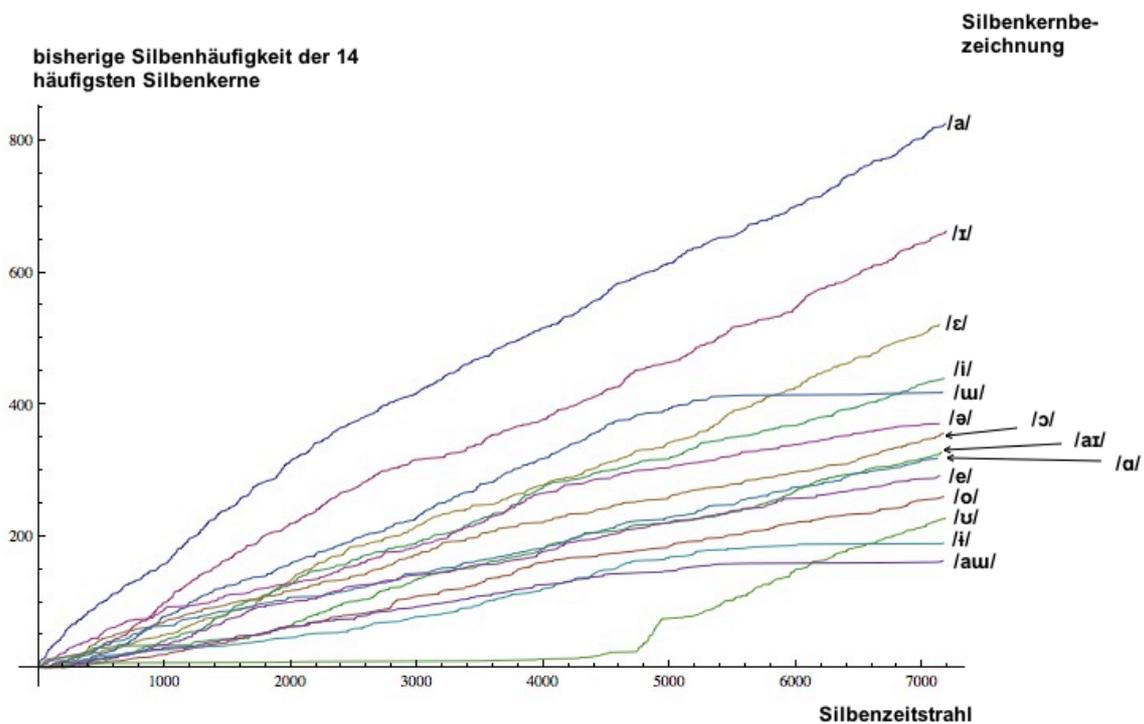
Annalena-Korpus			ISA-Korpus	
Zeitraum in Tagen	durchschnittliche Anzahl Silben		Zeitraum in Tagen	durchschnittliche Anzahl Silben
1-52	0,71		1-69	0,96
52-101	1,92		69-104	1,74
101-150	3,65		104-151	2,45
150-200	8,4		151-200	3,37
200-250	12,98		200-263	3,86
250-300	16,9		263-304	7,37
300-350	18,24			
350-400	18,86			
400-450	20,16			
450-500	17,58			
500-550	10,56			
550-600	8,22			
600-650	4,4			
650-702	0,46			
702-751	0,22			
751-803	0,27			
803-856	0,13			
856-934	0,12			

7.3.8.2. Wachstumsgeschwindigkeiten von Silbenbestandteilen:

Die Abb. 23 zeigt für jeden der 14 häufigsten Silbenkerne des Annalena-Korpus die kumulative Berechnung aller seiner Vorkommnisse (Addition einer 1) und aller seiner Nicht-Vorkommnisse (Addition einer 0) in allen unterschiedlichen Silben des Annalena-Korpus im Zeitverlauf, der durch den Silbenzeitstrahl repräsentiert wird (s.o.). Durch diese kumulative Berechnung erhält man für einen Silbenkern x für jeden Silbenzeitpunkt die Anzahl aller bis zu diesem Zeitpunkt aufgetretenen unterschiedlichen Silben, deren Silbenkern x ist, die auch mit *kumulative Silbenhäufigkeit* von x oder *bisherige Silbenhäufigkeit* von x bezeichnet wird. So zeigt der oben verlaufende Graph der Abb. 23 die kumulative Häufigkeit des Silbenkerns /a/: Bei Erreichen des Silbenzeitpunkts 1500, d.h.

des Auftretens der 1500-ten Silbe, die sich von allen vor ihr aufgetretenen Silben unterscheidet, weist der Silbenkern /a/ eine kumulative Silbenhäufigkeit von 241 auf oder zum Silbenzeitpunkt 3400 (d.h. dem Vorliegen der 3400-ten Silbe), eine kumulative Silbenhäufigkeit von 461.

Abbildung 23: Kumulative Häufigkeiten aller Nuklei

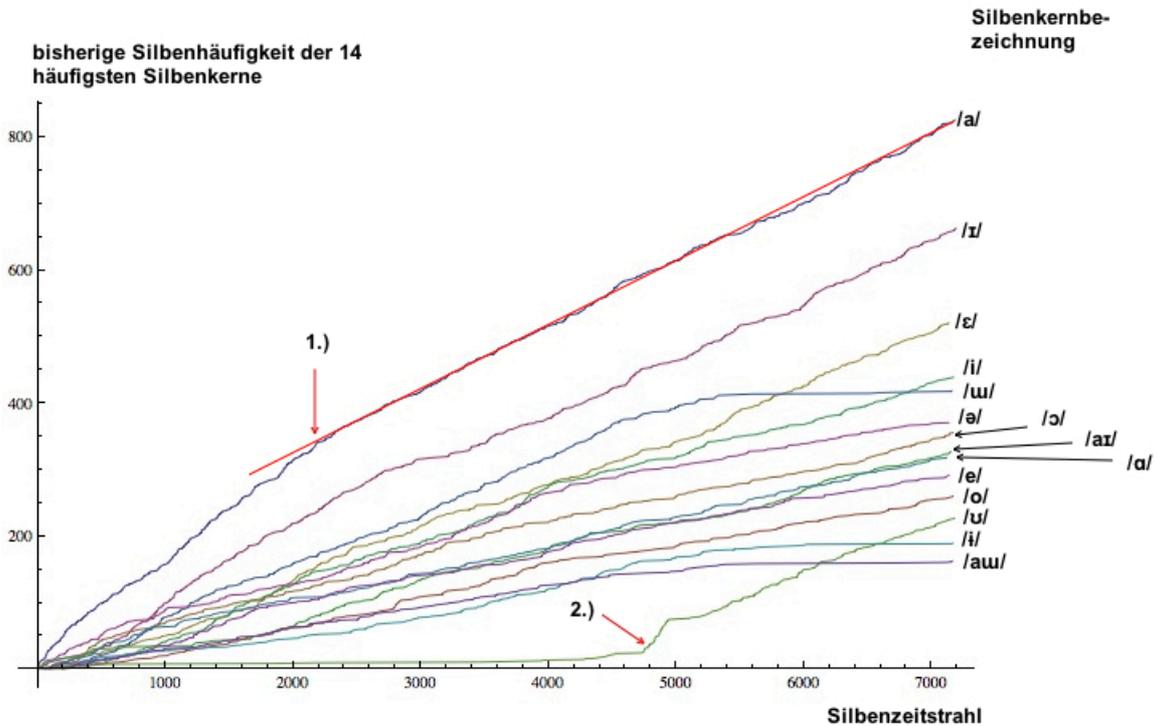


Die Abb. 24 ist mit der vorhergehenden identisch, nur dass per Hand zusätzliche Markierungen vorgenommen worden sind. Die Markierung "1.)" soll einen auffälligen Wechsel in der durchschnittlichen Zuwachsgeschwindigkeit des kumulativen Häufigkeitsgraphen des Silbenkerns /a/ anzeigen: Anfänglich ist diese größer als von dem Zeitpunkt ab, ab dem die per Hand eingezeichnete rote Linie eine gute Näherung⁷⁴ für den Graphen der kumulativen Häufigkeit des Nukleus /a/ ist. Dieser Wechsel tritt also um die 2200. Silbe auf. Zu diesem Zeitpunkt schwächt sich die globale Wachstumsgeschwindigkeit des /a/ ab und verbleibt bis zum Ende in dieser geringeren, jedoch zeitweise leicht variierenden Wachstumsgeschwindigkeit (Dies ist in der Abb. 24 mit "1.)" markiert.). Nimmt bei einem solchen Wechsel wie hier beim Silbenkern /a/ um die 2200-te Silbe die Wachstumsgeschwindigkeit ab, so bedeutet dies, dass der Silbenkern weniger an der Bildung neuer Silben beteiligt ist. Vergleichbare Wechsel sind auch für andere Silbenkerne zu beobachten, u.a. den zweithäufigsten Silbenkern /ɪ/ des Annalena-Korpus. Insbesondere sind neu

⁷⁴ bzw. besser als annähernd deckungsgleich mit dem Graphen der kumulativen Silbenhäufigkeit des Silbenkerns /a/ visuell wahrgenommen wird als zu einem früheren Zeitpunkt

oder sehr stark in Erscheinung tretende Silbenbestandteile (mit "2.") in Abb. 24 markiert), u.a. für den 14.-häufigsten Silbenkern des Annalena-Korpus, das /ɔ/, ca. um die 4800-te Silbe des Silbenzeitstrahls zu beobachten.

Abbildung 24: Kumulative Häufigkeiten aller Nuklei (mit Kennzeichnungen)

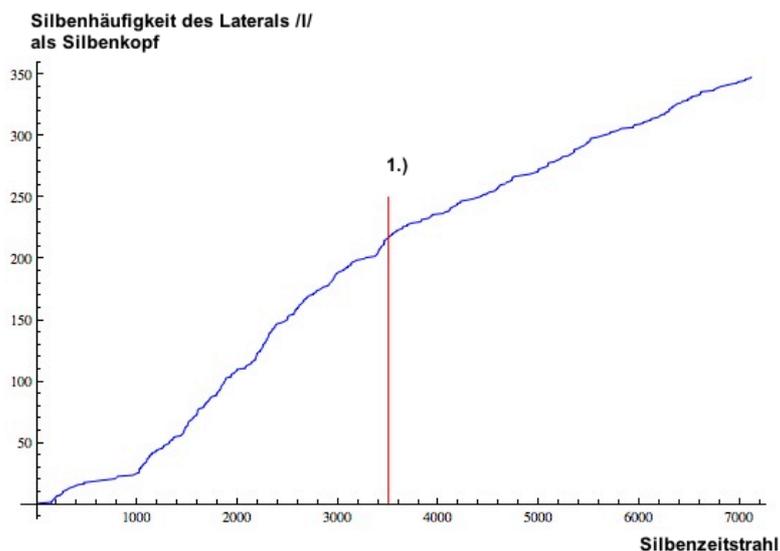


Die Kerne wachsen also phasenweise gleichmäßig. Diese Phasen werden als solche verstanden, die wie im Graphen der kumulativen Silbenhäufigkeit des Silbenkerns /a/ durch die annähernde Deckungsgleichheit desselben mit der roten Linie bzw. abschnittswisen Geraden (s.o.) festgelegt sind. Statt dieses phasenmäßig gleichmäßigen Auftretens der einzelnen Silbenkerne wären auch andere Verläufe denkbar, beispielsweise klar getrennte Zeiträume, in denen einzelne Silbenkerne fast ausschließlich auftreten, während andere nicht oder kaum an der Silbenbildung beteiligt sind. Diese hypothetische Alternative ist aber in dieser globalen, weil auf den Gesamtzeitraum bezogenen Betrachtung, nicht sichtbar. Stattdessen bedeutet das phasenweise gleichmäßige Wachstum der einzelnen Silbenkerne, das entsprechend der oben besprochenen Weise (s.o.) verstanden wird, dass ein ständiger Wechsel zwischen den unterschiedlichen Kernen stattfindet, ihre Aktivität bzw. ihr Auftreten in den Silben im Zeitverlauf nicht zeitlich klar voneinander getrennt ist, sondern sich durchmischt.

Derartige Wechsel in der durchschnittlichen Wachstumsgeschwindigkeit sind auch bei vielen Silbenrändern zu beobachten, beispielsweise beim Lateral /l/ als Kopf: Die Abb. 25 zeigt die kumulative Silbenhäufigkeit des Laterals /l/ als Kopf in allen unterschiedlichen

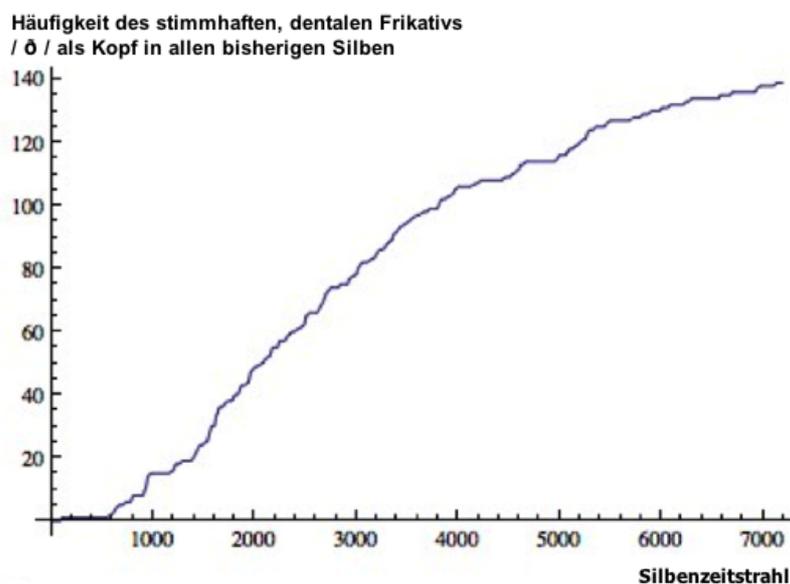
Silben des Annalena-Korpus. Der Wert des auffälligen Wechsels in der Wachstumsgeschwindigkeit des Auftretens neuer Silben mit dem Kopf /l/ ungefähr zum Zeitpunkt der 2500-ten Silbe ist in der folgenden Abbildung durch eine mit "1.)" gekennzeichnete rote senkrechte Linie markiert:

Abbildung 25: Kumulative Häufigkeit des Laterals /l/



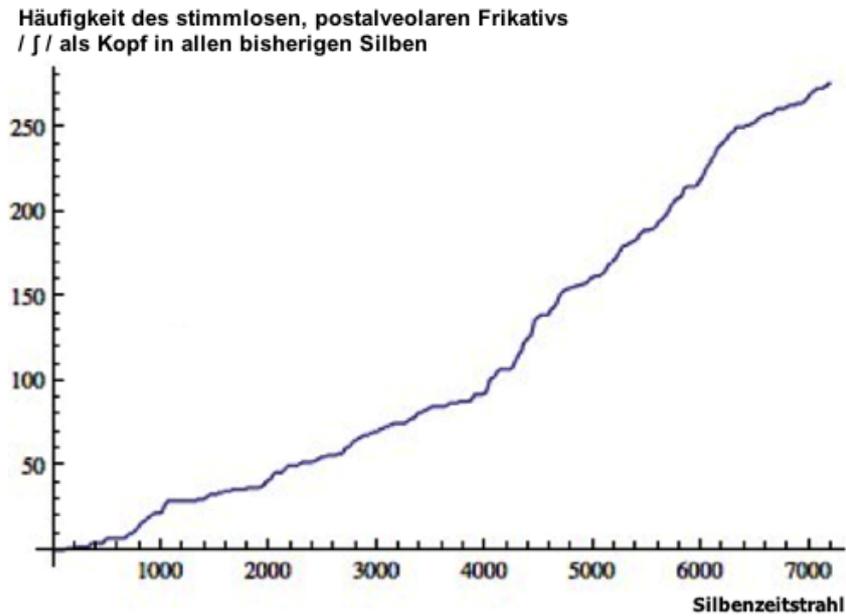
Es gibt auch Paare von Silbenkonstituenten, für die Zeitpunkte ungefähr übereinstimmen, an denen sich die Wachstumsgeschwindigkeiten ihrer kumulativen Silbenhäufigkeiten auffällig ändern: Für den im Englischen vorkommenden, stimmhaften, dentalen Frikativ /ð/ nimmt ab der 4000-ten Silbe die Geschwindigkeit des Wachstums seiner kumulativen Silbenhäufigkeit merklich ab, wie in Abb. 26 gezeigt ist.

Abbildung 26: Kumulative Häufigkeit des Kopfs /ð/



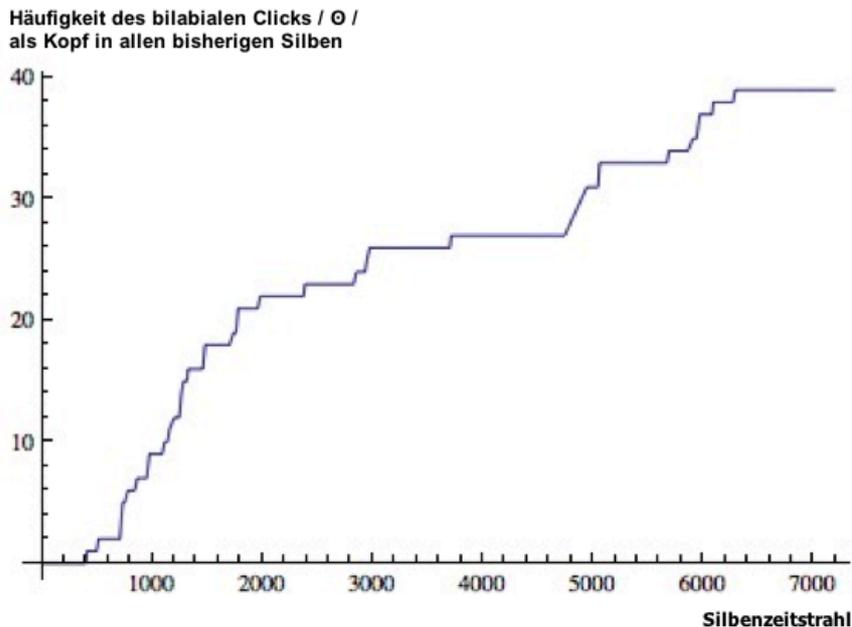
Ungefähr gleichzeitig verstärkt sich das Wachstum des nativen Frikativs /ʃ/.

Abbildung 27: Kumulative Häufigkeit des Kopfs /ʃ/



Ein noch stärkerer und früherer Wachstumseinbruch als beim oben gezeigten nicht-nativen Frikativ /ð/ ist beim nicht-nativen für das Deutsche exotischen bilabialen Click /ǀ/ zu beobachten:

Abbildung 28: Kumulative Häufigkeit des Kopfs /ǀ/

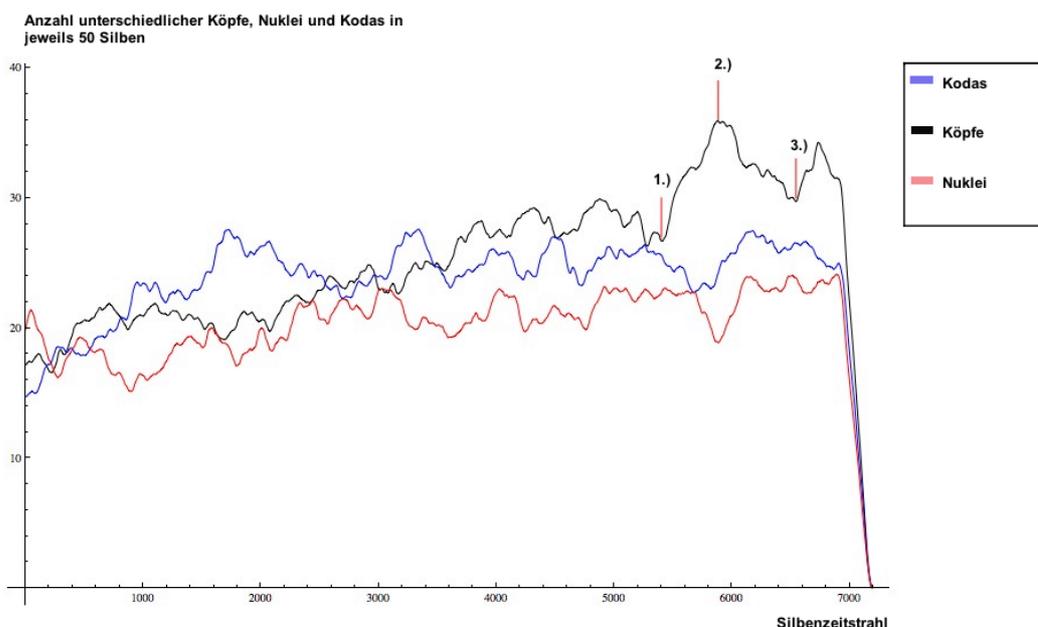


7.3.8.3. Anzahl unterschiedlicher Köpfe, Nuklei und Kodal in gleichgroßen Silbennengen im Zeitverlauf

Wie oben bereits in der Beschreibung des methodischen Vorgehens erwähnt, ist während der Untersuchungen der Verteilungen in den beiden Korpora klar geworden, dass die Betrachtung einzelner Silbenkonstituenten nicht primär im Vordergrund stehen kann, sondern dass vorrangig globale Maße das Problem des Auffindens des gesuchten Abstandsmaßes der Silbenbildung lösen können. Die in diesem Abschnitt beschriebene Untersuchung ermittelt ein solches globales Maß, weil nicht einzelne konkrete Konstituenten gemessen werden, sondern jeweils die Gruppierung aller Silbenköpfe, Silbennuklei und Silbenkodal. Die Abb. 29 zeigt die Anzahl unterschiedlicher Köpfe (schwarz), Nuklei (rot) und Kodal (blau) in jeweils 50 Silben (mit moving average 240). Genauer gesagt wurde für jeden Silbenzeitpunkt⁷⁵ t des Silbenzeitstrahls der folgende Wert berechnet:

Bezogen auf die Menge aus der Silbe dieses Silbenzeitpunkts und den 49 dieser Silbe folgenden Silben des Silbenzeitstrahls wurde die Anzahl x_t aller unterschiedlichen Köpfe (Kodal, Kerne) bestimmt. Anschließend wurde ausgehend von jedem Zeitpunkt der arithmetische Mittelwert von x_t und den diesem Zeitpunkt folgenden analogen 239 Werten x_j gebildet, d.h. der moving average 240 - Wert über den Wert x_t und die ihm folgenden analogen 239 Werte $x_{j+1}, \dots, x_{j+239}$.

Abbildung 29: Anzahl unterschiedlicher Köpfe, Nuklei und Kodal in je 50 Silben im Zeitverlauf (Annalena-Korpus)



⁷⁵ der ja einer Silbe der Menge aller Silben des Korpus entspricht. Siehe dazu die oben in Abschnitt 7.3.3. gegebene Begriffserklärung.

Für jede der drei Konstituentenarten, d.h. für die Köpfe (schwarzer Graph), Kodas (blauer Graph) und die Nuklei (roter Graph), ist der Verlauf der zuvor beschriebenen Anzahl nicht geradlinig, sondern schwankt im Zeitverlauf. Dies zeigt sich an wiederholten Auf- und Abwärtsbewegungen. Zunächst soll die Bedeutung der Einzelverläufe am Beispiel der Konstituentenart der Köpfe beschrieben werden. Die im Folgenden hinsichtlich des Einzelverlaufs vorgenommene Beschreibung gilt aber analog für die Silbenkonstituentenarten der Kodas und Nuklei. Im Anschluss an diese Beschreibung der Einzelverläufe soll auf die Unterschiede zwischen den drei Graphen (rot, blau, schwarz) der Konstituentenarten der Nuklei, Kodas und Köpfe eingegangen werden. Durch die Abfrage der eingangs beschriebenen Werte, d.h. wie viele unterschiedliche Köpfe in einem Silbenzeitraum von 50 Silben auftreten (und der Mittelung von je 240 solcher direkt aufeinanderfolgender für einen Werte (s.o.)), wird die Tendenz deutlich, inwieweit Köpfe in je 50 Silben wiederholt auftreten oder anders gesagt, wie groß das Ausmaß der in der betrachteten Menge von 50 Silben einmalig auftretenden Köpfe ist. So treten im Schnitt (moving average 240) im Zeitraum der 5400-5449-ten Silbe 26.72 unterschiedliche Köpfe auf und dieser Wert ist dem Zeitpunkt der 5400-ten Silbe zugeordnet, gekennzeichnet durch "1.)" und eine senkrechte rote Linie in der Abb. 29 (s.o.), während im Zeitraum der 5881-5930-ten Silbe 35.92 unterschiedliche Köpfe vorkommen, als Wert zugeordnet zum Zeitpunkt 5881 (s.o.) und gekennzeichnet durch eine "2.)" und eine senkrechte rote Linie in Abb. 29. Damit kommen im Zeitraum 5881-5930 ("2.")-Kennzeichnung) im Schnitt mehr unterschiedliche Köpfe als im Zeitraum 5400-5449 ("1.")-Markierung) vor. Es findet also von dem Zeitpunkt 5400 ("1.") zum Zeitpunkt 5881 ("2.") ein tendenzieller Anstieg der Anzahl unterschiedlicher Köpfe in Zeiträumen von je 50 Silben statt. Ein tendenzieller Anstieg der Anzahl unterschiedlicher Köpfe in gleich großen Silbenzeiträumen von je 50 Silben hat mehrere Bedeutungen: Erstens findet eine Streuung in dem Sinne statt, dass die Menge von 50 Silben sich nach dem Anstieg auf eine größere Anzahl unterschiedlicher Köpfe als vor dem Anstieg verteilt. Zweitens werden die Silben stärker über die Köpfe differenziert als dies für sie vor dem Anstieg der Fall war. Drittens kann gesagt werden, dass die Mengen von 50 Silben nach dem Anstieg mehr aus unterschiedlichen Köpfen als vor dem Anstieg gebildet werden. Im Gegensatz zum vorhergehenden Vergleich der Werte für die Zeitpunkte 5400 und 5881 findet für den Wert von durchschnittlich 35.92 unterschiedlichen Köpfen im Zeitraum 5881-5930, der dem Zeitpunkt 5881 zugeordnet ist, im Vergleich zu dem dem Silbenzeitpunkt 6543 zugeordneten Wert von

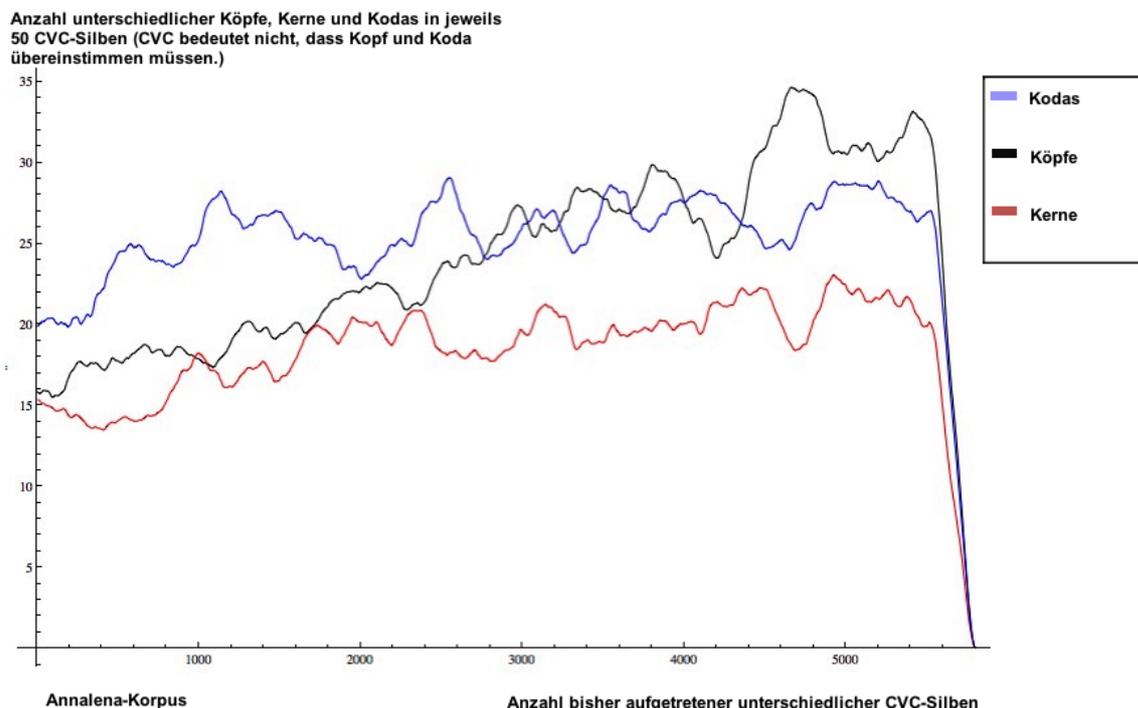
durchschnittlich 29.71 unterschiedlichen Köpfen im Zeitraum 6543-6592 ein tendenzieller Abstieg der Anzahl unterschiedlicher Köpfe in Zeiträumen von je 50 Silben statt.

In diesem Fall des Anstiegs können die folgenden zum Anstieg entgegengesetzten Aussagen getroffen werden: Erstens findet eine Konzentration in dem Sinne statt, dass die Menge von 50 Silben zum Zeitpunkt nach dem Abstieg sich auf weniger unterschiedliche Köpfe als vor dem Abstieg verteilt. Zweitens werden die Silben weniger über die Köpfe differenziert als dies für sie vor dem Abstieg der Fall war. Drittens kann gesagt werden, dass die Mengen von 50 Silben nach dem Abstieg von weniger unterschiedlichen Köpfen als vor dem Anstieg gebildet werden, also eher von gleichen Köpfen und damit die Wahrscheinlichkeit der Wiederholung eines Kopfes nach dem Abstieg größer ist als vor dem Abstieg. Wie bereits erwähnt gelten diese Aussagen zur Bedeutung des Einzelverlaufs der Köpfe analog für die Konstituentenarten der Kerne und Kodas.

Nachdem auf die Bedeutung der Einzelverläufe eingegangen worden ist, soll nun das Verhältnis der einzelnen Verläufe (d.h. des roten (Kerne), blauen (Kodas) und schwarzen (Köpfe) Graphen) zueinander besprochen werden: Für alle Paare von Graphen dieser drei Verläufe, d.h. für das Paar aus dem Kopf- (schwarz) und Koda-Graphen (blau), das aus dem Koda- und dem Nukleus- (rot) Graphen und das aus dem Kopf- und Nukleus - Graphen, sind zeitweilig, aber nicht während des Gesamtverlaufs, gegenläufige Bewegungen zu beobachten. D.h. es sind Verhältnisse des Verlauf von je zwei Graphen zu beobachten, bei denen der eine Graph eines dieser Paare ansteigt und der andere abfällt oder das umgekehrte quantitative Verhältnis zutrifft. So fällt der rote Verlauf der Kerne beispielsweise zum Zeitpunkt der 5400-ten Silbe ab, während der der Köpfe ansteigt. Ebenfalls zum gleichen Zeitpunkt steigt der Graph für die Kodas ab, während der für die Köpfe wie bereits erwähnt ansteigt. Besonders auffällig sind diese Gegenläufigkeiten beim Verlauf für das Koda-Kopf- und Nukleus-Kopf-Paar, während die Gegenläufigkeit für das Koda-Nukleus-Paar nicht so stark ausgeprägt zu sein scheint. Diese zeitweiligen Gegenläufigkeiten von je zwei Paaren lassen sich auch so ausdrücken, dass der eine Graph eines Paares stärker differenziert, während der andere dieses Paares gleichzeitig weniger differenziert. Neben einer Gegenläufigkeit weisen die drei Graphen für die Kern-, Kopf- und Koda-Gruppe auch unterschiedliche Niveaus ihrer Verläufe auf: Bis auf einen Anfangszeitraum, in dem der Graph für die Kerne (rot) über dem für die Köpfe liegt und der letztere über dem für die Kodas, verlaufen die Kerne fast durchgehend auf niedrigstem Niveau, wie in der Abb. 29 zu beobachten ist. Für die Kerne gibt es also bis auf den Anfangszeitraum die stärkste Tendenz zur Wiederholung in den jeweils betrachteten Zeit-

räumen von 50 Silben, wie sich aus der zuvor dargelegten Bedeutung der Einzelverläufe hinsichtlich ihrer An- und Abstiege ergibt (s.o.). Für den Kopf-Graph (schwarz) ist zudem ein ansteigender Trend in der Abb. 29 zu beobachten, weil sich gedanklich durch diesen eine von links unten nach rechts oben verlaufende Gerade mit positiver Steigung legen lässt. Diese liegt zunächst unterhalb des Verlaufs des Koda-Graphen (blau), übersteigt diesen aber schließlich und verbleibt bis zum Ende des Silbenzeitstrahls nahezu weitgehend oberhalb desselben. Die Gruppen der Köpfe, Kerne und Kodas unterscheiden sich deshalb entsprechend ihrer Gegenläufigkeit und ihrer unterschiedlichen Niveaus jeweils voneinander hinsichtlich der in der Abb. 29 betrachteten globalen Werte. Es wäre möglich, dass die gegenläufigen Bewegungen in den Graphen der Abb. 29 durch ungleichmäßiges Auftreten von CV- oder VC-Silben hervorgerufen werden. Um dies auszuschließen, sollen sie zusätzlich nur für die zeitlich geordneten Folgen der CVC-Silben⁷⁶, d.h. der Silben mit sowohl einem Kopf als auch einer Koda abgefragt werden.

Abbildung 30: Anzahl unterschiedlicher Köpfe, Nuklei und Kodas in je 50 CVC-Silben im Zeitverlauf (Annalena-Korpus)



Aber auch bei der Betrachtung der isolierten CVC-Silben des Annalena-Korpus bleiben die zuvor beobachteten Merkmale des Verhältnisses der Einzelverläufe wie der Gegenläufigkeit v.a. des Kern-Kopf-Paares und des Kopf-Koda-Paares erhalten ebenso wie der Verlauf der betrachteten drei Graphen auf unterschiedlichen Niveaus, auch wenn sich

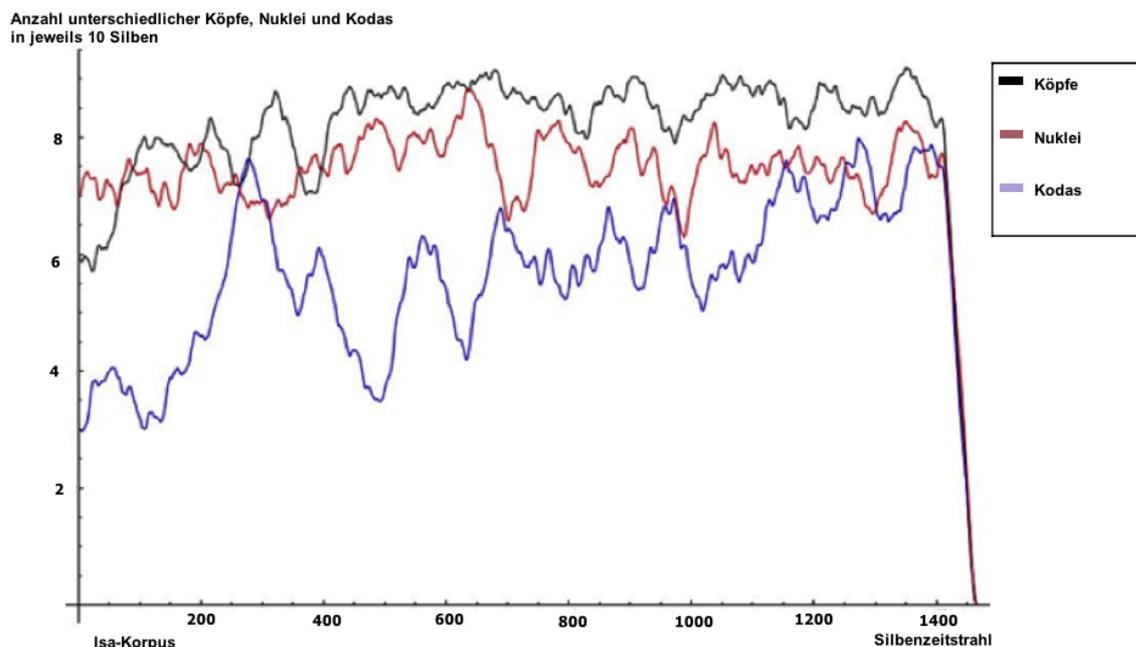
⁷⁶ Dabei bedeutet die Bezeichnung CVC-Silben nicht, dass der Kopf mit der Koda übereinstimmen muss, sondern nur, dass sowohl eine Koda als auch ein Kopf vorhanden ist.

hinsichtlich des Verhältnisses der Niveaus der Einzelverläufe Änderungen ergeben: So liegt der Graph für die Nuklei anfänglich nicht mehr oberhalb der Graphen der beiden anderen Konstituentenarten, sondern unterhalb derselben. Der Grund ist vermutlich darin zu suchen, dass anfänglich verstärkt CV-Silben auftreten, die in Abb. 30 herausfallen. Der schon in Abb. 29 beobachtete ansteigende Trend des Graphen der Kopf-Konstituentenart bleibt hingegen erhalten, auch wenn die im Zeitverlauf stattfindende Trennung des Kopf- und Koda-Graphen nicht mehr so deutlich wie in Abb. 29 ausfällt.

Für das Isa-Korpus wurden entsprechende Kopf- (schwarz), Koda- (blau) und Kern- (rot)-Graphen berechnet, aber mit einer moving average Mittelung für 10 Werte anstatt wie zuvor 240, was durch die erheblichen Unterschiede dieser beiden Korpora hinsichtlich ihrer Gesamtgröße gerechtfertigt scheint. Für diese Graphen können ähnliche Merkmale wie zuvor für das Annalena-Korpus beobachtet werden: Diese drei Graphen verlaufen auch für das Isa-Korpus auf unterschiedlichem Niveau mit dem Unterschied, dass der Koda-Graph (blau) unterhalb des Kern-Graphen (rot) verläuft, aber ansonsten wie zuvor für das Annalena-Korpus der Kern-Graph anfänglich oberhalb des Kopf-Graphen verläuft und sich nach einer kurzen Anfangsphase dieses Verhältnis umkehrt.

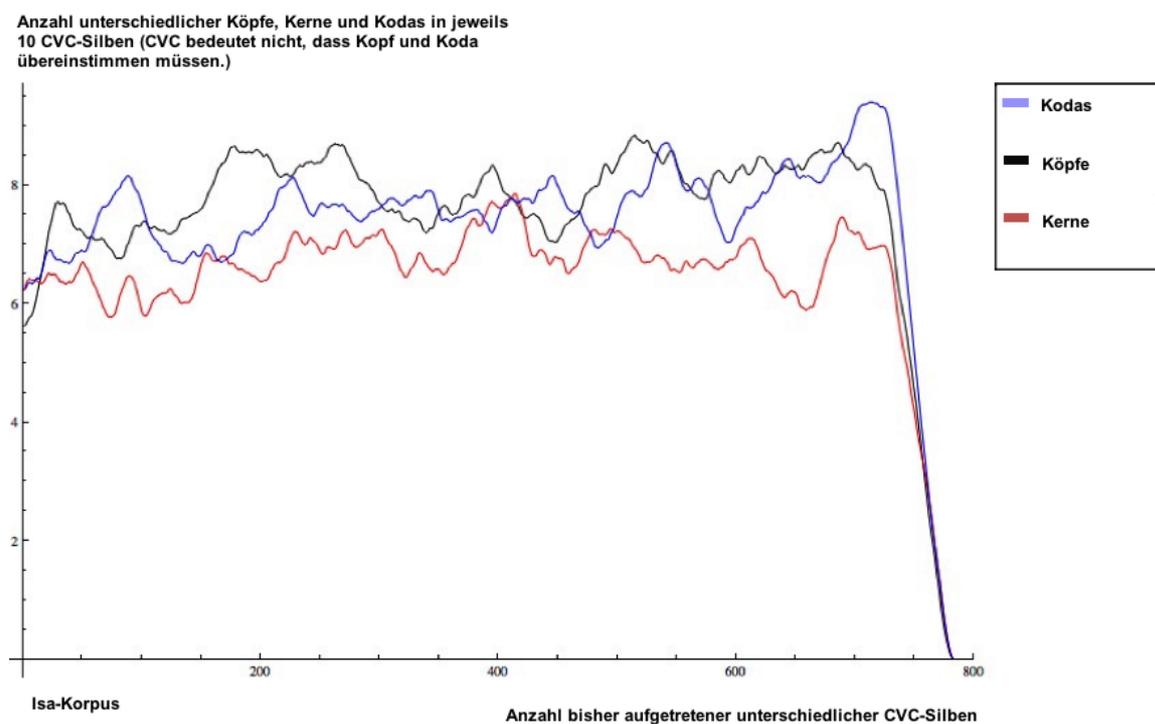
Auch für das ISA-Korpus zeigt sich teilweise dieses gegenläufige Verhalten für die Anzahl unterschiedlicher Köpfe (schwarz), Koda (blau) und Kerne (rot) in jeweils Zeiträumen von 10 Silben (moving average 50).

Abbildung 31: Anzahl unterschiedlicher Köpfe, Nuklei und Koda in je 10 Silben im Zeitverlauf (Isa-Korpus)



Wie zuvor für das Annalena-Korpus bleiben die beobachteten Merkmale des gegenseitigen Verhältnisses der drei Graphen bei ihrer Berechnung ausschließlich für alle CVC-Silben des Isa-Korpus wie in Abb. 32 geschehen erhalten: Die drei Graphen verlaufen auch in diesem Schaubild auf unterschiedlichen Niveaus und das Paar der Kopf-Kern-Graphen sowie das Paar der Kopf-Koda-Graphen zeigt ein zeitweilig gegenläufiges Verhalten.

Abbildung 32: Anzahl unterschiedlicher Köpfe, Nuklei und Koda in je 10 CVC-Silben im Zeitverlauf (Isa-Korpus)



7.3.8.4. Wiederverwendung von Silbenbestandteilen

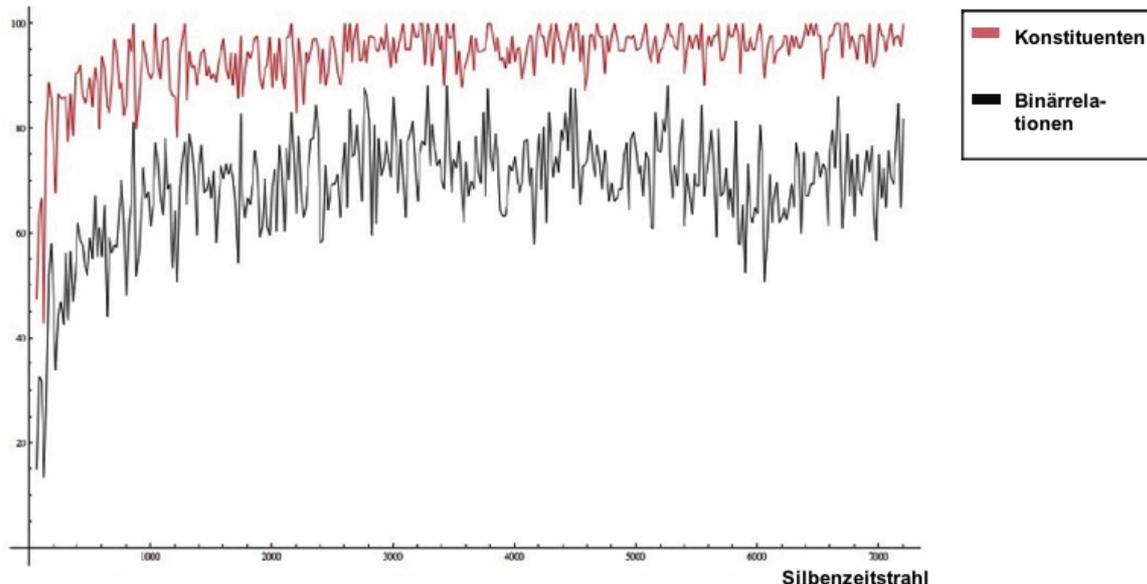
Die Abb. 33 zeigt für jeden Zeitpunkt des Silbenzeitstrahls den Prozentanteil der Konstituenten (rot) und Binärrelationen (schwarz) des diesem Zeitpunkt folgenden 20-Silben-Zeitraums (auch als *Block* bezeichnet), die bereits zu einem früheren Zeitraum in Silben aufgetreten sind und somit wiederverwendet werden⁷⁷. Für die Berechnung der beiden Graphen in der Abb. 33 findet zunächst eine Einteilung des Silbenzeitstrahls in direkt aufeinanderfolgende, überschneidungsfreie 20-Silbenzeiträume statt, die auch als (*Silben-)*Blöcke bezeichnet werden sollen. Daraufhin werden für jeden dieser Blöcke zwei Ver-

⁷⁷ Der Begriff der *Wiederverwendung* ist nicht als Annahme einer vollständigen Identität zwischen Silbenbestandteilen zu verstehen, die in unterschiedlichen Silbenkombinationen auftreten, sondern als Ergebnis einer Schnittmengenbildung, die von einem erwachsenen Sprecher der Zielsprache als starke Ähnlichkeit oder sogar Identität empfunden wird.

hältnisse gebildet: Erstens wird das Verhältnis aus der Anzahl aller Konstituenten, d.h. aller Köpfe, Kodas und Kerne, die bereits in früheren Blöcken aufgetreten sind zu der Anzahl aller Konstituenten diese Blockes in Beziehung gesetzt (und zur Überführung in eine Prozentangabe mit der Zahl 100 multipliziert) (der rote Graph im folgenden Schaubild). Zweitens wird das Verhältnis aus der Anzahl aller Binärrelationen, d.h. aller Schalen, Körper und Reime, die bereits in früheren Blöcken aufgetreten sind zu der Anzahl aller Binärrelationen diese Blockes in Beziehung gesetzt (und zur Überführung in eine Prozentangabe mit der Zahl 100 multipliziert) (der schwarze Graph im folgenden Schaubild). Beide so berechneten Werte werden jeweils am Anfangszeitpunkt des betrachteten Zeitraums abgetragen. Die Werte für die Konstituenten (roter Graph) als auch für die Binärrelationen (schwarzer Graph) liegen bis auf eine kurze Anfangsphase oberhalb der 50%-Grenze. Der Verlauf der Konstituenten liegt sogar bis auf diese Anfangsphase über der 80%-Marke.

Abbildung 33: Anteil bereits aufgetretener Konstituenten und Binärrelationen an 20-Silbenzeiträumen im Zeitverlauf

Der Silbenzeitstrahl wurde in direkt aufeinanderfolgende, überschneidungsfreie Blöcke von 20 Silben eingeteilt:
Welcher Prozentanteil aller Köpfe, Kodas und Nuklei eines Blockes ist t bereits in einem der vorhergehenden Blöcke aufgetreten?
Welcher Prozentanteil der Binärrelationen eines Blockes ist bereits in einem der vorhergehenden Blöcke aufgetreten?



7.3.8.5. Anteil der jeweils von den 3, 10 und 20 häufigsten Kernen gebildeten Silbenmenge im Zeitverlauf:

Die folgenden Abb. 34, 35 und 36 stellen für eine Einteilung des Silbenzeitstrahls in überschneidungsfreie, direkt aufeinanderfolgende Zeiträume von je 50 Silben, den Anteil der drei, fünf bzw. zwanzig hinsichtlich des Gesamtzeitraums häufigsten Silbenkerne an

diesen Zeiträumen dar. Für die Berechnung wurde in den Abb. 34, 35 und 36 der Silbenzeitstrahl in direkt aufeinanderfolgende, überschneidungsfreie Zeiträume von je 50 Silben unterteilt und für jeden dieser Zeiträume wurden analoge Verhältnisse berechnet, jedoch mit unterschiedlicher Anzahl der im Gesamtzeitraum des Annalena-Korpus häufigsten Silbenkerne: Die Abb. 34 zeigt für jeden der erwähnten Zeiträume den Anteil der Anzahl aller Silben des betrachteten Zeitraums, die einen der drei im gesamten Annalena-Korpus häufigsten Nuklei als Silbenkern haben, an der Anzahl aller Silben dieses Zeitraums. Diese Anzahl hat aber für alle Zeiträume des Silbenzeitstrahls jeweils konstant die Größe von 50 Silben.

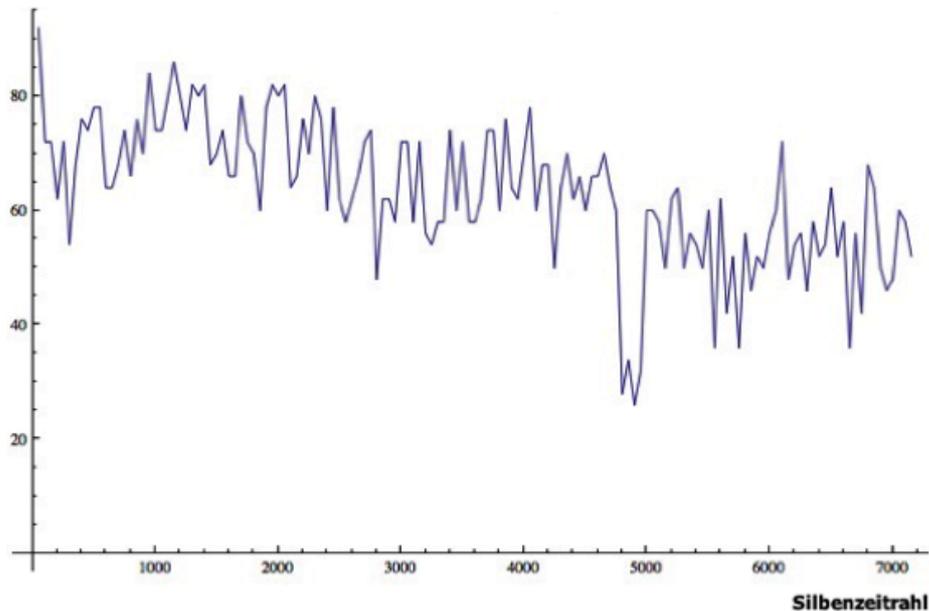
Abbildung 34: Anteil der 3 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben



Die häufigsten drei Silbenkerne im Gesamtzeitraum des Annalena-Korpus sind an der Bildung von 10 bis 50 Prozent aller Silben jedes der 50-Silbenzeiträume beteiligt, d.h. von allen Silben der betrachteten 50-Silbenzeiträume des Silbenzeitstrahls werden zwischen 5 bis 25 aller Silben von den drei häufigsten Silbenkerne im gesamten Annalena-Korpus gebildet. Die Abb. 35 unterscheidet sich von der Abb. 34 dadurch, dass der Anteil der 10 im gesamten Annalena-Korpus häufigsten Silbenkerne berechnet wird: Sie zeigt also für jeden der erwähnten Zeiträume den Anteil der Anzahl aller der Silben des betrachteten Zeitraums, die einen der 10 im gesamten Annalena-Korpus häufigsten Nuklei als Silbenkern haben, an der Anzahl aller 50 Silben dieses Zeitraums.

Abbildung 35: Anteil der 10 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben

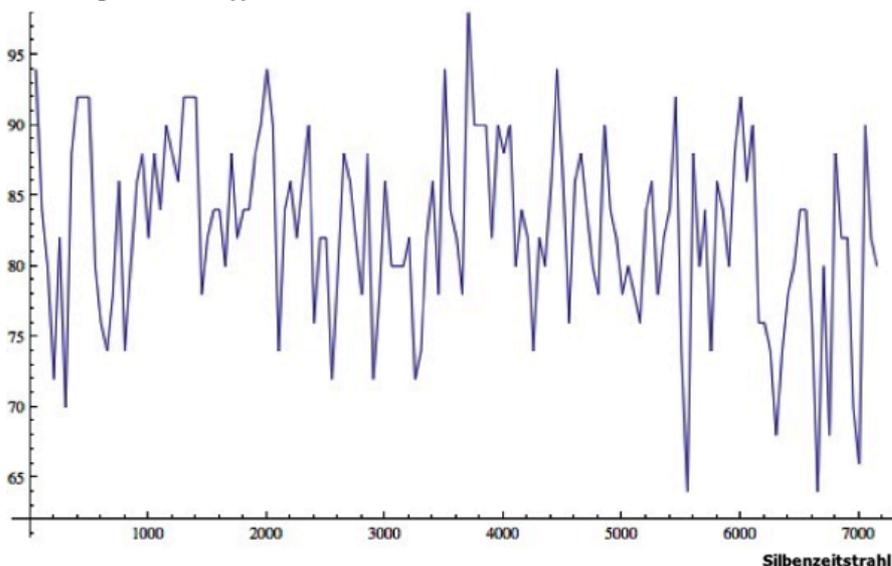
II. Anteil der 10 häufigsten Silbenkerne an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben



Es wird sichtbar, dass die häufigsten 10 Silbenkerne im Gesamtzeitraum an der Bildung von 15 bis 90 Prozent aller Silben beteiligt sind. Die Abb. 36 zeigt entsprechend den Anteil der Anzahl aller der Silben des betrachteten Zeitraums, die einen der 20 im gesamten Annalena-Korpus häufigsten Nuklei als Silbenkern haben, an der Anzahl aller 50 Silben dieses Zeitraums.

Abbildung 36: Anteil der 20 häufigsten Nuklei an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben

III. Anteil der 20 häufigsten Silbenkerne an der Silbenbildung in fortlaufend angeordneten Gruppen von 50 Silben



Die Beteiligung der 20 häufigsten Silbenkerne des gesamten Annalena-Korpus an der Bildung von jeweils 50 Silben des gesamten Annalena-Korpus liegt schließlich im Schnitt bei ca. 80, d.h. im Schnitt werden ca. 40 von 50 Silben in den betrachteten 50-Silben-Zeiträumen des Annalena-Korpus von den 20 im gesamten Annalena-Korpus häufigsten Nuklei gebildet.

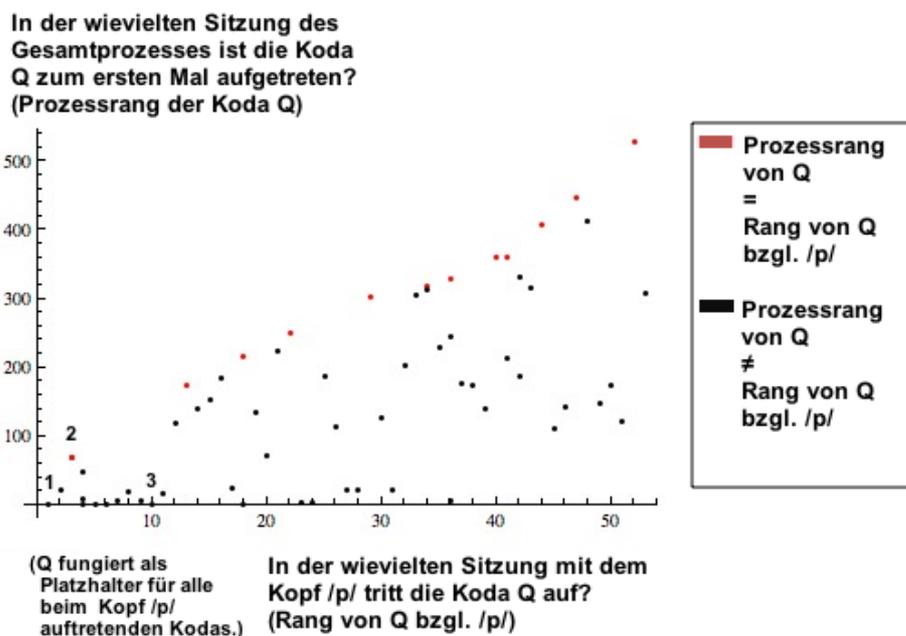
7.3.9. Dynamische Verteilungen (relational)

7.3.9.1. Zeitliches Kombinationsverhalten konkreter Köpfe mit Kodas:

In diesem Abschnitt soll der Frage nachgegangen werden, wie die zeitliche Verteilung aller Kodas im Gesamtsilbenbildungsprozess geregelt ist, mit denen sich ein bestimmter Kopf verbindet. Es wäre denkbar, dass ein bestimmter Kopf in seinem Kombinationsverhalten auf zunächst vor seinem ersten Auftreten vorkommende Kodas zurückgreift, bevor er sich mit neuen Kodas verbindet. Kodas werden dabei hinsichtlich eines bestimmten Kopfes *x* als *neu* bezeichnet, wenn die erste Silbe, in der sie überhaupt während des gesamten Silbenbildungsprozesses auftreten, den Kopf *x* aufweist.

Dass dies nicht der Fall ist, zeigt die Abb. 37 für den Kopf /p/:

Abbildung 37: Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /p/:



Für das Verständnis der Abb. 37 und der damit verbundenen Überprüfung der zuvor geäußerten Vermutung ist die begriffliche Unterscheidung des *Prozessrangs* einer Koda und des *Rangs* einer Koda *bzgl.* eines bestimmten Kopfes hilfreich. Diese beiden Begriffe beziehen sich auf ein konkretes Korpus eines Kindes und sind wie folgt definiert:

Für die Ermittlung des Prozessranges einer Koda wird die Menge aller Sitzungen eines Korpus in eine zeitliche Rangfolge gebracht, in der eine zeitlich frühere Sitzung einen niedrigeren Rang hat als eine zeitlich spätere Sitzung. Der *Prozessrang* einer Koda ist der Rang der Sitzung, in der sie zum ersten Mal während des Gesamtprozesses überhaupt in einer Sitzung auftritt. Analog zur Definition des Prozessrangs einer Koda wird für den Begriff des *Rangs* einer Koda *bzgl.* eines Kopfes die Menge aller Sitzungen, in der dieser Kopf in einer Silbe vorkommt, in eine Rangordnung gebracht, indem eine zeitlich früher platzierte einen niedrigeren Rang erhält als eine zeitlich später platzierte. Die Abb. 37 entspricht nun einem Streudiagramm⁷⁸ des Prozessrangs einer Koda *x* und des Rangs dieser Koda *x* bzgl. dem Kopf /p/ für alle Kodas, die mit dem Kopf /p/ in einer Silbe vorkommen. Es stellt für jede Koda *x*, die mit dem Kopf /p/ in einer Silbe zusammen auftritt, einen Punkt⁷⁹ dar, der für diese Koda die Relation repräsentiert zwischen ihrem Prozessrang (d.h. in der wievielten Sitzung des Gesamtprozesses ist diese Koda zum ersten Mal in einer Silbe aufgetreten?), der auf der senkrechten Achse des Koordinatensystems abgetragen wird, wobei die aufsteigende Zählung mit 0 beginnt, und ihrem Rang bzgl. des Kopfes /p/, der auf der waagerechten Achse des Koordinatensystems markiert ist, wobei die aufsteigende Zählung mit 1 beginnt. So steht beispielsweise der mit einer "1" gekennzeichnete schwarze Punkt in der Abb. 37 für die Koda, die zum ersten Mal überhaupt im Silbenbildungsprozess in der 1-ten Sitzung in einer Silbe auftritt (Prozessrang dieser Koda) (markiert auf der senkrechten Achse) und zudem das erste Mal in der 1-ten Sitzung aller Sitzungen vorkommt, in denen eine Silbe mit dem Kopf /p/ vorhanden ist (Rang bzgl. dieser Koda bzgl. des Kopfes /p/) (markiert auf der waagerechten Achse des Koordinatensystems). Dass die oben geäußerte Vermutung nicht zutreffend ist, dass ein bestimmter Kopf in seinem Kombinationsverhalten auf zunächst vor seinem ersten Auftreten vorkommende Kodas zurückgreift, bevor er sich mit neuen Kodas verbindet, zeigt die Art der Verteilung der roten und schwarzen Punkte in der Abb. 37: Rote Punkte stehen in der Abb. 37 für die Kodas *x* des Annalena-Korpus mit der folgenden Eigenschaft: Die Koda *x* tritt das erste Mal überhaupt in einer Silbe mit dem Kopf /p/ auf. Für diese Kodas *x* stimmt also ihr Prozessrang mit dem Rang bzgl. des Kopfes /p/ überein. Dies ist in der Legende des vorhergehenden Schaubildes entsprechend durch ein Gleichheitszei-

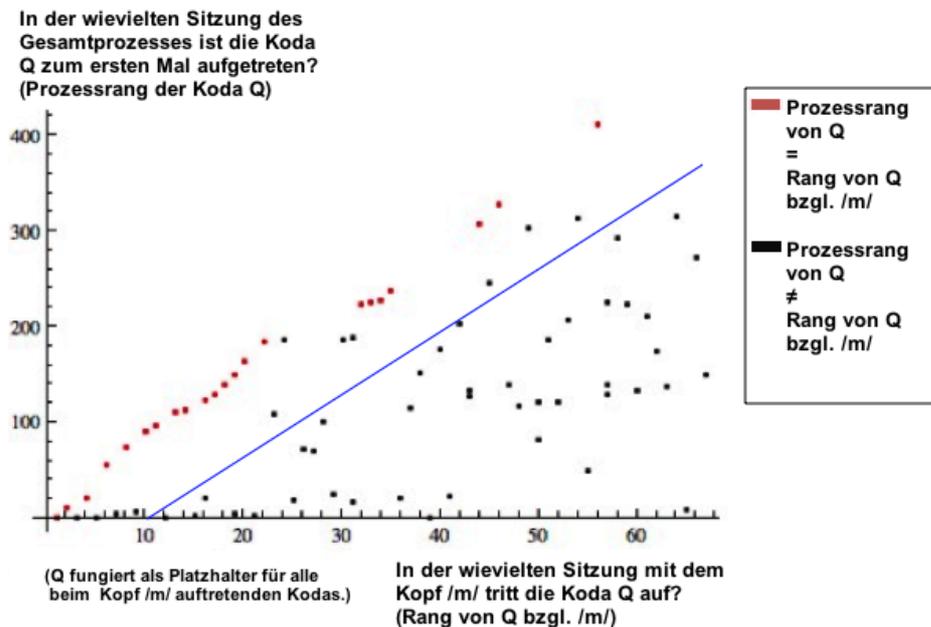
⁷⁸ siehe dazu auch Abschnitt 7.3.2., in dem die den Abbildungen zu Grunde liegenden quantitativen Grundverhältnisse und deren Relationen erläutert werden

⁷⁹ der entweder rot oder schwarz markiert ist. Diese Farbsymbolik wird im Anschluss noch näher erläutert.

chen ("=") gekennzeichnet. Schwarz markierte Punkte verweisen in der Abb. 37 hingegen auf die Kodas x des Annalena-Korpus mit der folgenden entgegengesetzten Eigenschaft: Die Koda x tritt nicht das erste Mal überhaupt in einer Silbe mit dem Kopf /p/ auf. Für diese Kodas x stimmt also ihr Prozessrang mit dem Rang bzgl. des Kopfes /p/ nicht überein. Dies ist in der Legende des vorhergehenden Schaubildes entsprechend durch ein Ungleichheitszeichen ("≠") gekennzeichnet. Würde nun die eingangs aufgestellte Vermutung zutreffen, so dürfte nach dem Auftreten der ersten neuen (s.o.) Koda y des Kopfes /p/ keine Koda mehr auftreten, deren Prozessrang kleiner als der von y ist. Dies ist jedoch der Fall: Der mit einer "2" gekennzeichnete rote Punkt steht für die erste neue Koda des Kopfes /p/. Diese erste neue Koda des Kopfes /p/ hat den Prozessrang 70 und den Rang 3 bzgl. /p/. In späteren Sitzungen, die den Kopf /p/ in einer ihrer Silben enthalten, treten aber wiederum Kodas auf, deren Prozessrang kleiner als 70 ist und deren Rang bzgl. des Kopfes /p/ größer als 3 ist und somit später als die durch den mit "2" gekennzeichneten roten Punkt symbolisierte neue Koda des Kopfes /p/ auftreten: So weist beispielsweise die durch den mit "3" markierten schwarzen Punkt symbolisierte Koda den Prozessrang 0 und den Rang 10 bzgl. des Kopfes /p/ auf. Damit widerspricht die Verteilung der neuen (rote Punkte) und wiederholten (schwarze Punkte) Kodas bzgl. des Kopfes /p/ der eingangs aufgestellten Vermutung, dass ein bestimmter Kopf in seinem Kombinationsverhalten auf zunächst vor seinem ersten Auftreten vorkommende Kodas zurückgreift, bevor er sich mit neuen Kodas verbindet.

Die vorhergehende Vermutung ist ein Spezialfall der allgemeineren Aussage, dass der Prozessrang in der Menge aller Kodas, die mit einem bestimmten Kopf x in Silben auftreten, innerhalb einer gewissen Bandbreite mit dem Rang dieser Kodas bzgl. dieses Kopfes x zunimmt bzw. wächst. Dies würde bedeuten, dass die neuen und wiederholten Kodas in der Abb. 38 (entsprechend für den Kopf /m/ zum Schaubild für den Kopf /p/) oberhalb des Graphen einer Geraden verlaufen, ähnlich der blauen Linie, wie sie in Abb. 38 per Hand eingezeichnet wurde. Auch wenn man für diese Gerade (blau) niedrigere y-Achsenabschnitte wählen würde, verblieben immer noch Punkte unterhalb.

Abbildung 38: Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /m/:



Eine ähnlich Überlegung gilt auch für eine in die Darstellung der zeitlichen Verteilung der Kodas in allen Silben mit dem Kopf /b/ eingezeichnete analoge blaue Gerade:

Abbildung 39: Streudiagramm des Zeitpunktes des ersten Auftretens einer Koda im Gesamtprozess und ihrem Zeitpunkt des ersten Auftretens beim Kopf /b/:

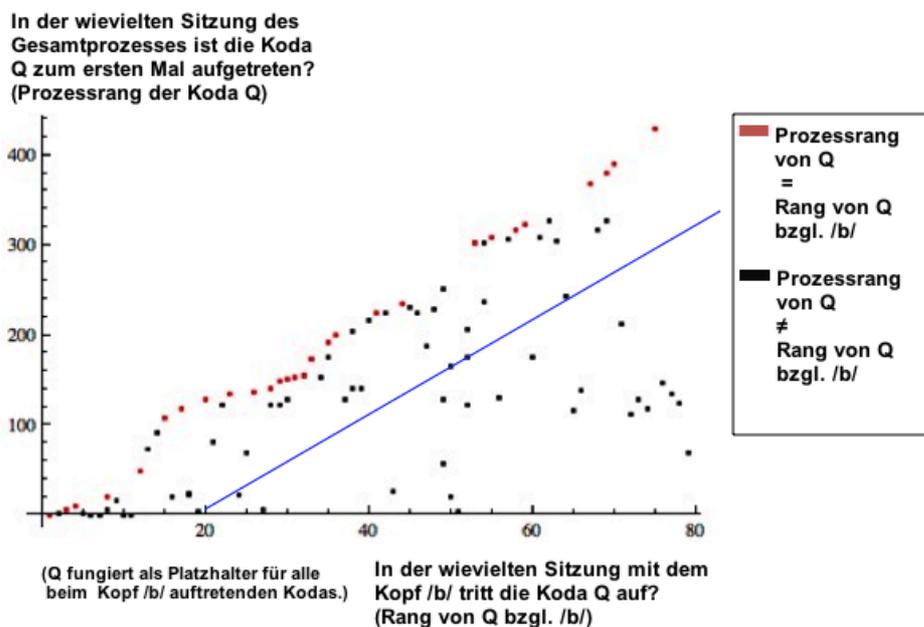


Abb. 37, 38 und 39 verdeutlichen, dass ein Kopf sich im Zeitverlauf sowohl mit neuen Kodas (rot) als auch mit zeitlich früher auftretenden Kodas (schwarz) zu Schalen verbindet und der Prozessrang der wiederholten Kodas eines Kopfes weit im Prozessverlauf bis zur ersten Sitzung des Korpus zurückreichen kann und im Allgemeinen der Prozessrang

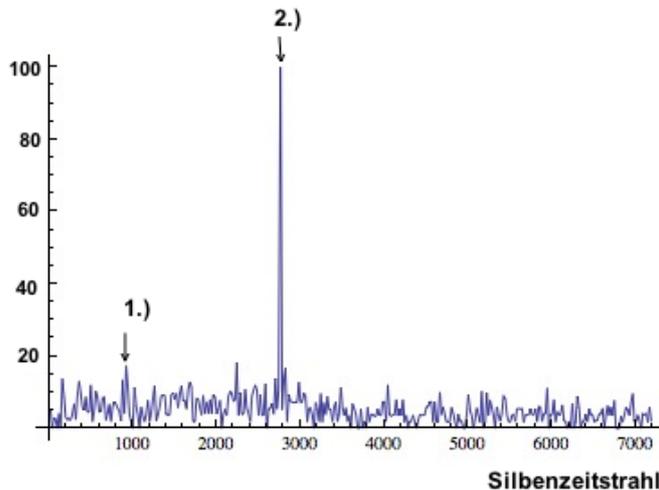
einer Koda nicht in Abhängigkeit von ihrem Rang bzgl. dieses Kopfes in einer gewissen Bandbreite zunimmt.

7.3.9.2. Vergleich zeitbezogener Binärrelationsmengen

In diesem Abschnitt werden Quantitäten von zeitbezogenen Silbenmengen miteinander verglichen wie in Abschnitt 7.3.2., Punkt 4.) "Mengenmäßiger Vergleich unterschiedlicher gleich großer Silbenzeiträume" als eine von 6 möglichen quantitativen Grundverhältnissen für die in dieser Arbeit dargestellten Abbildungen beschrieben wurde. Für den Vergleich der noch näher zu bestimmenden Quantitäten wurde der Silbenzeitstrahl, wie bereits an anderer Stelle auch schon geschehen, in direkt aufeinanderfolgende, überschneidungsfreie, gleich große Zeiträume von je 20 Silben (auch *Blöcke* genannt) eingeteilt. Die gleiche Größe der durch Einteilung des Silbenzeitstrahls erhaltenen Silbenzeiträume ermöglicht so ihre Vergleichbarkeit hinsichtlich ausgesuchter Quantitäten. Anschließend wurde die Binärrelationsmenge jedes Zeitraums mit der jedes anderen verglichen, indem die Anzahl der gemeinsamen Binärrelationen dieser beiden Zeiträume als Prozentanteil auf die Anzahl aller Binärrelationen eines der beiden Zeiträume bezogen wurde. Da die Unterteilung des Silbenzeitstrahls von 7190 Silben eine Unterteilung in 359 direkt aufeinanderfolgende, überschneidungsfreie, gleich große 20-Silbenzeiträume zulässt, wurden $359 \cdot 359 = 128.881$ Vergleiche durchgeführt, von denen in den Abb. 40, 41 und 42 die Vergleiche aller 20-Silbenzeiträume jeweils mit dem Silbenzeitraum 2760-2780, 3260-3280 oder 6060-6080 als Bezugszeitraum aufgeführt sind. In Abb. 40 sind die Vergleiche mit dem Zeitraum 2760-2779 als Bezugszeitraum aufgeführt, d.h. die Anzahl der Binärrelationen, die ein 20-Silbenzeitraum x mit dem Zeitraum 2760-2779 gemeinsam hat, wird auf die Anzahl aller Binärrelationen des Zeitraums 2760-2779 als Prozentangabe y bezogen. Diese Prozentangabe ist am Anfangszeitpunkt des Silbenzeitraums x abgetragen, und alle notierten Werte sind durch Geradenstücke miteinander verbunden worden. So zeigt der mit "1.)" und einem Pfeil markierte Wert (Abb. 40) den zuvor beschriebenen Prozentanteil y (hier: 17,39%) für den Silbenzeitraum 920-939 mit dem Bezugszeitraum 2760-2780. Der mit "2.)" und einem hinweisenden Pfeil gekennzeichnete Wert zeigt hingegen den zuvor beschriebenen Anteilswert y des Vergleichs des 2760-2779-Silbenzeitraums mit sich selbst und stellt somit gegenüber allen anderen Werten der Abb. 40 den einzigen 100-Prozentanteil dar.

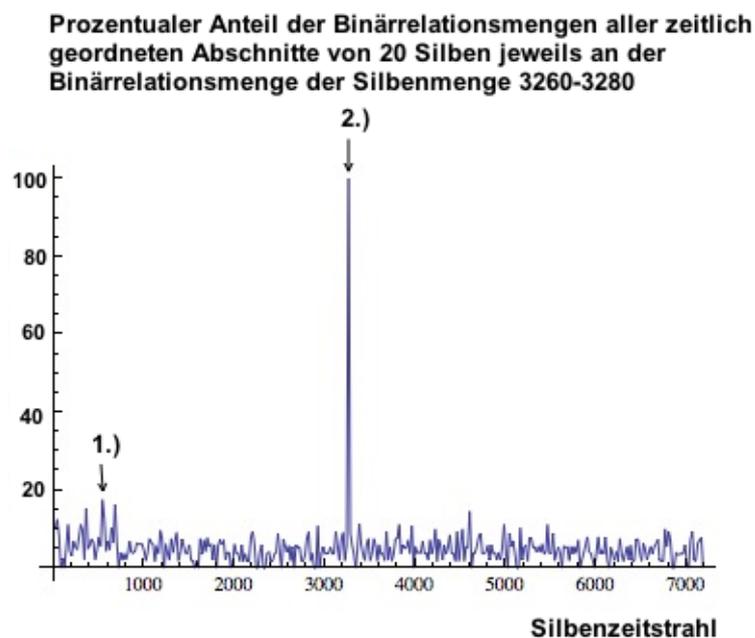
Abbildung 40: Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 2760-2780)

Prozentualer Anteil der Binärrelationsmengen aller zeitlich geordneten Abschnitte von 20 Silben jeweils an der Binärrelationsmenge der Silbenmengen 2760-2780



Bis auf den Prozentanteilswert von 100% (gekennzeichnet mit "2.")), der sich aus dem Vergleich des für die Abb. 40 gewählten Bezugszeitraums 2760-2780 mit sich selbst ergibt, liegen die Prozentanteilswerte für die Vergleiche des Bezugszeitraums 2760-2780 mit allen anderen, von ihm verschiedenen Zeiträumen unterhalb der Marke von 18,18%, die für den Vergleich des Zeitraums 2240-2260 mit dem Bezugszeitraum erreicht wird. Es ist also in der Abb. 40 zu beobachten, dass sich in keinem anderen, vom Bezugszeitraum verschiedenen Zeitabschnitt die Menge aller Binärrelationen des 20-Silbenzeitraums 2760-2780 vollständig wiederholt. Der Prozess ist von einer vollständigen Wiederholung weit entfernt, stattdessen liegt der Durchschnittswert aller Übereinstimmungen bzw. der durch die Vergleiche ermittelten Prozentanteilswerte bei 2.895%. Auch bei der Wahl aller anderen 358 möglichen Bezugszeiträume, die sich aus der bereits erwähnten Unterteilung des Silbenzeitstrahls in 20-Silbenzeiträume ergeben, ist eine vergleichbare Beobachtung zu machen wie bei der Abb. 40 mit dem Bezugszeitraum 2760-2780. Es sollen beispielhaft die zur Abb. 40 analogen Abbildungen für die Bezugszeiträume 3260-3280 und 6060-6080 gezeigt werden. Die Abb. 41 zeigt die Darstellung der Prozentanteilswerte y (s.o.) aller Vergleiche des Bezugszeitraums 3260-3280 mit sich selbst und allen anderen 20-Silbenzeiträumen der Unterteilung des Silbenzeitstrahls.

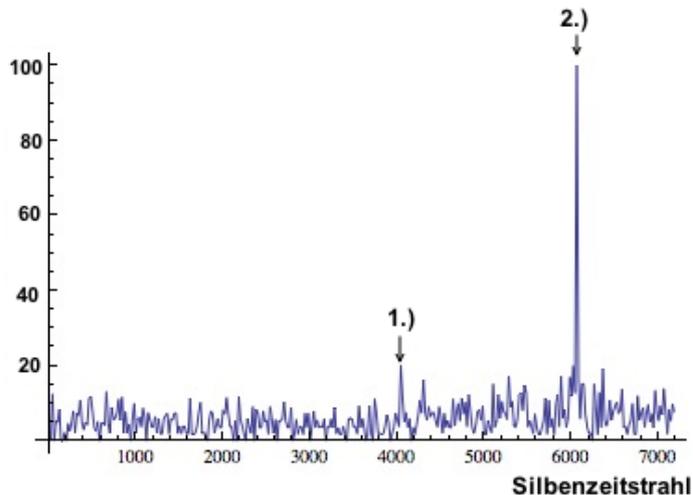
Abbildung 41: Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 3260-3280)



Durch "2.)" und den hinweisenden Pfeil ist in der Abb. 41 der 100%-Anteilswert des Vergleichs des Bezugszeitraums mit sich selbst markiert. Dieser übersteigt die Prozentanteilswerte aller Vergleiche der Binärrelationsmenge des Bezugszeitraums mit der aller anderen, von ihm verschiedenen 20-Silbenzeiträume wie zuvor in der Abb. 40 um mehr als 80%. Der maximale Prozentanteil der Vergleiche des Bezugszeitraums mit allen anderen, von ihm verschiedenen 20-Silbenzeiträumen liegt lediglich bei 17,5% und wird für den Vergleich der Binärrelationsmenge des Zeitraums 540-560 mit dem Bezugszeitraum erreicht. Dieser Maximalwert ist in der Abb. 41 durch "1.)" gekennzeichnet. Der Durchschnittswert aller in der Abb. 41 dargestellten Prozentanteilswerte liegt bei nur 2.512%. Die Abb. 42 zeigt die zu den Abb. 40 und Abb. 41 analoge Darstellung der Prozentanteilswerte y (s.o.) aller Vergleiche des Bezugszeitraums 6060-6080 mit sich selbst und allen anderen 20-Silbenzeiträumen der Unterteilung des Silbenzeitstrahls.

Abbildung 42: Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Bezugszeitraum 6060-6080)

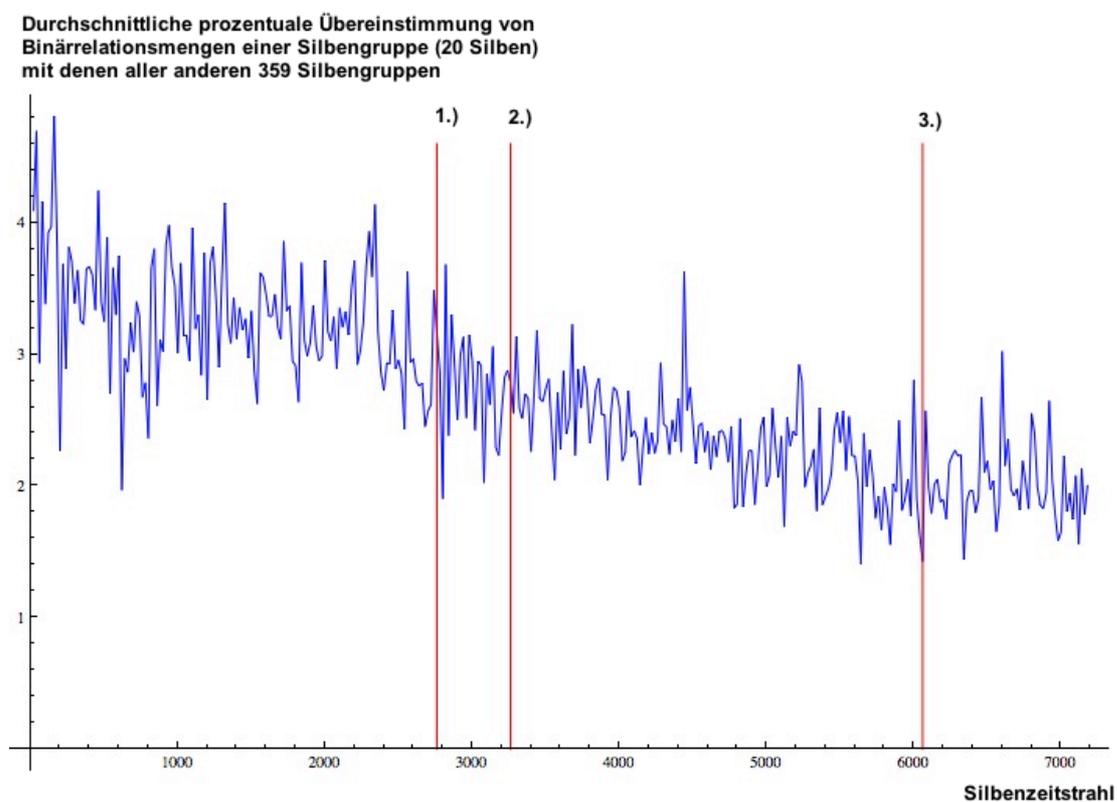
Prozentualer Anteil der Binärrelationsmengen aller zeitlich geordneten Abschnitte von 20 Silben jeweils an der Binärrelationsmenge der Silbenmenge 6060-6080



Durch "2.)" und den hinweisenden Pfeil ist in der Abb. 42 der 100%-Anteilswert des Vergleichs des Bezugszeitraums mit sich selbst markiert. Dieser übersteigt die Prozentanteilswerte aller Vergleiche der Binärrelationsmenge des Bezugszeitraums mit der aller anderen, von ihm verschiedenen 20-Silbenzeiträume wie zuvor in der Abb. 40 und Abb. 41 um mehr als 80%. Der maximale Prozentanteil der Übereinstimmungen des Bezugszeitraums mit allen anderen, von ihm verschiedenen 20-Silbenzeiträumen liegt lediglich bei 20% und wird für den Vergleich der Binärrelationsmenge des Zeitraums 4040-4060 mit dem Bezugszeitraum 6060-6080 erreicht. Dieser Maximalwert ist in der Abb. 42 durch "1.)" gekennzeichnet. Der Durchschnittswert aller in der vorhergehenden Abbildung dargestellten Prozentanteilswerte liegt bei nur 1.145%. Die bisher für jede der Abb. 40, 41 und 42 gegebenen durchschnittlichen Prozentanteilswerte, die die durchschnittliche Gemeinsamkeit der Binärrelationsmengen aller der jeweils mit dem Bezugszeitraum verglichenen 20-Silbenzeiträume der Silbenzeitstrahleinteilung angeben, sind nun in der Abb. 43 neben den drei bereits präsentierten 20-Silbenzeiträumen 3260-3280, 2760-2780 und 6060-6080 für jeden anderen 20-Silbenzeitraum der verbliebenen 356 (s.o.) 20-Silbenzeiträume als Bezugszeitraum berechnet worden und am Anfangszeitpunkt jedes dieser Zeiträume in der Abb. 43 abgetragen worden. So markieren die drei senkrechten roten Linien in Abb. 43 diese bereits bei der Besprechung zu den Abb. 40, 41 und 42 genannten Durchschnittsprozentanteilswerte: Die mit der "1.)" gekennzeichnete rote senkrechte Linie gibt den durchschnittlichen Wert der Prozentanteilswerte aller Vergleiche

mit dem 20-Silbenzeitraum 2760-2780 als Bezugszeitraum an, die mit "2.)" markierte den entsprechenden Anteilswert mit dem Bezugszeitraum 3260-3280 und die mit "3.)" gekennzeichnete ebenfalls den analogen Wert für die entsprechenden Vergleiche mit 6060-6080 als Bezugszeitraum.

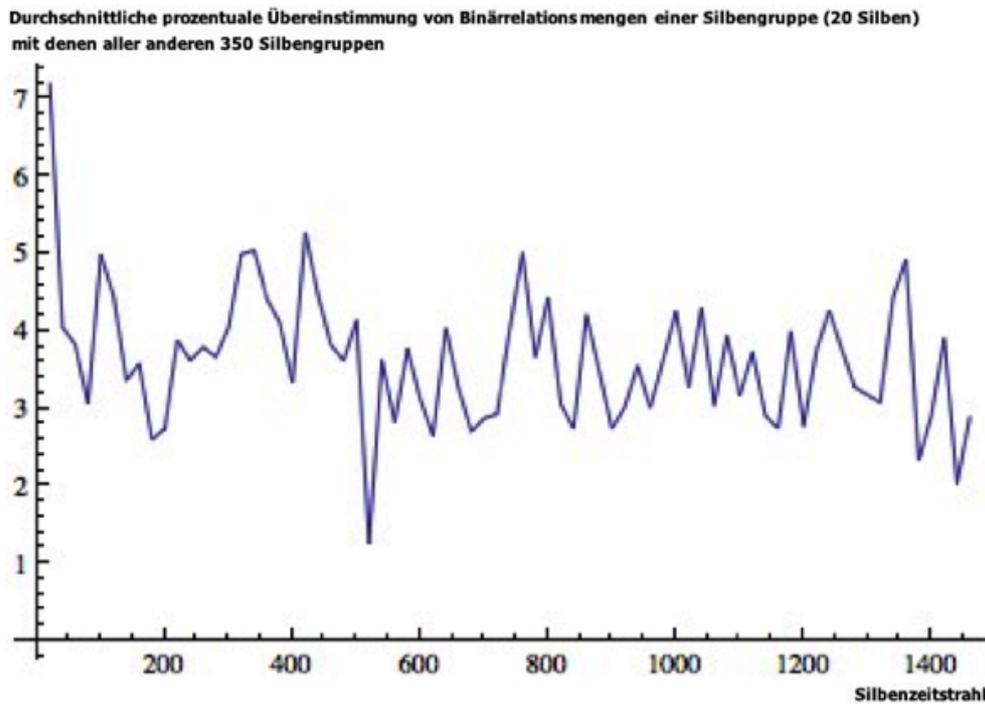
Abbildung 43: durchschnittliche Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Annalena-Korpus)



Über die Betrachtung von einzelnen Durchschnittswerten wie den mit "1.)", "2.)" und "3.)" markierten zeigt die Abb. 43 auch den Bereich an, in dem diese 359 Durchschnittswerte streuen: Es wird deutlich, dass sich die Vergleichswerte jedes einzelnen Silbenzeitraums mit allen anderen Silbenzeiträumen bei einer Bandbreite von ca. 1,2% - ca. 4,2% auf einem geringen Niveau befinden⁸⁰. Die zur Abb. 43 für das Isa-Korpus analoge Abbildung zeigt, dass auch für das ISA-Korpus diese Durchschnittswerte niedrig sind und zwischen maximal ca. 7.2% und minimal ca. 1% streuen:

⁸⁰ Obwohl auf dem Zeitstrahl die 7000 Silben abgetragen sind, beziehen sich die Werte auf 359 einzelne Gruppen von jeweils 20 Silben, die entsprechend der zeitlichen Gruppenverteilung eingeordnet sind.

Abbildung 44: durchschnittliche Binärrelationsmengengemeinsamkeiten von Silbenzeiträumen (Isa-Korpus)



7.3.9.3. Das Verhältnis aus einer Silbenproduktivitätszunahme und dem Schalenwachstum:

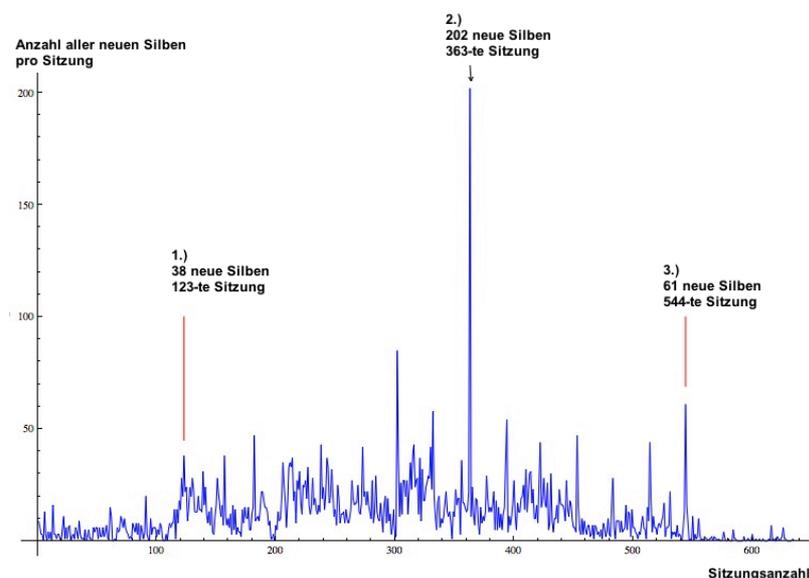
In diesem Abschnitt soll ein möglicher Zusammenhang zwischen der beobachteten merklichen Silbenproduktivitätssteigerung und dem Zuwachs neuer Schalen aufgezeigt werden. Die Abb. 45 zeigt die Anzahl neuer Silben pro Sitzung. Um die 110-te Sitzung findet ein merklicher Anstieg der durchschnittlichen sitzungsweisen Silbenanzahl statt, der in der 123-ten Sitzung mit einem Wert von 38 neuen Silben sein vorläufiges Maximum erreicht⁸¹. Als Anhaltspunkt für diesen Anstieg der durchschnittlichen sitzungsbezogenen Anzahl neuer Silben können die beiden folgenden arithmetischen Mittelwerte dienen:

Erstens wird das arithmetische Mittel über die Anzahl aller neuen Silben des Anfangszeitraums gebildet, der sich von der 1-ten bis einschließlich der 122-ten Sitzung erstreckt, deren Wert direkt links von der mit "1.)" gekennzeichneten roten Linie der Abb. 45 liegt. Der Mittelwert für diesen Anfangszeitraum beträgt $(598 \text{ Silben}) / (122 \text{ Sitzungen}) = 4.902 \text{ Silben/Sitzung}$, d.h. es treten in den 122 Sitzungen dieses Anfangszeitraums im Schnitt 4.902 neue Silben pro Sitzung auf. Zweitens wird das arithmetische Mittel über die Anzahl aller neuen Silben des dem Anfangszeitraums folgenden Zeitraums gebildet, der sich

⁸¹ Der Wert von 38 neuen Silben in der 123-ten Sitzung ist in Abb. 45 durch eine direkt über diesem Wert senkrecht verlaufende, mit "1.)" gekennzeichnete rote Linie markiert.

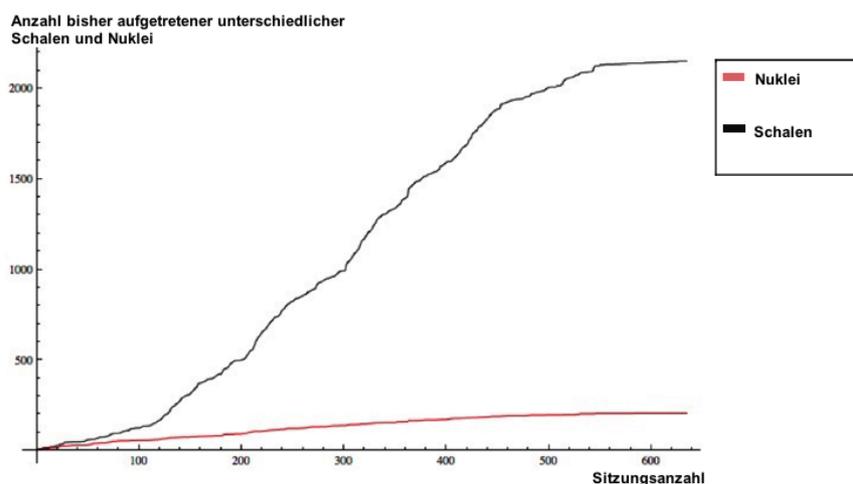
von der 123-ten Sitzung (gekennzeichnet durch "1.") und eine senkrechte rote Linie) bis einschließlich der 544-ten Sitzung erstreckt (gekennzeichnet durch "3.") und eine senkrechte rote Linie). Dieser Mittelwert für diesen Folgezeitraum beträgt $(6475 \text{ Silben}) / (422 \text{ Sitzungen}) = (15.344 \text{ Silben/Sitzung})$, d.h. es treten in den 422 Sitzungen dieses Folgezeitraums im Schnitt 15.344 neue Silben pro Sitzung auf. Die durchschnittliche Anzahl der neuen Silben pro Sitzung nimmt also vom zuvor genannten Anfangszeitraum zum beschriebenen Folgezeitraum so zu, dass sie sich mehr als verdreifacht. Zusätzlich zu den für die Darstellung der beschriebenen Silbenproduktivitätszunahme eingefügten Markierungen ist in der Abb. 45 ein weiterer Wert durch "3.)" und einen hinweisenden Pfeil gekennzeichnet. Er ist verglichen mit allen anderen Werten der folgenden Abbildung extrem groß und maximal und beträgt 202 neue Silben in der 363-ten Sitzung. Dieser Wert ist ein Beispiel für die starken Schwankungen bzw. Ausreißer bei der Verteilung der Anzahl neuer Silben im Zeitverlauf, die dadurch erklärt werden, dass bei der Digitalisierung des Annalena-Korpus neben der ebenfalls ungenauen Zeitangabe in Form von Tagen für die vom Kind geäußerten Lautformen auf noch ungenauere Zeitangaben in Form von größeren Zeitbereichen wie beispielsweise Monatsangaben (Elsen 1991: 136ff.) zurückgegriffen werden musste. Das in dieser Arbeit vorgestellte Modell geht an sich davon aus, dass die absolute Auftretensreihenfolge der Lautformen wichtig ist und deshalb eine zeitliche Ordnung der Lautformen notwendig ist, die diese anstatt sie zu gruppieren in eine absolute Ordnung bringt, in der jede neue Lautform eine von allen anderen verschiedene Position hat. Dennoch wird mit diesen Ungenauigkeiten gearbeitet, weil detailliertere Phonemerwerkskorpora als das Elsen-Korpus nicht verfügbar waren.

Abbildung 45: Sitzungsbezogene Anzahl neuer Silben (Annalena-Korpus)



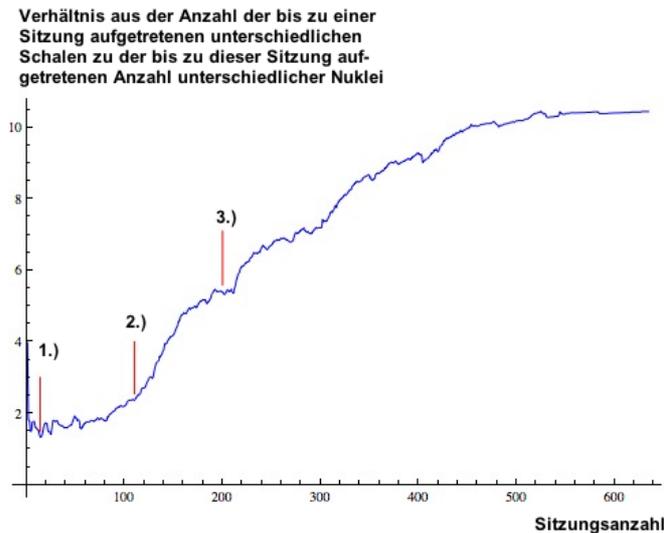
Mit diesem Anstieg der Silbenproduktivität geht auch eine starke Zunahme der Anzahl neuer Schalen (schwarz) einher. Demgegenüber steigt die Anzahl neuer Silbenkerne (rot) kaum an. Dies ist in Abb. 46 zu sehen: Eine senkrechte blaue Linie in der Abbildung dient der Unterteilung des Zeitverlaufs in zwei Zeiträume von der 1-ten bis zur 110-ten Sitzung, im folgenden auch *Anfangszeitraum* genannt, sowie von der 111-ten Sitzung bis zum zeitlichen Ende des Annalena-Korpus reichend, im Folgenden auch als *Folgezeitraum* bezeichnet. Im Anfangszeitraum beträgt die durchschnittliche Anzahl neuer Nuklei pro Sitzung 0,53 und die neuer Schalen 1,78. Demgegenüber nimmt die durchschnittliche Anzahl neuer Schalen pro Sitzung im Folgezeitraum stark zu und beträgt 4,22, während sich die durchschnittliche Anzahl neuer Nuklei pro Sitzung im Folgezeitraum nicht stark von dem entsprechenden, zuvor genannten Wert für den Anfangszeitraum unterscheidet und sogar ihre Größe auf nur noch 0,28 nahezu halbiert.

Abbildung 46: Kumulative sitzungsbezogene Anzahl unterschiedlicher Schalen und Nuklei (Annalena-Korpus)



Diese im Vergleich zu der Anzahl neuer Kerne starke Zunahme der Anzahl neuer Schalen zeigt sich auch deutlich in dem Anstieg des entsprechenden Verhältnisses in der Abb. 47. In dieser ist für jede Sitzung des Annalena-Korpus das Verhältnis aus der Anzahl aller bisher aufgetretenen unterschiedlichen Schalen zu der Anzahl aller unterschiedlichen Nuklei eingetragen und die so eingetragenen Verhältnisse durch Geradenstücke miteinander verbunden worden: So entspricht dem mit "1.)" (und einer senkrechten roten Linie) gekennzeichnete Wert das für die Abb. 47 minimale Verhältnis von 1,33, das in der 14-ten Sitzung gemessen wurde.

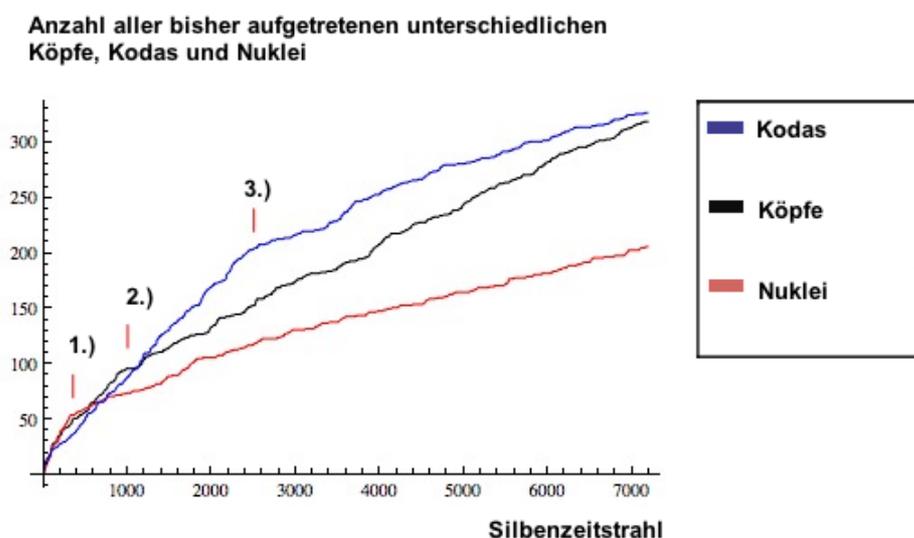
Abbildung 47: Verhältnis aus der kumulativen Anzahl unterschiedlicher Schalen zu der unterschiedlicher Nuklei (Annalena-Korpus)



Schon vor Beginn der 110-ten Sitzung vergrößert sich das Verhältnis aus der Anzahl bisher vorgekommener unterschiedlicher Schalen zu der der bisher aufgetretenen unterschiedlichen Silbenkerne ausgehend von dem zuvor genannten minimalen Wert von 1,33 ("1.)"-Kennzeichnung) bis auf über 10 gegen Ende des Annalena -Korpus: Bei der Markierung "2.)" ist ein Verhältnis von 2,36 für die 110-te Sitzung dargestellt, der schon größer ist als die analogen Verhältnisse für die vor der 110-ten Sitzung liegenden Sitzungen, insbesondere den mit "1.)" in der Abb. 47 markierten Wert. Bereits in der 200-ten Sitzung ("3.)"-Markierung) hat sich dieses Verhältnis im Vergleich zu dem mit "1.)" gekennzeichneten Wert mehr als verdoppelt und beträgt 5,39. In der Folge nimmt dieses Verhältnis für die der 200-ten Sitzung nachfolgenden Sitzungen immer weiter zu, jedoch mit durchschnittlich abnehmender Geschwindigkeit. Als eine mögliche Ursache des Anstiegs der Schalenanzahl und damit einhergehend der Silbenproduktivitätssteigerung sollte die Wachstumsverringering bei der Anzahl neuer Silbenkerne angesehen werden. Eine erste auffällige Wende im Wachstum ist für die Kerne in der Abb. 48 ungefähr nach der 300-ten Silbe gegeben. Entsprechend der dargestellten Graphik wächst die Zahl der neuen Köpfe, Kerne und Kodas in der initialen Phase ungefähr gleich schnell. Ab ca. der 100. Silbe verlangsamt sich das Wachstum der Kodas und ab der 225-ten das Wachstum der Köpfe. Dadurch übertrifft das Wachstum der Kerne für einen kurzen Zeitraum das der Kodas und der Köpfe. Jedoch verlangsamt sich die Wachstumsgeschwindigkeit neuer Kerne ab ca. der 350-ten Silbe (Markierung "1.") und zusätzlicher senkrechter roter Linie). Und schließlich übersteigt die Anzahl der Köpfe und Kodas die der Kerne. Da sich

die unterschiedlichen Wachstumsraten der drei Silbenkonstituenten auf eine absolute Silbenzahl beziehen, bedeutet ein relativ geringeres Wachstum einer Konstituente ihre stärkere Beteiligung am Silbenbildungsprozess durch die Kombination vorhandener Elemente mit neuen oder vorhandenen Elementen der beiden anderen Silbenkonstituentengruppen. Umgekehrt bedeutet ein relativ stärkeres Wachstum einer Gruppe im Verhältnis zu den beiden anderen, dass in ihr eine stärkere Differenzierung stattfindet. Das Wachstum der Gruppe der Kodas flacht ungefähr bei der 2500. Silbe ab (Markierung "3.") und zusätzlicher senkrechter roter Linie), während dies entsprechend für die Gruppe der Köpfe ab der 1000. Silbe (Markierung "2.") und zusätzlicher senkrechter roter Linie) gilt.

Abbildung 48: Kumulative Anzahl unterschiedlicher Köpfe, Kodas und Nuklei (Annale-na-Korpus)

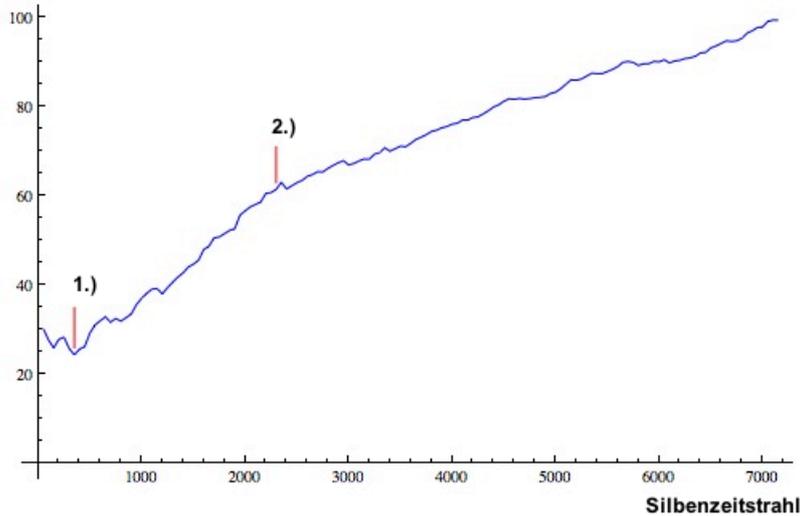


Neue Schalen entstehen aber nicht nur durch das stärkere Wachstum der Anzahl neuer Köpfe oder Kodas, da vor dem Beginn des Wachstumseinbruchs neuer Kerne oder fast gleichzeitig eine wachsende Kombination der bei den einzelnen Silbenkernen aufgetretenen Köpfe und Kodas beginnt, wie die folgenden Abb. 49 und Abb. 50 zeigen. In diesen wurde jeweils die Frage untersucht, welche Anzahl an Schalen ein Silbenkern tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der Silben verglichen wird, die aus den Schalenkonstituenten gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden.

Die Abb. 49 und Abb. 50 zeigen diese Anteile beispielhaft für die zwei häufigsten Silbenkerne /a/ und /ɪ/, wobei diese Anteile nur für jede 50-te Silbe des Silbenstrahls berechnet und dort abgetragen worden ist: Die "1.)"-Markierung (mit zusätzlicher hinweisender senkrechter, roter Linie) in der Abb. 49 zeigt den minimalen Wert des dargestellten Verhältnis, und zwar entspricht er 24.39. Er ergibt sich aus dem Verhältnis, dass zu diesem Zeitpunkt der 350-ten Silbe für den Silbenkern bereits insgesamt 20 unterschiedliche Köpfe und 23 unterschiedliche Kodas aufgetreten sind. 20 Köpfe und 23 Kodas lassen sich rein theoretisch zu 460 Schalen "kombinieren". Von diesen treten aber nur 205 wirklich im gesamten Annalena-Korpus auf, deren Zahl 205 als Bezugsgröße für das betrachtete Verhältnis genommen wird. Zu diesem Zeitpunkt sind aber erst 50 Schalen aufgetreten. Setzt man die Anzahl von 50 der zu diesem Zeitpunkt bereits aufgetretenen Schalen zu der der erwähnten 205 im Gesamtverlauf des Prozesses realisierten Schalen ins Verhältnis (und multipliziert für die Überführung des Verhältnisses in eine Prozentangabe mit 100), so erhält man den zu diesem Silbenzeitpunkt abgetragenen Verhältniswert von 24.39 ("1.)"-Markierung). Von diesem minimalen Verhältniswert zum Zeitpunkt der 350-ten Silbe ausgehend steigt das betrachtete Verhältnis, abgesehen von einigen kleineren Schwankungen, zunehmend an. So hat es bereits zum Zeitpunkt der 2300-ten Silbe die Größe 61.49 erreicht ("2.)"-Kennzeichnung) und sich damit im Verhältnis zum mit "1.)" gekennzeichneten Wert mehr als verdoppelt. Dieser Wert des Verhältnisses ergibt sich auf analoge Weise wie der mit "1.)" gekennzeichnete Wert: Es sind bis einschließlich zum Zeitpunkt der 2300-ten Silbe mit dem Silbenkern /a/ bereits 41 unterschiedliche Köpfe und 82 unterschiedliche Kodas in allen Silben aufgetreten. Die Anzahl der maximalen Kopf-Koda-Kombinationen würde $41 \cdot 82 = 3362$ betragen. Von diesen treten aber nur 457 Kombinationen tatsächlich im gesamten Annalena-Korpus auf. Das mit "2.)" gekennzeichnete Verhältnis 61.49 ergibt sich nun als $(281/457) \cdot 100 = 61.49$, d.h. als Verhältnis aus der Anzahl 281 aller bereits in allen Silben mit dem Silbenkern /a/ aufgetretenen Schalen zu der Anzahl aller aus den bereits aufgetretenen Köpfen und Kodas rein theoretisch bildbaren 3362 Kombinationen, eingeschränkt auf die Menge der 457 Kombinationen von diesen, die im gesamten Annalena-Korpus auch wirklich als Schalen in Silben mit dem Silbenkern /a/ vorkommen.

Abbildung 49: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /a/ im Zeitverlauf
(Annalena-Korpus)

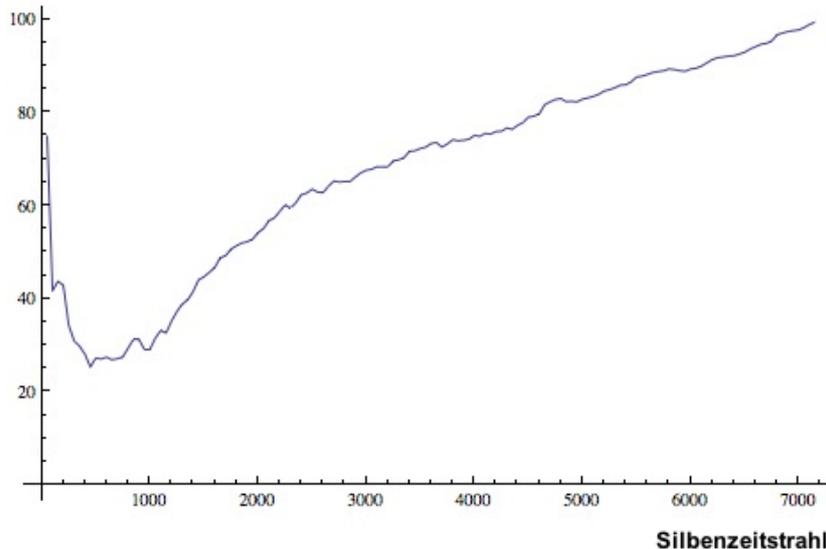
Anzahl der Schalen, die der Nukleus /a/ tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der /a/-Silben verglichen wird (als Prozentangabe), die aus den bis zu diesem Zeitpunkt in /a/-Silben aufgetretenen Silbenrändern gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden



Eine zu der für den Silbenkern /a/ analoge Beobachtung lässt sich auch für den Silbenkern /ɪ/ festhalten:

Abbildung 50: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /ɪ/ im Zeitverlauf
(Annalena-Korpus)

Anzahl der Schalen, die der Nukleus /ɪ/ tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der /ɪ/-Silben verglichen wird (als Prozentangabe), die aus den bis zu diesem Zeitpunkt in /ɪ/-Silben aufgetretenen Silbenrändern gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden.



Es lässt sich also zusammenfassend sagen, dass im Anfangszeitraum bis mindestens zur 550-ten Silbe aus den bei einem Silbenkern (hier /a/, /I/) aufgetretenen Köpfen und Kodas wenige Schalen "kombiniert" werden. Eine verstärkte "Kombination" findet erst frühestens nach diesem Zeitpunkt statt⁸². Obwohl diese Schalenkonstituenten vorhanden sind, wird aber nur ein geringer Prozentsatz der möglichen Schalenkombinationen zeitnah realisiert. Die zuvor für das Annalena-Korpus gegebene Darstellung der zu beobachtenden quantitativ-zeitlichen Korrelationen soll auch durch die Abb. 51 - Abb. 56 in verkürzter Form für das Isa-Korpus entwickelt werden. Für ausführlichere Erläuterungen können die folgenden Abbildungen mit den analogen, zuvor für das Annalena-Korpus gegebenen verglichen werden, dazu insbesondere die Abb. 45 mit der Abb. 51, Abb. 46 mit der Abb. 52, Abb. 47 mit der Abb. 53, Abb. 48 mit der Abb. 54 und schließlich die Abb. 49, 50 mit den Abb. 55, 56.

Auch das ISA-Korpus⁸³ zeigt eine starke Zunahme der Silbenproduktivität, d.h. der Anzahl der neuen Silben pro Sitzung, wie die Abb. 51 zeigt (vgl. Abb. 45 des Annalena-Korpus):

Abbildung 51: Sitzungsbezogene Anzahl neuer Silben (Isa-Korpus)

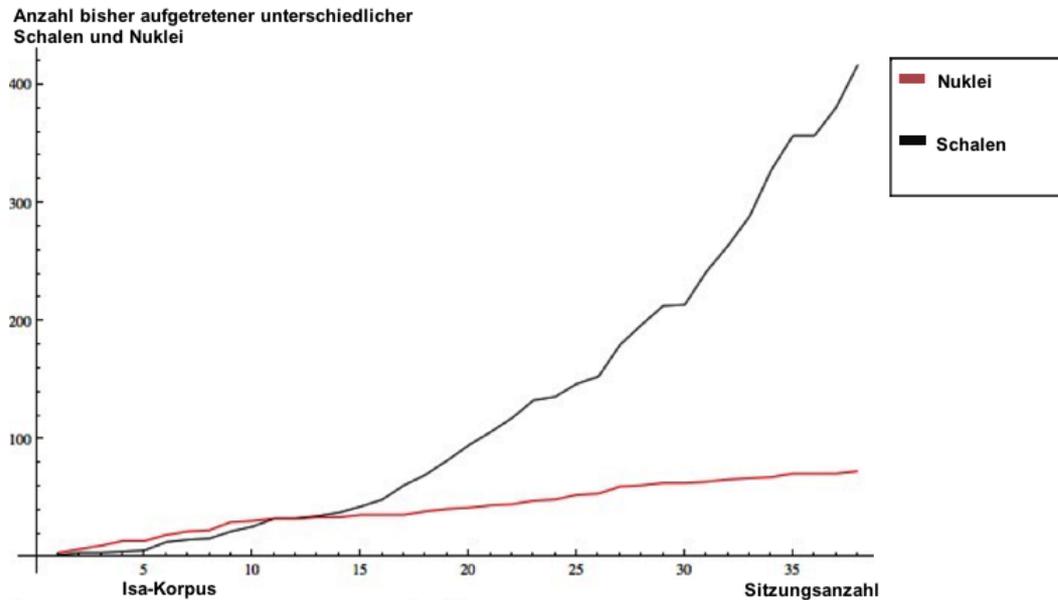


⁸² Das Ausmaß der Kombination aus den bei einem Silbenkern zu einem Zeitpunkt T bereits aufgetretenen Köpfen und Kodas wird dabei als Anteil dieser Kombinationen an der Menge aller Schalen gemessen, die im Gesamtkorpus bei diesem Silbenkern auftreten und aus den bis zu diesem Zeitpunkt aufgetretenen Köpfen und Kodas kombiniert werden können.

⁸³ Der direkte Vergleich ist jedoch wegen der unterschiedlichen zeitlichen Detailliertheit und der Länge des Beobachtungszeitraums der Korpora nicht ganz unproblematisch: Das Kind ISA wurde nur jeden siebten Tag über ein Jahr lang besucht, während das Kind Annalena tagtäglich über 1 1/2 Jahre lang beobachtet wurde. Dadurch können bei ISA zwischenzeitlich geäußerte Silben verlorengegangen sein. Dennoch zeigt auch das ISA-Korpus die oben beobachteten Merkmale.

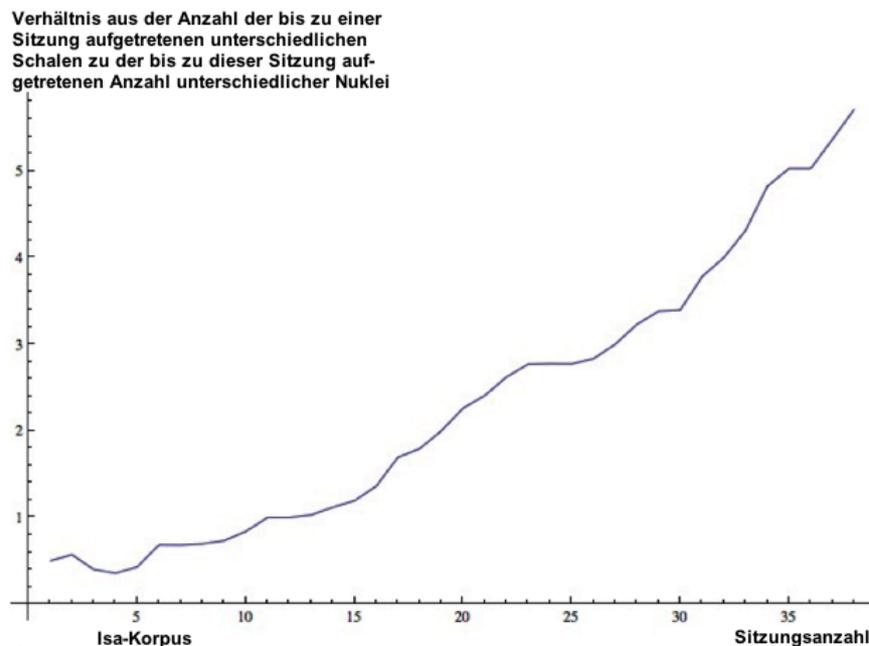
Im zeitlichen Bereich der 16-ten Sitzung (entspricht ca. der 250-ten Silbe) ist ein starker Zuwachs neuer Schalen zu verzeichnen (vgl. Abb. 46 des Annalena-Korpus),

Abbildung 52: Kumulative sitzungsbezogene Anzahl unterschiedlicher Schalen und Nuklei (Isa-Korpus)



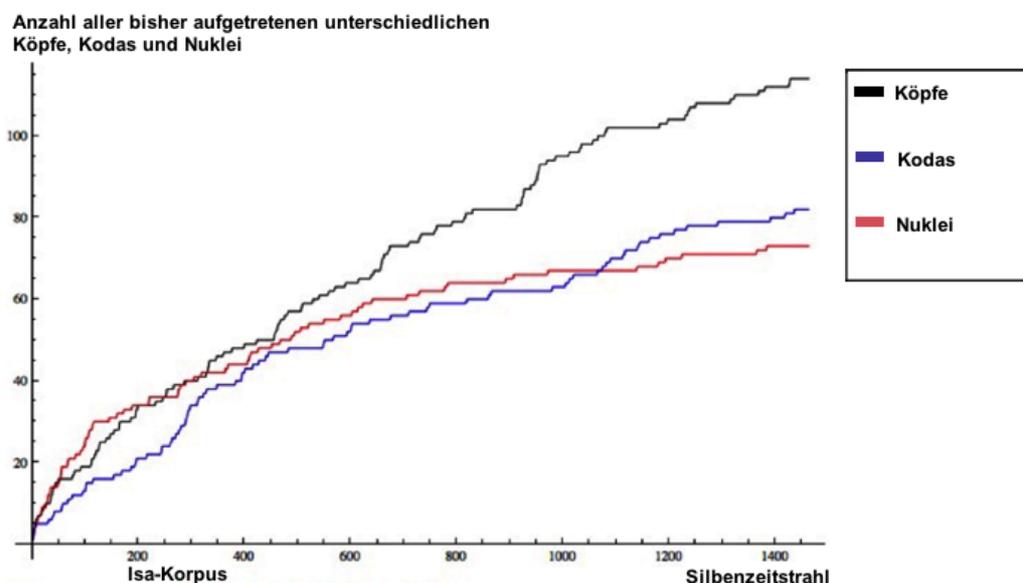
die erheblich stärker pro Sitzung wachsen als die Silbenkerne (vgl. Abb. 47 des Annalena-Korpus):

Abbildung 53: Verhältnis aus der kumulativen Anzahl unterschiedlicher Schalen zu der unterschiedlicher Nuklei (Isa-Korpus)



Auch beim ISA-Korpus wird die Silbenkernanzahl nach einem gewissen Zeitraum von der der Köpfe und Kodas übertroffen (vgl. Abb. 48):

Abbildung 54: Kumulative Anzahl unterschiedlicher Köpfe, Kodal und Nuklei (Isa-Korpus)



Außerdem werden ebenfalls die bei einem Silbenkern bereits aufgetretenen Köpfe und Kodal stärker miteinander zu Schalen kombiniert⁸⁴:

Dies ist im Folgenden für die beiden häufigsten Silbenkerne des ISA-Korpus, das /a/ und das /ɪ/, aufgeführt (vgl. die Abb. 49 und Abb. 50 des Annalena-Korpus):

Dazu wurde⁸⁵ in den folgenden Graphiken, wie bereits oben für das Annalena-Korpus erwähnt, jeweils die Frage untersucht, welche Anzahl an Schalen ein Silbenkern tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der Silben verglichen wird, die aus den Schalenkonstituenten gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden (vgl. auch die Besprechung zu den Abb. 49 und Abb. 50 des Annalena-Korpus).

⁸⁴ Das Ausmaß der Kombination aus den bei einem Silbenkern zu einem Zeitpunkt T bereits aufgetretenen Köpfen und Kodal wird dabei als Anteil dieser Kombinationen an der Menge aller Schalen gemessen, die im Gesamtkorpus bei diesem Silbenkern auftreten und aus den bis zu diesem Zeitpunkt aufgetretenen Köpfen und Kodal kombiniert werden können.

⁸⁵ wie oben bereits für das Annalena-Korpus beschrieben

Abbildung 55: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /a/ im Zeitverlauf (Isa-Korpus)

Anzahl der Schalen, die der Nukleus /a/ tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der /a/-Silben verglichen wird (als Prozentangabe), die aus den bis zu diesem Zeitpunkt in /a/-Silben aufgetretenen Silbenrändern gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden.

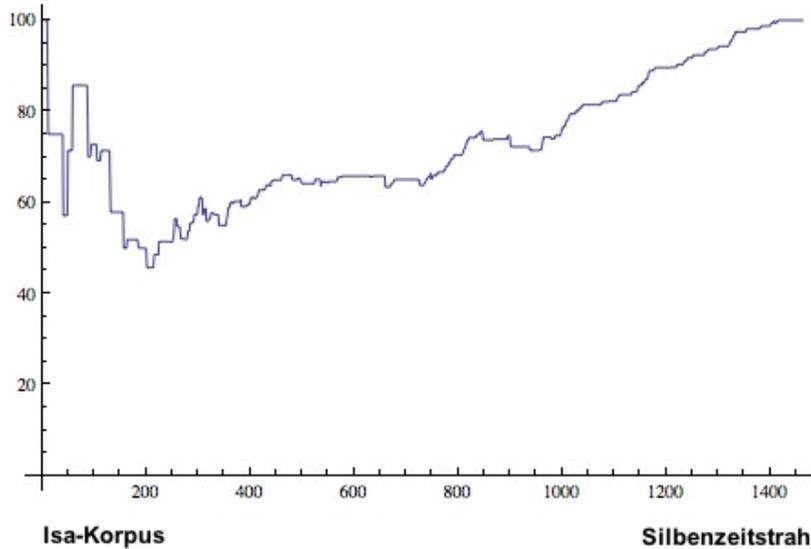
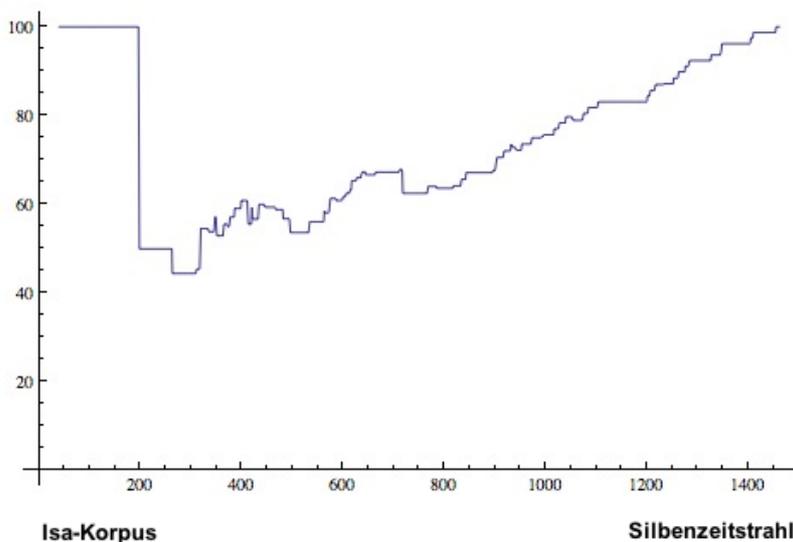


Abbildung 56: Ausmaß der Kopf-Koda-Kombinationen beim Nukleus /ɪ/ im Zeitverlauf (Isa-Korpus)

Anzahl der Schalen, die der Nukleus /ɪ/ tatsächlich bis zu einem bestimmten Zeitpunkt in einer Silbe realisiert, wobei diese Zahl mit der Menge der /ɪ/-Silben verglichen wird (als Prozentangabe), die aus den bis zu diesem Zeitpunkt in /ɪ/-Silben aufgetretenen Silbenrändern gebildet werden könnten und tatsächlich im Gesamtkorpus zu irgendeinem Zeitpunkt realisiert werden.



Die Silbenproduktivitätssteigerung weist also auf eine wichtige Entwicklung im Korpus hin, die mit der anfänglichen Konzentration auf die Kerne einhergeht und der späteren Trennung im Wachstumstempo der Köpfe und Koda.

7.3.9.4. Schalengemeinsamkeiten von Silbenkernen

Die Menge aller Silbenkerne des Gesamtkorpus lässt sich in zwei Gruppen einteilen, die sich hinsichtlich ihrer Schalenschnittmengen unterscheiden: Eine Menge A von Silbenkernen mit einer Schalengesamthäufigkeit⁸⁶, die größer oder gleich 11 (Annalena) (bzw. 12 (Isa)) ist und eine Menge B von Silbenkernen, deren Schalengesamthäufigkeit kleiner oder gleich 10 (Annalena) (bzw. 11 (Isa)) ist. Die Gruppe A besteht aus 45 (Annalena) (bzw. 24 (Isa)) unterschiedlichen Silbenkernen, die Gruppe B aus 162 (Annalena) (bzw. 49 (Isa)). Wenn beide Gruppen in einer Häufigkeitsrangfolge entsprechend der Schalengesamthäufigkeiten geordnet werden, kann man feststellen, dass in der Gruppe A die paarweisen Schalenschnittmengen keine Identität mit der Schalengesamtmenge⁸⁷ des rangniedrigeren zeigen. Im Gegensatz zur Gruppe A kann man bei einer analogen Betrachtung der Verhältnisse in der Gruppe B, die im Anschluss an die Gruppe A die Rangfolge nach unten fortsetzt, für paarweise Schnittmengen von Silbenkernen innerhalb der Gruppe B als auch für Schnittmengen eines Silbenkerns aus der Menge B mit einem aus der Gruppe A vollständige Übereinstimmungen mit der Schalengesamtmenge des selteneren Silbenkerns der beiden beobachten. Die Abb. 57 zeigt die prozentuale Schalenübereinstimmung zweier Silbenkerne aus der Gesamtmenge von A und B gemessen an der Schalenmenge des selteneren⁸⁸ (y-Achse), abgetragen gegen seine Schalengesamthäufigkeit (x-Achse). Diese Messwerte sollen an drei Beispielen von Silbenkernen der Gruppe A erläutert werden: So zeigt der mit "1.)" gekennzeichnete blaue Punkt in der Abb. 57 den Prozentanteil 41.87% der Anzahl 278 aller gemeinsamen Schalen des häufigsten Silbenkerns /a/ mit dem zweithäufigsten Silbenkern /ɪ/ des Annalena Korpus an der Anzahl 664 aller Schalen des Silbenkerns /ɪ/ im gesamten Annalena-Korpus. Der mit "2.)" gekennzeichnete blaue Punkt in der Abb. 57 zeigt analog den Prozentanteil 49.43% der Anzahl 254 aller gemeinsamen Schalen des häufigsten Silbenkerns /a/ mit dem dritthäufigsten Silbenkern /ɛ/ des Annalena Korpus an der Anzahl 522 aller Schalen des Silbenkerns /ɛ/ im gesamten Annalena-Korpus. Analog zu den beiden vorhergehend besprochenen

⁸⁶ Die Schalengesamthäufigkeit eines Silbenkerns ist die Menge all seiner Schalen im Gesamtkorpus und stimmt mit der Silbengesamthäufigkeit überein, sofern dieser Silbenkern nicht in einer gleichzeitig sowohl nackten (d.h. VC₂) als auch offenen (d.h. C₁V), d.h. einer V-Silbe vorkommt.

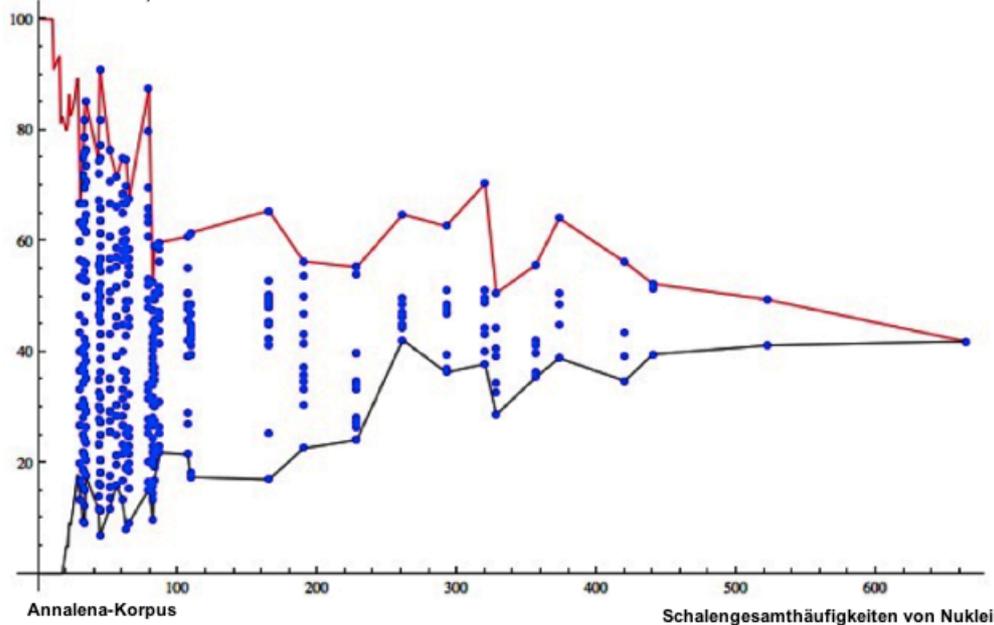
⁸⁷ d.h. die Schalenmenge des Gesamtkorpus eines Silbenkerns

⁸⁸ Bsp.: Ein rangniedrigerer Silbenkern soll 23 Schalen haben, ein ranghöherer eine Gesamtmenge von 40 Schalen, die gemeinsame Schalenmenge soll 10 betragen. Bezogen auf die Schalenmenge des rangniedrigeren beträgt deshalb der prozentuale Anteil der paarweisen Schalenschnittmenge 43,478%, bezogen auf die des ranghöheren beträgt er 25%. In der Abb. 57 wurde der Bezug zur Schalenmenge des rangniedrigeren gewählt.

Kennzeichnungen markiert "3.)" in der Abb. 57 den Prozentanteil 41.19% der Anzahl 214 aller gemeinsamen Schalen des zweithäufigsten Silbenkerns /i/ mit dem dritthäufigsten Silbenkern /ε/ des Annalena Korpus an der Anzahl 522 aller Schalen des Silbenkerns /ε/ im gesamten Annalena-Korpus. Die Abb. 57 stellt zudem durch eine Minimum- (schwarz) und eine Maximum-Linie (rot) dar, in welchem Bereich diese Anteile streuen⁸⁹ und dass die Maximum-Linie nur für Schalengesamthäufigkeiten der Silbenkerne der Gruppe B eine vollständige Übereinstimmung (100%) zeigt.

Abbildung 57: Schalengemeinsamkeiten von Silbenkernpaaren (Annalena-Korpus)

Prozentuale Übereinstimmung gemeinsamer Schalenmengen jedes Nukleus mit jeweils jedem anderen als Bandbreitendarstellung (Minima, Maxima, Zwischenwerte)
 (Für jedes Paar von Nuklei der Prozentanteil ihrer gemeinsamen Schalen an der Schalengesamthäufigkeit des selteneren Nukleus des Paares jeweils abgetragen gegen die Schalengesamthäufigkeit des selteneren.)

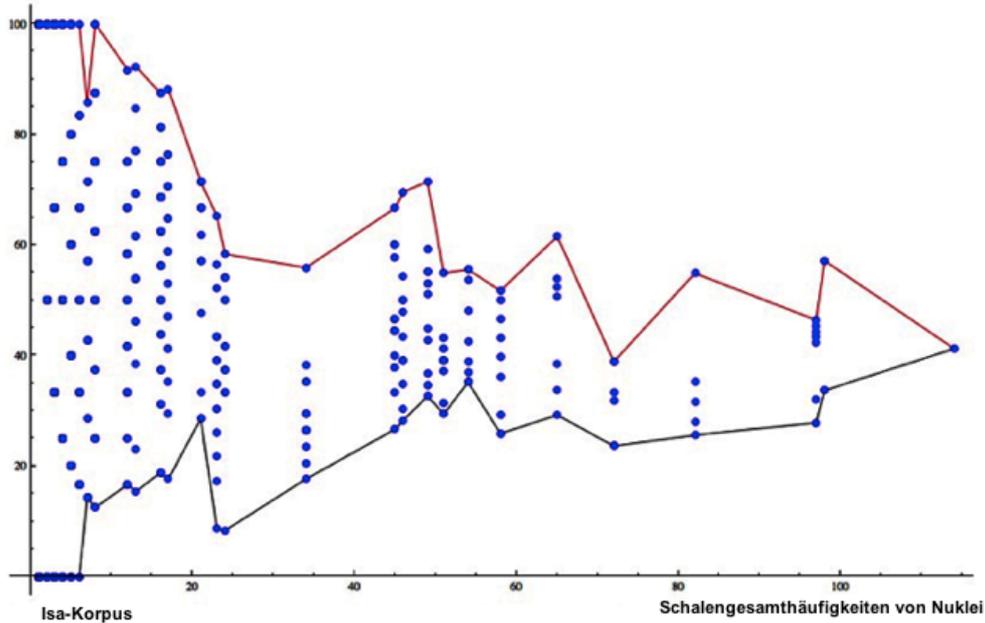


Das analoge Schaubild für das ISA-Korpus bestätigt diese Beobachtungen:

⁸⁹ Es sind Punkte für die Schalenprozentanteile der 30 häufigsten Silbenkerne eingezeichnet, die Max- und Min-Linien beziehen sich aber auf den gesamten Schalengesamthäufigkeitsbereich.

Abbildung 58: Schalengemeinsamkeiten von Silbenkernpaaren (Isa-Korpus)

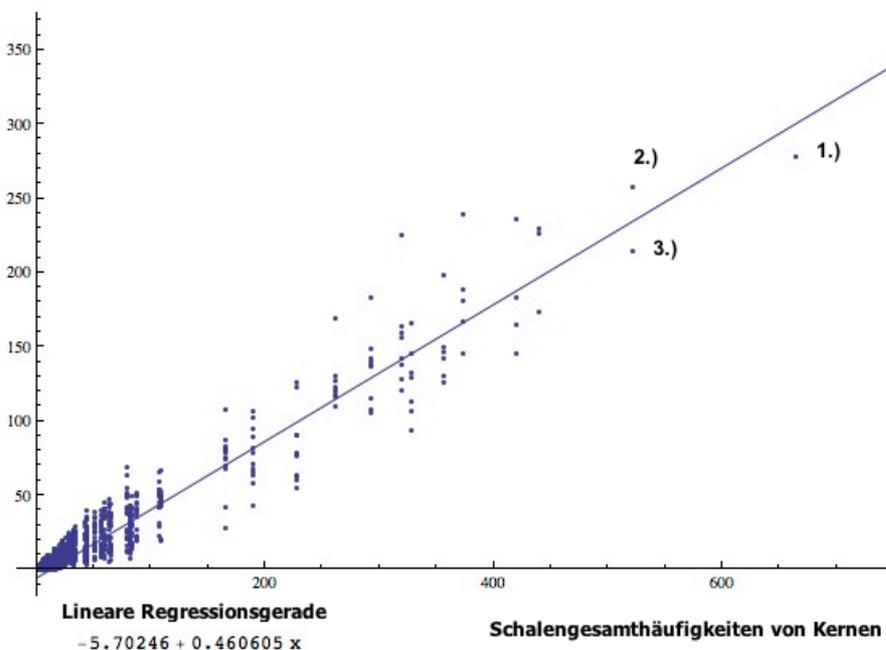
Prozentuale Übereinstimmung gemeinsamer Schalenmengen jedes Nukleus mit jeweils jedem anderen als Bandbreitendarstellung
 (Für jedes Paar von Nuklei der Prozentanteil ihrer gemeinsamen Schalen an der Schalengesamthäufigkeit des selteneren Nukleus des Paares jeweils abgetragen gegen die Schalengesamthäufigkeit des selteneren.)
 (Minima, Maxima, Zwischenwerte)



In der folgenden Graphik sind die gemeinsamen Schalenmengen zweier Silbenkerne statt wie zuvor als Prozentanteil als absolute Zahlen gegen die Schalengesamthäufigkeit des selteneren zweier Silbenkerne abgetragen:

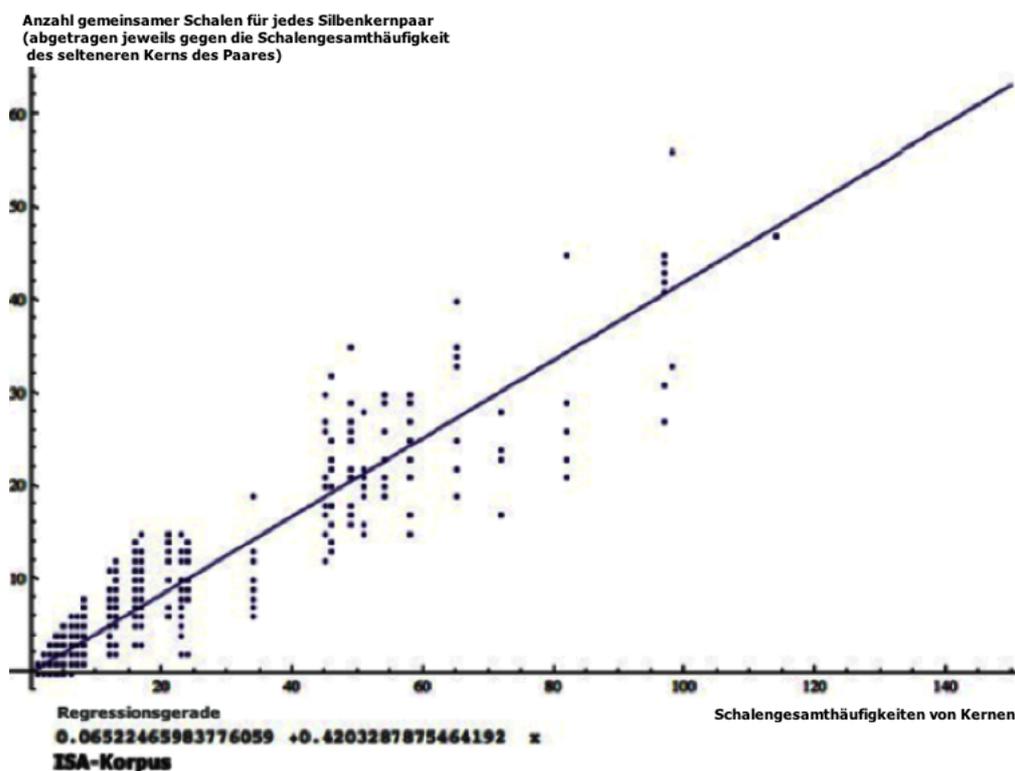
Abbildung 59: Streudiagramm der Schalengemeinsamkeit und des Minimums der Silbengesamthäufigkeiten aller Silbenkernpaare (Annalena-Korpus)

Anzahl gemeinsamer Schalen für jedes Silbenkernpaar
 (abgetragen jeweils gegen die Schalengesamthäufigkeit des selteneren Kerns des Paares)



Die Anzahl der gemeinsamen Schalen zweier Silbenkerne nimmt tendenziell⁹⁰ mit den Schalengesamthäufigkeiten der beiden Silbenkerne zu: So zeigt beispielsweise der mit "1.)" markierte Punkt der Abb. 59 durch seinen y-Achsenwert die Anzahl 278 der Schalen an, die der zweithäufigste Silbenkern /ɪ/ mit Schalengesamthäufigkeit 665 (= x-Achsenwert) mit dem häufigsten Silbenkern /a/ mit einer Gesamthäufigkeit von 830 Schalen gemeinsam hat. Die mit "2.)" und "3.)" markierten Punkte zeigen analog die Anzahl der Schalen an, die der dritthäufigste Silbenkern /ɛ/ (Gesamthäufigkeit 522 Schalen) mit dem häufigsten Silbenkern, dem /a/ (Markierung "2."), nämlich 254, und mit dem zweithäufigsten Silbenkern, dem /ɪ/ (Markierung "3."), nämlich 214, gemeinsam hat. Die anderen Punkte haben eine entsprechende Bedeutung. Die Abb. 60 zeigt die analogen Verhältnisse für das ISA-Korpus:

Abbildung 60: Streudiagramm der Schalengemeinsamkeit und des Minimums der Silben-
gesamthäufigkeiten aller Silbenkernpaare (Isa-Korpus)



Der Unterschied zwischen den Gruppen A und B hinsichtlich der Identität von Schnittmengen wird im Prinzip auch deutlich, wenn man die Silbenkerne entsprechend der Rangfolge ihrer Schalengesamthäufigkeiten progressiv⁹¹ in Gruppen ordnet:

⁹⁰ Die im Schaubild angegebene lineare Regressionsgerade ist nur für Silbenkernpaare mit mindestens 100 gemeinsamen Schalen berechnet worden.

⁹¹ d.h. $S_1, S_2, S_3, \dots, S_i, \dots$ seien die Schalenmengen der Silbenkerne $x_1, x_2, x_3, \dots, x_i, \dots$ mit $|S_j| \geq |S_k|$ für $j \geq k$, ("|" bezeichnet die Mächtigkeit der Mengen) dann werden die progressiven Gruppen G_s aus den Silbenker-

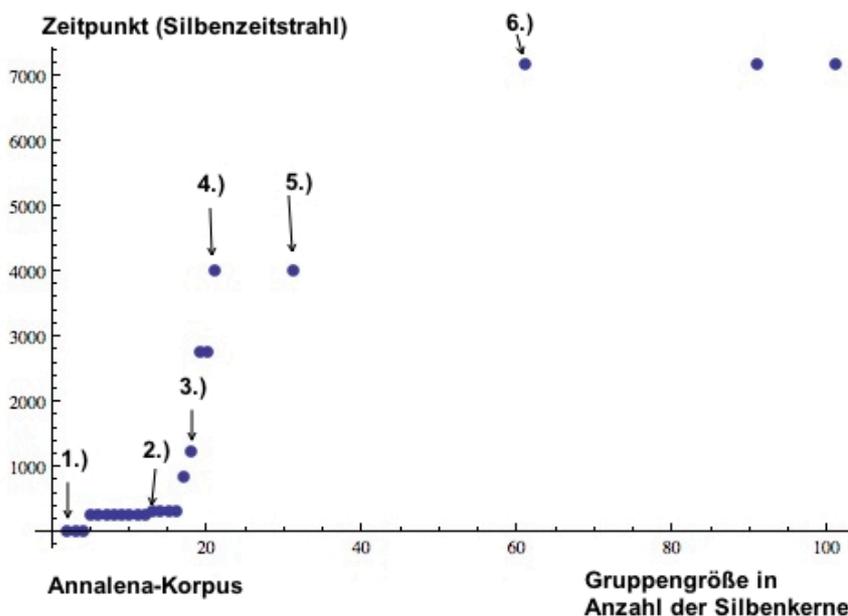
In der Graphik der Abb. 55 wird an der senkrechten Achse des Koordinatensystems der Zeitpunkt angezeigt, von dem ab diese progressiv erweiterten Gruppen keine vollständige Identität von Teilmengen mehr aufweisen. Die Gruppengröße wird hingegen an der waagerechten Achse des Koordinatensystems der Abb. 61 angezeigt. Die Zeitpunkte wurden jedoch nicht für alle Gruppengrößen berechnet, sondern nur für die Gruppen der Größe 1, 2, ..., 20, 21 sowie die Gruppengrößen 31, 61, 91 und 101. Das Ergebnis zeigt einen Zusammenhang zwischen der Gruppengröße und dem Zeitpunkt: Je geringer die Gruppengröße ist, umso früher tritt eine Differenzierung ein, die eine vollständige Übereinstimmung (100%) von Schalenmengen ausschließt. Aus der Abb. 61 geht durch den mit "1.)" gekennzeichneten Punkt (2,16) hervor, dass für die Gruppe der zwei häufigsten Silbenkerne, also der Gruppe aus dem im Annalena-Korpus häufigsten Silbenkern /a/ und dem in diesem Korpus zweithäufigsten Nukleus /ɪ/, eine Identität der Schalenmenge, die /a/ und /ɪ/ gemeinsam haben, mit der Schalenmenge von /ɪ/ ab einschließlich der 16-ten Silbe beendet ist. Betrachtet man nicht nur die im Gesamtzeitraum beiden häufigsten Silbenkerne, sondern nimmt einige der im Gesamtzeitraum nächst häufigeren zu der untersuchten Menge der häufigsten Silbenkerne hinzu, so dauert es tendenziell länger, bis der Zeitpunkt erreicht ist, ab dem keine der zuvor beschriebenen Identitäten mehr existieren. Allgemeiner ist zu beobachten: Je mehr Silbenkerne, beginnend mit den im Gesamtzeitraum häufigsten und der Hinzunahme des jeweils noch nicht in der Menge enthaltenen nächst häufigeren zu einer Gruppe zusammengefasst werden, desto länger dauert es tendenziell bis der Zeitpunkt erreicht ist, ab dem keine der paarweise gemeinsamen Schalenmengen dieser Nukleigruppe keine Identität mehr mit der Schalenmenge des jeweils selteneren Kerns des Paares aufweist: Während für die Gruppe der beiden häufigsten Kerne eine solche Identität bereits ab der 16-ten Silbe nicht mehr vorliegt (der mit "1.)" gekennzeichnete Punkt der Abb. 61), so dauert es für die Gruppe der 13 häufigsten Silbenkerne länger bis zu dem Zeitpunkt, ab dem keine solchen Identitäten mehr auftreten: Sie erreicht diesen Zeitpunkt mit der 327-ten Silbe ("2.")-Markierung), die der 18 häufigsten ab der 1231-ten Silbe (Kennzeichnung "3."), die der 21 häufigsten aber der 4023-ten Silbe ("4.")-Markierung). U.a. an dieser Gruppengröße wird deutlich, dass der Zusammenhang zwischen Gruppengröße und späteren zeitlichen Positionierung des beschriebenen Zeitpunkt nur tendenziell ist: Denn auch die Gruppe der 31 im Annalena-Korpus häufigsten Nuklei weist den Zeitpunkt der 4023-ten Silbe auf (Markierung "5."), ab dem die

nen x_i , $t \geq 1$, als die Vereinigungen $G_s = \cup \{x_d\}$ für alle $s \geq 2$ angesehen ("U" bezeichnet die mengentheoretische Vereinigung), wobei die Vereinigung G_s über alle $\{x_d\}$ mit $1 \leq d \leq s$ läuft.

gemeinsamen Schalenmengen aller Paare der in ihr enthaltenen Silbenkerne keine Identitäten mehr mit der Schalenmenge des selteneren Nukleus des jeweiligen Paares aufweisen. Schließlich sei als letztes erläuterndes Beispiel zur Abb. 61 die Gruppe der 101 häufigsten Silbenkerne genannt, die mit der 7190-ten und damit letzten Silbe des Silbenzeitstrahls zu keinem Zeitpunkt des Silbenzeitstrahls keine Übereinstimmungen der analogen, zuvor beschriebenen Schalenmengen für alle Paare ihrer Silbenkerne erreicht.

Abbildung 61: Streudiagramm des ersten Zeitpunkts des Erreichens und Beibehaltens von Nicht-Inklusionsverhältnissen der Schalenmengen von Silbenkernpaaren und der Größe jeweils von Gruppen der häufigsten Silbenkerne (Annalena-Korpus)

Verhältnis aus der Größe aller entsprechend der Schalengesamthäufigkeitsrangfolge der Kerne progressiv zusammengefassten Gruppen von Kernen (gemessen in der Anzahl der Kerne) und dem Zeitpunkt, ab dem keine Identitäten mehr für alle ihre paarweisen Schalenschnittmengen jeweils mit der Schalengesamtmenge des selteneren Kerns vorliegen

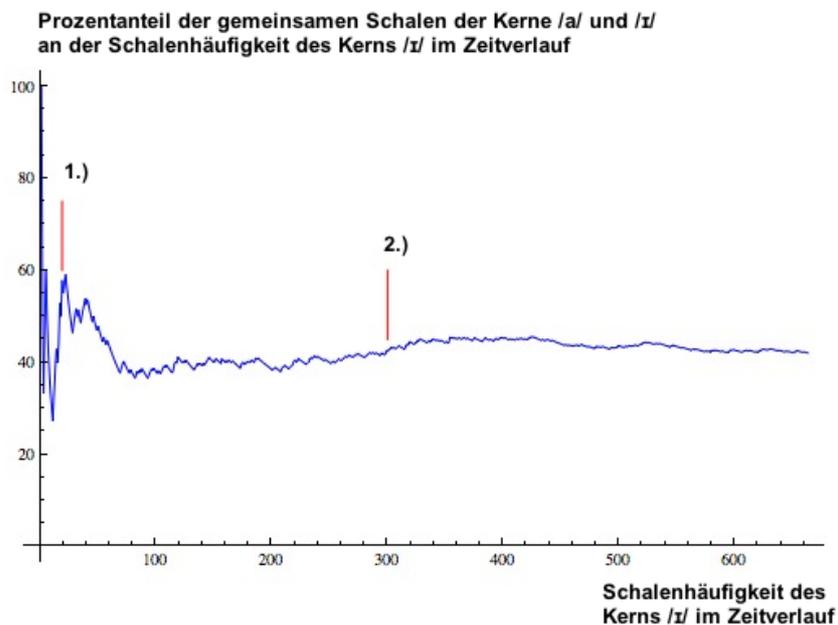


Der Verlauf des Schnittmengenanteils des Silbenkernpaars /a/ und /ɪ/ in der Zeit⁹² ist in der Abb. 62 zu sehen: Der mit "1.)" und einer senkrechten roten Linie gekennzeichnete Wert in Abb. 62 ist das prozentuale Verhältnis aus der Anzahl 11 der Schalen, die zum Zeitpunkt der 267-ten Silbe bereits sowohl in allen Silben mit dem Silbenkern /a/ als auch in allen Silben mit dem Silbenkern /ɪ/ vorgekommen sind (auch als *gemeinsame Schalenmenge* von /a/ und /ɪ/ zum Zeitpunkt der 267-ten Silbe bezeichnet) und der Anzahl 19 aller für den Silbenkern /ɪ/ bereits zu dem Zeitpunkt der 267-ten Silbe aufgetretenen

⁹² abgetragen gegen die im Zeitverlauf auftretende Schalenhäufigkeit des Silbenkerns /ɪ/

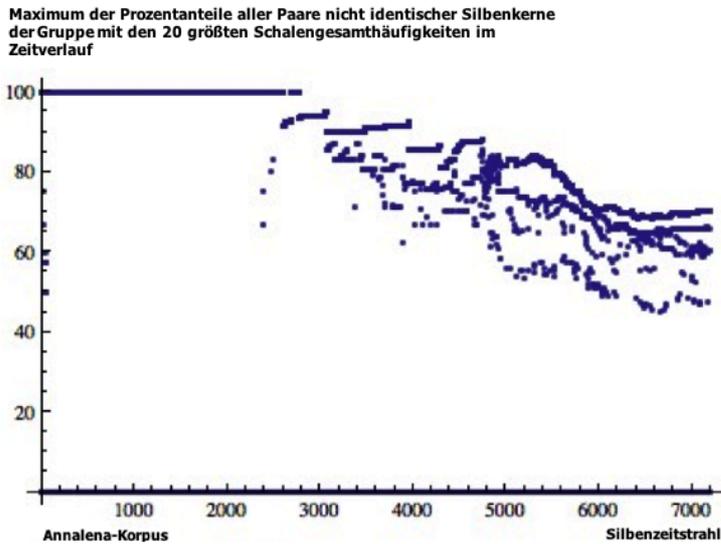
Schalen. Also berechnet sich der Wert der Abb. 62 zum Zeitpunkt der 267-ten Silbe als $(11/19) \cdot 100 = 57,89$. Die "2.)"-Kennzeichnung (mit zusätzlicher roter Linie) verweist auf das Verhältnis $127/300 = 42,33$, wobei 127 die Anzahl aller Schalen in der gemeinsamen Schalenmenge der Silbenkerne /a/ und /ɪ/ zum Zeitpunkt der 2771-ten Silbe angibt und 300 die Anzahl aller bereits in Silben mit dem Nukleus /ɪ/ aufgetretenen unterschiedlichen Schalen.

Abbildung 62: Ausmaß der Schalengemeinsamkeiten der Silbenkerne /a/ und /ɪ/



Die Abb. 63 zeigt für die 20 häufigsten Silbenkerne den Verlauf des Maximums ihrer paarweisen Schalenschnittmengen im Zeitverlauf. In Übereinstimmung mit dem Punkt (20, 2763) der Abb. 61 zeigt nun die Abb. 63 ab dem Silbenzeitpunkt 2763 keine 100-Prozentanteile mehr von paarweisen Schalenschnittmengen an der Schalengesamthäufigkeit des selteneren der beiden.

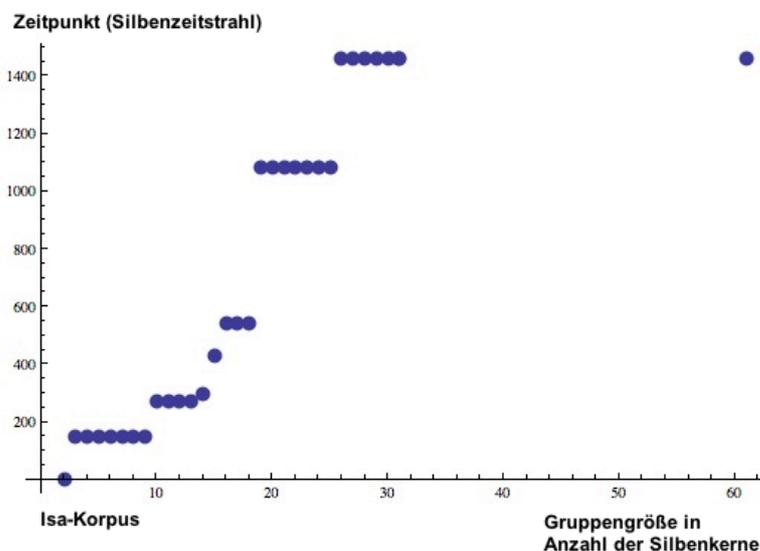
Abbildung 63: Zeitlicher Verlauf der maximalen paarweisen Schalengemeinsamkeiten der 20 häufigsten Silbenkerne



Die zu Abb. 61 analoge Graphik Abb. 64 für das ISA -Korpus bestätigt dieses Ergebnis. Auch hier gilt, dass je früher eine Differenzierung eintritt, die eine vollständige Übereinstimmung (100%) von Schalenmengen ausschließt, die Gruppengröße um so geringer ist (vgl. die zur Abb. 61 des Annalena-Korpus gegebenen Erläuterungen).

Abbildung 64: Streudiagramm des ersten Zeitpunkts des Erreichens und Beibehaltens von Nicht-Inklusionsverhältnissen der Schalenmengen von Silbenkernpaaren und der Größe jeweils von Gruppen der häufigsten Silbenkerne (Isa-Korpus)

Verhältnis aus der Größe aller entsprechend der Schalengesamthäufigkeitsrangfolge der Kerne progressiv zusammengefassten Gruppen von Kernen (gemessen in der Anzahl der Kerne) und dem Zeitpunkt, ab dem keine Identitäten mehr für alle ihre paarweisen Schalenschnittmengen jeweils mit der Schalengesamtmenge des selteneren Kerns vorliegen



7.3.9.5. Zunehmende "Wiederverwendung" von Konstituenten bei der Silbenbildung

Mit dem Silbenbildungsprozess ist auch eine zunehmende Wiederverwendung von Konstituenten verbunden, d.h. die durchschnittliche Wiederverwendung⁹³ ein und derselben Konstituente nimmt zu. Insbesondere nimmt die Zahl der mehrmals verwendeten Konstituenten zu, weshalb sich das Verhältnis aus der Anzahl dieser zu der der einmal verwendeten Konstituenten im Zeitverlauf vergrößert: In der Abb. 65 wird die Menge aller Köpfe, Kodas und Kerne betrachtet, deren Elemente ohne Differenzierung nach Konstituentenarten (d.h. Köpfen, Kodas, Nuklei) als *Konstituenten* bezeichnet werden. In der Abb. 65 ist dann zu jedem Zeitpunkt t des Silbenzeitstrahls das Verhältnis aus der Anzahl der zu diesem Zeitpunkt t bisher schon in mehr als einer Silbe vorgekommenen zu der Anzahl der bis zu diesem Zeitpunkt t in nur einer einzigen Silbe aufgetretenen Konstituenten bei diesem Zeitpunkt abgetragen worden. Der in der Abb. 65 mit "1.)" gekennzeichnete Wert 1.46, der dem Zeitpunkt der 189-ten Silbe entspricht, berechnet sich beispielsweise wie folgt: Von allen bis zur 189-ten Silbe aufgetretenen Konstituenten (d.h. der Menge aller Köpfe, Nuklei und Kodas) sind 57 bereits in mehr als einer Silbe aufgetreten, hingegen 39 in nur einer einzigen Silbe. Der Quotient aus 57 zu 39 ergibt dann den mit "1.)" gekennzeichneten Anteilswert. Der Verlauf dieses Anteilswertes in der Zeit ist dadurch gekennzeichnet, dass er zunächst zwischen lokalen Maximal- und Minimalwerten schwankt, von denen einige mit den Kennzeichnungen "1.)" (Wert 1.46), "2.)" (Wert 0.97), "3.)" (Wert 1.26), "4.)" (Wert 1.02), "5.)" (Wert 1.14) und "6.)" (Wert 0.99) versehen sind, bevor er nach diesem letzten lokalen Minimum "6.)" nicht mehr auf diese Weise oszilliert, sondern einen ansteigenden Trend bis auf den Anteilswert "7.)" (Wert 1.41) zeigt.

⁹³ d.h. die durchschnittliche Anzahl der Silben, in der eine Konstituente vorkommt

Abbildung 65: Verhältnis aus der Anzahl der mehrmals zu der der einmal verwendeten Konstituenten im Zeitverlauf



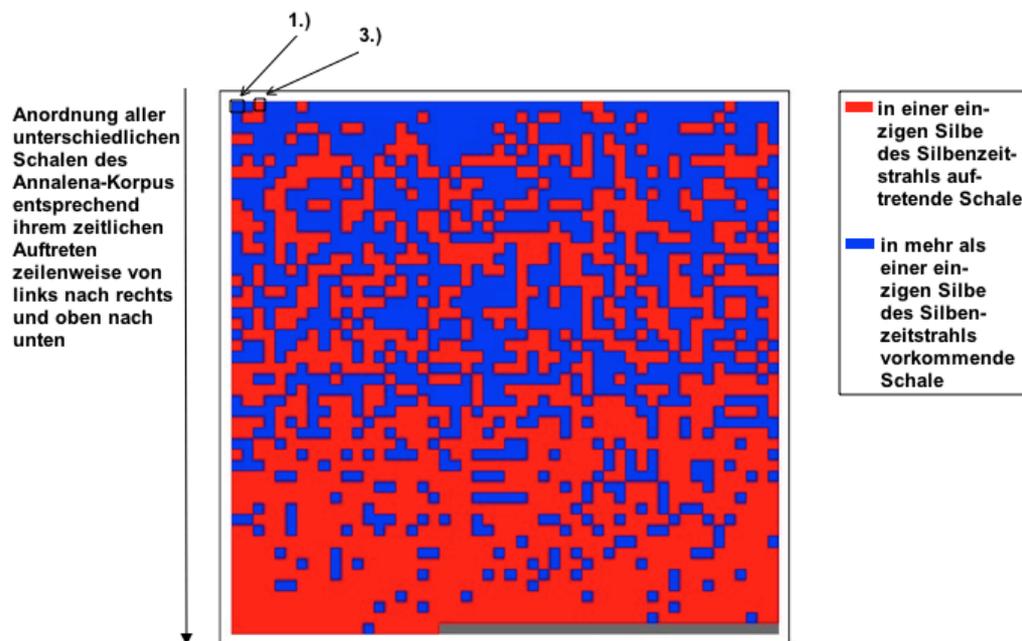
Das Verhältnis der mehrmals verwendeten zu den einmal verwendeten Konstituenten kehrt sich also durch den ab dem mit "6.)" gekennzeichneten Wert zugeordneten Zeitpunkt zu beobachtenden ansteigenden Trend im Zeitverlauf tendenziell um, was als zunehmende Wiederverwendung von Konstituenten in der Silbenmenge aufgefasst werden kann.

7.3.9.6. Zunehmende Streuung bei der Silbenbildung

In der Abb. 66 ist die Menge aller Schalen des gesamten Annalena-Korpus entsprechend des Zeitpunkts des ersten Auftretens der einzelnen Schalen im Zeitverlauf angeordnet, und zwar zeilenweise von links nach rechts und oben nach unten. Dabei entspricht jeder einzelnen Schalen ein rotes oder blaues farbiges Quadrat in einer dieser Zeilen: So ist das mit "1.)", einem Pfeil und einem schwarzen Rahmen markierte blaue Quadrat an der ersten Position in der obersten Zeile positioniert und stellt die erste im gesamten Annalena-Korpus auftretende Schale dar, während das mit "2.)", einem hinweisenden Pfeil und einem schwarzen Rahmen gekennzeichnete rote Quadrat an dritter Position, aber ebenfalls in der ersten Zeile der folgenden Graphik angeordnet ist und somit die als drittes auftretende aller im Annalena-Korpus vorkommender Schalen repräsentiert. Im Gegensatz zur als drittes auftretenden Schale des Annalena-Korpus ("2.)"-Kennzeichnung), die nur in einer einzigen Silbe der Menge aller unterschiedlichen Silben des gesamten Annalena-Korpus vorkommt, was durch die Rotfärbung des Quadrats in der Abb. 66 angezeigt wird, kommt das mit "1.)" markierte Quadrat der Abb. 66 in mehr als einer einzigen aller

unterschiedlichen Silben des Korpus vor und dies wird durch eine Blaufärbung des Quadrats angezeigt. Für die Schalen in der Abb. 66 ist eine zunehmende Streuung zu beobachten, die sich insbesondere in dem starken Überwiegen der roten Quadrate in den Zeilen des unteren Drittels der Abb. 66 zeigt.

Abbildung 66: Zeitliche Anordnung aller Schalen und Darstellung ihrer einmaligen oder mehrfachen Verwendung



7.3.9.7. Silbenkernpaarbasierte Phasenbildung

Verschiedene Beobachtungen legen die Vermutung nahe, dass die Silbenbildung in engen phonologischen und phonetischen Bahnen verläuft, die in verschiedener Hinsicht durch starke Konzentrationen gekennzeichnet sind. Deshalb soll in diesem Abschnitt der Frage nachgegangen werden, inwieweit das Silbenwachstum sich phasenweise als Hauptaktivität zweier Silbenkerne begreifen lässt. Dazu wird der Silbenzeitstrahl in Zeiträume von je 30 Silben eingeteilt und die Silbenkerne für jeden dieser Zeiträume in eine Rangordnung entsprechend ihrer Gesamthäufigkeit in diesem gebracht. Ausgehend von diesen Zeiträumen und den für sie bestimmten Silbenkernrängen soll jeweils für die Silbenkerne sowohl des ersten und zweiten als auch des dritten und vierten Rangs der einzelnen 30-Silbenzeiträume untersucht werden, ob Phasen entstehen, wenn die folgende Regel angewendet wird:

- Gruppierung von 30-Silben-Zeiträumen beginnend mit einem Anfangszeitraum zu Folgen von Zeiträumen, soweit mindestens in jedem zweiten dieser Zeiträume einer der bei-

den Silbenkerne des ersten und zweiten Rangs des Anfangszeitraums den ersten oder zweiten Rang besetzt.

Die unter dieser Voraussetzung entstehenden Folgen von 30-Silben-Zeiträumen werden als *Phase* betrachtet, wenn jeder der beiden häufigsten Silbenkerne des Anfangszeitraums mindestens in einem der Folgezeiträume des Anfangszeitraums wiederholt wird und die entsprechend gebildete Phase wird diesem Silbenkernpaar terminologisch zugeordnet⁹⁴.

Eine analoge Phaseneinteilung wird für die Silbenkerne des dritten und vierten Rangs der 30-Silbenzeiträume angewendet. Die Abb. 67, 68 und 69 zeigen die unter diesen Voraussetzungen gewonnene Phaseneinteilung für die ersten 74 30-Silben-Zeiträume bzw. 2220 Silben des Silbenzeitstrahls. Dabei entsprechen die Spalten den Zeiträumen und die Zeilen den Rängen⁹⁵ der Silbenkerne in den einzelnen 30-Silben-Zeiträumen, wobei die erste Zeile, mit *Beg* gekennzeichnet, den Beginn des jeweiligen Zeitraums⁹⁶ angibt:

Abbildung 67: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 1-720

Beg 1	Beg 31	Beg 61	Beg 91	Beg 121	Beg 151	Beg 181	Beg 211	Beg 241	Beg 271	Beg 301	Beg 331	Beg 361	Beg 391	Beg 421	Beg 451	Beg 481	Beg 511	Beg 541	Beg 571	Beg 601	Beg 631	Beg 661	Beg 691	
a (31)	ə (8)	a (10)	a (7)	ə (4)	a (6)	a (8)	a (8)	a (5)	ə (4)	ε (5)	ε (7)	ɔ (6)	ɔ (6)	ɑ (5)	a (7)	ə (5)	ə (6)	a (7)	ɪ (8)	u (5)	a (5)	a (6)	ɪ (6)	
ai (5)	a (6)	ɪ (4)	ɑ (5)	a (4)	ə (4)	l (3)	ə (7)	au (3)	ɑ (3)	l (5)	a (5)	a (6)	ə (4)	ɪ (5)	ɔ (5)	ε (4)	a (4)	l (3)	ɪ (3)	ə (3)	ɪ (4)	ə (4)	u (3)	
ɪ (5)	ai (5)	e (3)	ɪ (3)	ei (3)	ɔ (4)	ɔu (2)	ɔ (4)	ɔ (3)	u (3)	u (4)	i (4)	ai (5)	ai (4)	au (4)	ɑ (5)	ɔ (4)	ɑ (3)	i (3)	a (3)	au (3)	u (3)	ε (2)	ə (2)	
ə (4)	ε (3)	ə (2)	ɔ (3)	ɔ (3)	l (3)	aua (2)	ɔa (2)	ɑ (2)	a (3)	a (4)	ə (3)	i (3)	u (3)	u (3)	ə (3)	e (4)	e (3)	ia (2)	ɔ (2)	ɔ (3)	ɔ (2)	u (2)	ɑ (2)	
	əɪ (2)	ɔa (1)	ai (2)	ɑ (3)	ɪ (3)	ai (2)	ɑɪ (1)	o (2)	au (2)	ɔ (2)	u (3)	ə (2)	o (3)	a (3)	al (2)	u (2)	u (2)	ea (2)	u (2)	l (3)	ɛ (2)	ɑ (2)	o (2)	
		ɔ (1)	e (2)	ε (2)	ε (2)	ɪ (1)	ε (1)	ɪ (2)	ɬ (2)	æ (2)	ɔ (2)	ɑ (2)	a (3)	ɔ (2)	i (2)	ɪ (2)	ɪ (2)	ea (2)	l (2)	ia (2)	ə (2)	l (2)	i (2)	
		ea (1)	ɔ (1)	i (2)	e (2)	ɔua (1)	ɔu (1)	æ (2)	u (2)	əɔ (1)			ɑ (2)	i (2)	e (2)	i (2)	ε (2)	ai (1)	a (2)	l (2)	ɪ (2)	ɪ (2)	ɪ (2)	
		ɪl (1)	au (1)	e (2)	ε (1)		ɑɔ (1)	u (1)	ə (1)						au (1)	ɔa (1)	ɔɪ (1)	ε (1)	uɪ (1)	e (2)	ai (2)	ɑɪ (2)	a (2)	
		ɔɑ (1)		ɔ (1)	ia (1)		ɜ (1)	ɔ (1)										ei (1)	ɔ (1)	ə (1)	ε (1)	æɪ (1)	ɔ (1)	ɪə (1)
					ɔ (1)		ε (1)	ε (1)										ə (1)		ɪɑ (1)	ɑ (1)	ɔu (1)	ε (1)	
					ə (1)		ɔu (1)	ε (1)										u (1)				o (1)	ε (1)	
					œ (1)		ə (1)	l (1)																

⁹⁴ Beginnen neue Phase während noch nicht abgeschlossener Phasen, so werden sie nur dann genannt, wenn ihr Ende über das der Phasen hinausreicht, in deren Verlauf ihr Beginn liegt.

⁹⁵ Um den Blick auf die quantitativ maßgeblichen Verteilungen zu richten, sind nur die häufigsten Phonyme in einem 30-Silbenzeitraum berücksichtigt worden, die etwa 80 % der Silbenmenge entsprechen.

⁹⁶ in Form des Silbenzeitpunkts, d.h. der Silbe, die dem Anfang der Silbenmenge entspricht

Abbildung 68: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 721-1290

Beg 721	Beg 751	Beg 781	Beg 811	Beg 841	Beg 871	Beg 901	Beg 931	Beg 961	Beg 991	Beg 1021	Beg 1051	Beg 1081	Beg 1111	Beg 1141	Beg 1171	Beg 1201	Beg 1231	Beg 1261
ɪ (7)	ɛ (4)	ʊ (4)	a (6)	ɪ (5)	ɪ (5)	ə (5)	ɪ (8)	ʊ (6)	ə (5)	ɪ (6)	ɪ (9)	a (8)	ɛ (5)	a (6)	ʊ (8)	e (5)	e (4)	a (8)
ɑ (5)	ʊ (4)	ʌ (4)	ʊ (5)	ɑ (4)	ʊ (4)	ɔ (4)	ʊ (7)	ɪ (5)	ɔ (4)	a (6)	a (7)	ɑ (6)	ɪ (4)	ɪ (5)	a (7)	aɪ (5)	ʊɔ (3)	ɪ (6)
ʊ (3)	ɪ (4)	ɪ (4)	i (3)	ʊ (3)	a (4)	ɪ (4)	a (6)	au (4)	ʌ (4)	e (4)	ʌ (3)	ʊ (4)	i (4)	ɛ (2)	ɪ (4)	ɛ (3)	ɔ (3)	ʊ (4)
ɛ (2)	a (3)	ɑ (3)	ɛ (2)	e (3)	ɔ (3)	e (4)	ə (3)	a (4)	i (3)	ɔ (2)	ʊ (2)	ɛ (2)	a (3)	ɔ (2)	ɛ (3)	ɪ (3)	æ (3)	ɑ (3)
au (2)	au (2)	i (3)	ɑ (2)	ə (2)	ɑ (3)	ʊ (2)		ə (3)	ɛ (2)	ʊ (2)	e (2)	i (2)	ɔ (3)	ɑ (2)	i (2)	a (3)	ʊ (3)	ə (2)
a (2)	ɔ (2)	au (2)	ʌ (2)	i (2)	au (2)	i (2)		ɛ (2)	ʊ (2)	o (2)	ɪə (1)	æ (2)	e (3)	ʊ (2)		ʊ (2)	ɪ (3)	ɛ (1)
ɔ (1)	ʌ (2)	a (2)	i (2)	a (2)	i (2)	æ (2)		ɪ (2)	i (2)				ɑ (2)	e (2)		n (2)	ɛ (2)	
ə (1)	i (2)	əɔ (1)	e (2)	ɛ (1)	o (2)	ɛ (1)		o (2)					o (2)			o (2)	au (2)	
ɔʊ (1)	ə (1)	ɛ (1)		ea (1)													i (2)	
				ɔʊ (1)														

Abbildung 69: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 1-4 und den Zeitraum 1291-1800

Beg 1291	Beg 1321	Beg 1351	Beg 1381	Beg 1411	Beg 1441	Beg 1471	Beg 1501	Beg 1531	Beg 1561	Beg 1591	Beg 1621	Beg 1651	Beg 1681	Beg 1711	Beg 1741	Beg 1771
a (5)	a (8)	ɪ (5)	i (5)	a (6)	ɪ (5)	ɔ (3)	i (6)	ɪ (7)	a (7)	a (6)	ɪ (6)	a (7)	a (7)	ɛ (7)	ɪ (8)	ɔ (4)
ɛ (4)	i (4)	i (5)	a (5)	ə (4)	e (5)	ɛ (3)	a (4)	i (6)	ʊ (3)	ʊ (4)	e (4)	i (4)	ɛ (4)	ɪ (4)	ɛ (7)	au (3)
ɔ (4)	e (3)	e (4)	ʊ (4)	ɪ (4)	ə (3)	o (3)	ɛ (3)	au (3)	ɔ (2)	ɛ (3)	o (3)	aɪ (4)	ə (2)	a (4)	i (3)	ɪ (3)
i (4)	o (3)	ə(3)	ɪ (3)	au (2)	ɪ (2)	ʊ (3)	ɪ (3)	e (2)	œɛ (2)	ɪ (3)	ə (2)	e (3)	ʊ (2)	ɑ (3)	aɪ (3)	i (3)
ɪ (3)	ʊ (2)	ʊ (3)	ɔa (2)	ɔ (2)	ʊ (2)	i (3)	ɔ (2)	o (2)	au (2)	e (3)	ʊə (2)	ə (2)	e (2)	e (3)	ɔ (2)	ʊ (2)
ʊ (2)	ɔ (1)	ɑ (3)	æɛ (2)	ɑ (2)	i (2)	a (3)	ɑ (2)	a (2)	ea (2)	i (2)	ea (2)	ʊ (2)	œɛ (1)	ə (2)	ap (2)	i (2)
æ (2)	ɛ (1)	œ (2)	ɛ (2)	e (2)	eɪ (2)	ə (2)	ɔa (1)	ɑɔ (1)	i (2)	i (2)	ɑ (2)	ɸ (1)	ɸ (1)	ɸ (1)		a (2)
	au (1)		ə (1)	ɔa (1)	a (2)	i (2)	eɪ (1)	ɔ (1)	ɪ (2)	ɸ (1)	i (2)	au (1)	ʊə (1)			ɔa (1)
	ɑ (1)			ɛ (1)	ɸ (1)	e (2)	ə (1)		o (2)		œɛ (2)		au (1)			e (1)
							θ (1)						u (1)			θ (1)
													ɸ (1)			ə (1)

In der zweiten und dritten Zeile ist die Phaseneinteilung für den ersten und zweiten Rang der Silbenkerne in den 30-Silben-Zeiträumen farbig markiert wie in der vierten und fünften Zeile die für den dritten und vierten Rang. Für die Markierung der Phasen des jeweiligen Silbenkernpaares werden abwechselnd die Farben in der Reihenfolge Grün, Gelb,

Blau und Rot verwendet, wobei jeweils nur der häufigste Silbenkern der jeweiligen Phase markiert ist. Die Phaseneinteilung für den ersten und zweiten Rang (dargestellt in der zweiten und dritten Zeile) ergibt als erste Phase nach den oben genannten Regeln die des Silbenkernpaars /a/ (grün markiert) und /ə/ und beginnt mit dem 30-Silbenzeitraum, der mit der 31-ten Silbe anfängt. Die zweite Phase (/u/, /ə/) fängt mit der 601-ten Silbe noch während der /a/-/ə/-Phase an und endet mit der 990-ten Silbe, usw. Die Tab. 7 zeigt alle Phasen für den ersten und zweiten Rang im gesamten Zeitverlauf, die sich, wie eben am Beispiel der /a/-/ə/- und /u/-/ə/-Phasen gesehen, überschneiden können, sofern eine Phase während einer anderen beginnt und über deren Ende hinausreicht:

Tabelle 7: Silbenkernpaarbasierte Phaseneinteilung der ersten beiden Ränge

Phasen-Nr.	Phasenzeitraum		Kerne, die zu Beginn einer Phase die beiden häufigsten Silbenkerne darstellen und im Anschluss die Kriterien für die Phasenbildung erfüllen										
	Anfang	Ende	i	ε	a	ai	ɪ	ɨ	ə	o	u	u	ɔ
	1	31	690			x				x			
2	601	1020							x		x		
3	691	1200					x				x		
4	1261	1770			x		x						
5	1291	2460		x	x								
6	2521	2790					x		x				
7	2731	3150		x							x		
8	2911	3270		x									x
9	3121	3360			x						x		
10	3661	3750		x					x				
11	3691	3930	x						x				
12	3691	4110					x	x					
13	4111	4170			x				x				
14	4261	4350		x	x								

			i	ε	a	ai	ɪ	ĩ	ə	o	u	u	ɔ
15	4351	4530						x			x		
16	4441	4560			x						x		
17	4651	4740		x			x						
18	4741	4950								x		x	
19	5011	5190			x						x		
20	5071	5250			x		x						
21	5131	5370					x				x		
22	5191	5520		x			x						
23	5581	5670		x	x								
24	5911	6270			x	x							
25	6091	6480		x						x			
26	6181	6510		x	x								
27	6301	6570	x				x						
28	6421	6840		x			x						
29	6931	7050	x							x			
30	6991	7110		x	x								
			i	ε	a	ai	ɪ	ĩ	ə	o	u	u	ɔ
Gesamtanzahl Phasen jeder Silbenkern			3	12	13	1	10	2	6	3	8	1	1

In Tab. 7 ist jedem Kern, der eine Phase bilden kann, eine Spalte zugeordnet worden. Jeder Phase entspricht eine Zeile, in der die beiden Silbenkerne gekennzeichnet sind, die die Phase bilden. Die analog erstellte Tabelle für die Phaseneinteilung im dritten und vierten Rang ergibt folgende Verteilung:

Tabelle 8: Silbenkernpaarbasierte Phaseneinteilung für die Häufigkeitsränge 3 u. 4

Phasen-Nr.	Phasen-zeitraum		Kerne, die zu Beginn einer Phase die 3-t u. 4-t-häufigsten Silbenkerne darstellen und im Anschluss die Kriterien für die Phasenbildung erfüllen														
	Anfang	Ende	i	e	ε	a	ai	ɪ	ï	ə	ɛ	au	u	o	ʌ	ɔ	ɑ
1	1	120						x		x							
2	271	330				x							x				
3	481	720														x	x
4	661	750			x								x				
5	871	1050				x										x	
6	991	1140	x													x	
7	1081	1290			x								x				
8	1141	1320			x											x	
9	1381	1530						x					x				
10	1741	1890	x				x										
11	1831	2190	x										x				
12	2161	2340						x					x				
13	2431	2580				x			x								
14	3121	3210			x					x							
			i	e	ε	a	ai	ɪ	ï	ə	ɛ	au	u	o	ʌ	ɔ	ɑ
15	3301	3390	x														x
16	3391	3540				x				x							
17	3451	3600								x	x						
18	3781	3930				x									x		
19	3961	4140	x	x													
20	4081	4260			x								x				
21	4561	4800				x			x								
22	5131	5280			x	x											
23	5971	6120			x		x										
24	6091	6270					x	x									
25	6391	6600				x											x
26	6811	6990	x								x						
			i	e	ε	a	ai	ɪ	ï	ə	ɛ	au	u	o	ʌ	ɔ	ɑ
Gesamtanzahl Phasen jeder Silbenkern			6	1	7	8	3	4	2	4	1	1	7	1	1	4	2

Bei der Analyse der Verteilungen fällt auf, dass mit insgesamt 11 unterschiedlichen Silbenkernen nur sehr wenige Kerne überhaupt eine Phase bilden können und dass von diesen 11 Silbenkernen (/ i,ε, a, ai, ɪ, ï, ə, u, u, ɔ/) wiederum eine geringe Anzahl untereinander Kombinationen in einer Phase bildet. Dieser Zusammenhang kann der Tab. 9 entnommen werden, die die Häufigkeit von Phasenbildungen eines Silbenkernpaares⁹⁷ verzeichnet:

⁹⁷ als Kombination aus einer der rechten Spalten- und obersten Zeilen-Laut-Bezeichnungen

Tabelle 9: Häufigkeit der Phasenbildung (1. u. 2. Rang) eines Kernpaares

	a	ɛ	ɪ	ʊ	ə	i	o	ɨ	ai	u	ɔ
a		5	2	3	2				1		
ɛ	5		3	1	1		1				1
ɪ	2	3		2	1	1		1			
ʊ	3	1	2		1			1			
ə	2	1	1	1		1					
i			1		1		1				
o		1	1			1				1	
ɨ				1							
ai	1										
u							1				
ɔ		1									

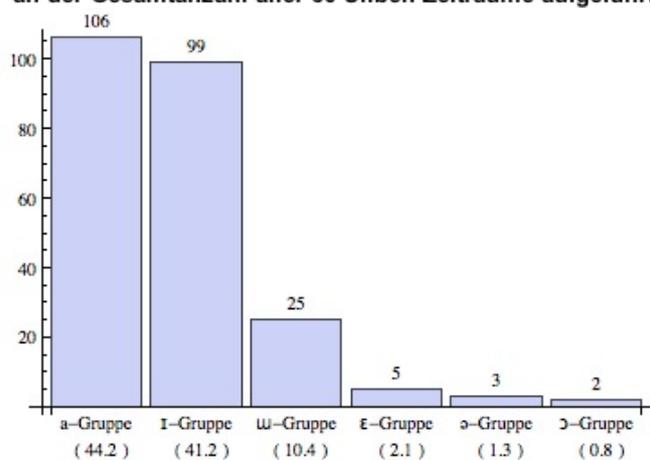
Um eine weitere Bündelung der in den Tab. 7 und 8 der oben gezeigten Phaseneinteilung für die Ränge 1 und 2 bzw. 3 und 4 gegebenen Laute zu erreichen und somit eine möglicherweise anzunehmende Konzentration des Silbenbildungsprozesses auf Hauptsilbenkerngruppenpaare ausfindig zu machen, wurde eine Gruppierung der Phoneme nach angenommener phonetisch-auditiver Ähnlichkeit wie folgt vorgenommen:

Die /a/-artigen Vokale /ɑ, ɐ, æ/ wurden mit dem Vokal /a/ zusammengruppiert und somit ihr Vorkommen quantitativ als Vorkommen einer /a/-Gruppe gerechnet, analog wurden die Phoneme /i, ɨ/ mit dem Vokal /ɪ/ zur /ɪ/-Gruppe zusammengefasst, die Vokale /u, ʊ/ mit dem Vokal /ʊ/ zur /ʊ/-Gruppe, der Vokal /ʌ/ mit dem Vokal /ɔ/ zur /ɔ/-Gruppe, der Vokal /æ/ mit /ɛ/ zur /ɛ/-Gruppe, /ø/ mit /œ/ zur /œ/-Gruppe und /ə/ mit /ə/ zur /ə/-Gruppe. Außerdem wurden die Laute mehrphonemiger Silbenkerne quantitativ anteilmäßig den jeweils in ihnen enthaltenen Phonemen zugerechnet: Beispielsweise wurden für den Diphthong /ai/ der Gesamthäufigkeit des Vokals /a/ eine Häufigkeit von 0,5 hinzuaddiert und ebenso dem Vokal /ɪ/ eine gleichgroße Häufigkeit (0,5). Im Anschluss an diese Gruppierung wurde erneut eine Phaseneinteilung nach dem oben beschriebenen Kriterium für die Ränge 1 und 2 durchgeführt: Diese ergab eine einzige Phase, die aus einem Wechsel der /a/-Vokalgruppe mit der /ɪ/-Vokalgruppe besteht. Dies kann so verstanden werden, dass die /a/- und /ɪ/- Vokalgruppen nach dem obigen Phasenkriterium in einem

ständigen gegenseitigen Verhältnis stehen. Aber auch wenn nach dieser phonetisch-auditiven Gruppierung der Silbenkerne die /a/- und /ɪ/-Gruppe zusammen eine einzige Phase bilden können, so werden die Ränge 1- 4 auch von weiteren Silbenkern(grupp)en besetzt. Deshalb soll hier die Häufigkeitsverteilung der Silbenkerngruppen für jeden der ersten vier Ränge dargestellt werden:

Abbildung 70: Häufigkeitsverteilung der Silbenkerngruppen auf dem 1-ten Rang

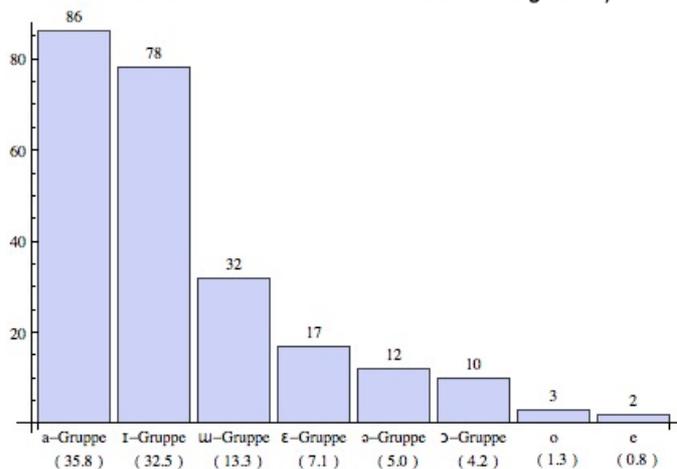
Häufigkeitsverteilung der Silbenkern(grupp)e(n) auf dem 1-ten Rang in jedem der 30-Silbenzeiträume der Unterteilung des Silbenzeitstrahls
(Unter der Kern(gruppen)-Bezeichnung ist in Klammern der Anteil der Häufigkeit der Kerngruppe bzw. des Kerns an der Gesamtanzahl aller 30-Silben-Zeiträume aufgeführt.)



Annalena-Korpus

Abbildung 71: Häufigkeitsverteilung der Silbenkerngruppen auf dem 2-ten Rang

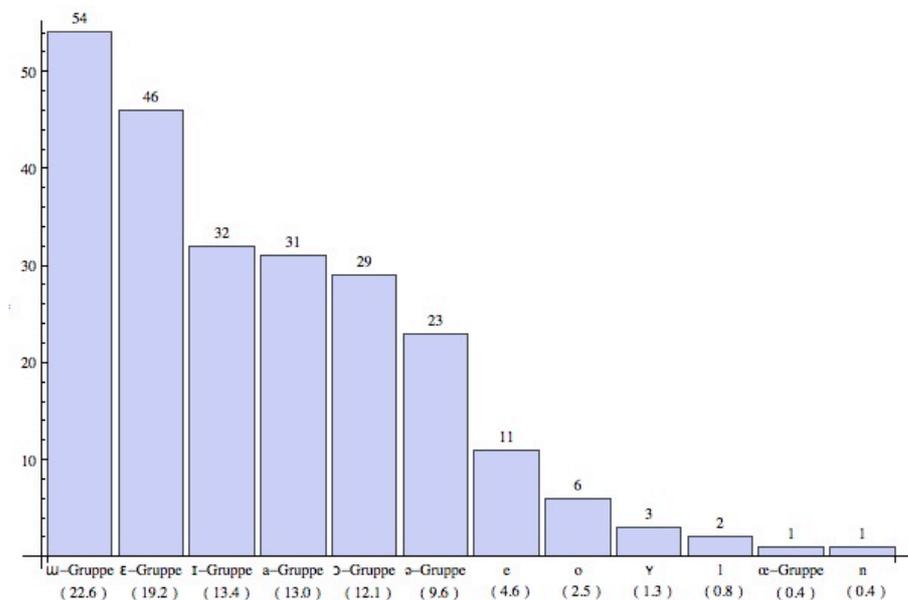
Häufigkeitsverteilung der Silbenkern(grupp)e(n) auf dem 2-ten Rang in jedem der 30-Silbenzeiträume der Unterteilung des Silbenzeitstrahls
(Unter der Kern(gruppen)-Bezeichnung ist in Klammern der Anteil der Häufigkeit der Kerngruppe bzw. des Kerns an der Gesamtanzahl aller 30-Silben-Zeiträume aufgeführt.)



Annalena-Korpus

Abbildung 72: Häufigkeitsverteilung der Silbenkerngruppen auf dem 3-ten Rang

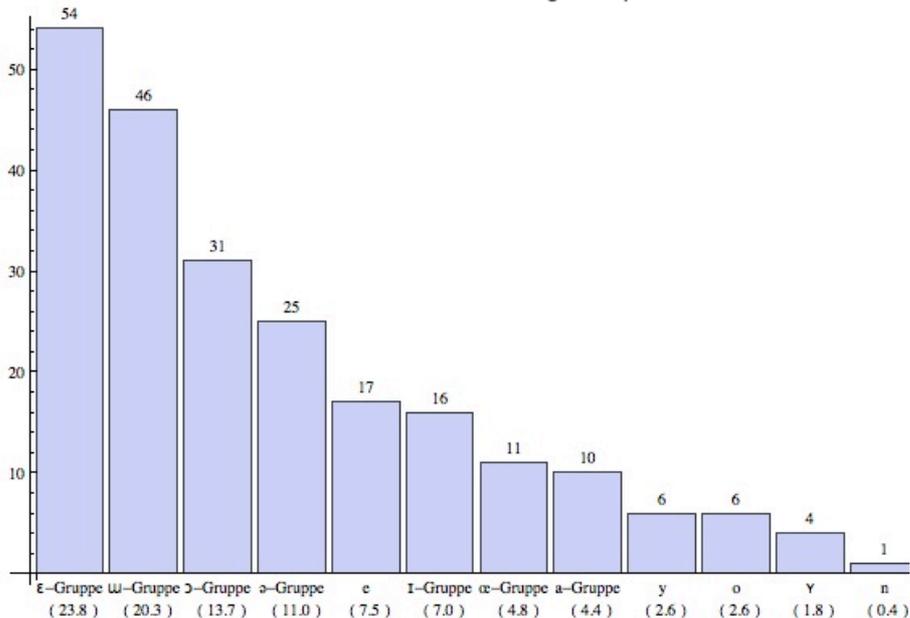
Häufigkeitsverteilung der Silbenkern(grupp)e(n) auf dem 3-ten Rang in jedem der 30-Silbenzeiträume der Unterteilung des Silbenzeitstrahls
(Unter der Kern(gruppen)-Bezeichnung ist in Klammern der Anteil der Häufigkeit der Kerngruppe bzw. des Kerns an der Gesamtanzahl aller 30-Silben-Zeiträume aufgeführt.)



Annalena-Korpus

Abbildung 73: Häufigkeitsverteilung der Silbenkerngruppen auf dem 4-ten Rang

Häufigkeitsverteilung der Silbenkern(grupp)e(n) auf dem 4-ten Rang in jedem der 30-Silbenzeiträume der Unterteilung des Silbenzeitstrahls
(Unter der Kern(gruppen)-Bezeichnung ist in Klammern der Anteil der Häufigkeit der Kerngruppe bzw. des Kerns an der Gesamtanzahl aller 30-Silben-Zeiträume aufgeführt.)



Annalena-Korpus

Sowohl auf dem ersten als auch dem zweiten Rang der 30-Silben-Zeiträume dominieren die /a/- und die /ɪ/-Gruppe, d.h. sie überwiegen anteilmäßig. Diesen Rängen folgen jeweils die /u/- und /ɛ/-Gruppe, die auf den Rängen 3 und 4 dominieren. So erhält man eine Menge von vier Kerngruppen, deren Elemente zusammen oft auch Phasen bilden und somit für die zeitliche Interaktion der Silbenkerne relevant zu sein scheinen. Die häufige⁹⁸ Bildung von Phasen für Paare aus Kernen einer der vier /a/-, /ɪ/-, /u/- oder /ɛ/-Silbenkerngruppen deutet möglicherweise auch auf eine zeitliche Konzentration auf auditive Vokalkontraste hin.

7.3.9.8. Die Bildung von Köpfen und Kodas unter dem Gesichtspunkt phonetischer Merkmale und in Bezug auf die Silbenkernentwicklung

Da menschliche Sprache sich eines spezifischen Instruments bedient, beeinflusst und bedingt seine Ausgestaltung und Anwendung die innere Struktur und äußere Wahrnehmbarkeit der erzeugten Laute, ohne dass jedoch dieses Werkzeug als autonomer Verursacher der erwähnten Strukturen angesehen wird. Vielmehr soll die Frage geprüft werden, wie sich die im Silbenkorpus realisierte Phonemverteilung unter phonetischen Aspekten ordnet, wobei angenommen wird, dass systematische Beziehungen zwischen den phonetischen Möglichkeiten der Artikulationsorgane und der inneren Entwicklung des Silben- und Phonemsystems bestehen. Dazu werden die Verteilungen der konsonantischen Silbenkonstituenten im Zeitverlauf in den Köpfen und Kodas der Silben nach den Merkmalen der Artikulationsart, Stimmhaftigkeit bzw. -losigkeit und dem Artikulationsort betrachtet und zusätzlich für die drei bereits in Abschnitt 7.3.9.7 gebildeten /a/-, /ɪ/- und /u/-Vokalgruppen differenziert:

⁹⁸ Diese Häufigkeitsaussage bezieht sich jeweils auf eine der Silbenkerngruppen, die auf mindestens einem der für die Ränge 1 bis 4 (vgl. Abb. 70 - 73) dargestellten Häufigkeitsverteilungen eine Häufigkeit von der Größe haben, dass diese bei der sukzessiven Summation der Häufigkeiten der Gruppen in einer der Abb. 70, 71, 72 oder 73, beginnend mit der Häufigkeit der häufigsten Gruppe, noch zur Überschreitung der 50%-Anteilsmarke der bereits gebildeten Summe benötigt werden. So werden in der Abb. 70 für den Rang 1 nur die Häufigkeiten der beiden auf diesem Rang (1) häufigsten Silbenkerngruppen zur Überschreitung der 50% Marke benötigt; in der Abb. 72 für den dritten Rang ist hingegen die Summe (41.8%) der beiden häufigsten Nuklei-Gruppen, der häufigsten Gruppe, der /u/-Gruppe (22.6%-Anteil) und der zweithäufigsten Gruppe, der /ɛ/-Gruppe (19.2-Prozentanteil) noch nicht größer als 50%, so dass noch der Anteil (13.4%) der auf dem zweiten Rang (Abb. 71) dritthäufigsten Nukleus-Gruppe, der /ɪ/-Gruppe, hinzugenommen werden muss, nicht aber auch noch der der vierthäufigsten (d.h. der /a/-Gruppe), weil bereits mit der dritthäufigsten eine Summe der Prozentanteile von 55.2% erreicht ist. Für keinen der ersten vier Ränge wird bei dieser Summenbildung der Anteil der /ə/-Silbenkerngruppe benötigt, so dass bei der getroffenen Aussage beispielsweise die /ə/-Silbenkerngruppe auf Grund dieses Häufigkeitskriteriums nicht erwähnt wird.

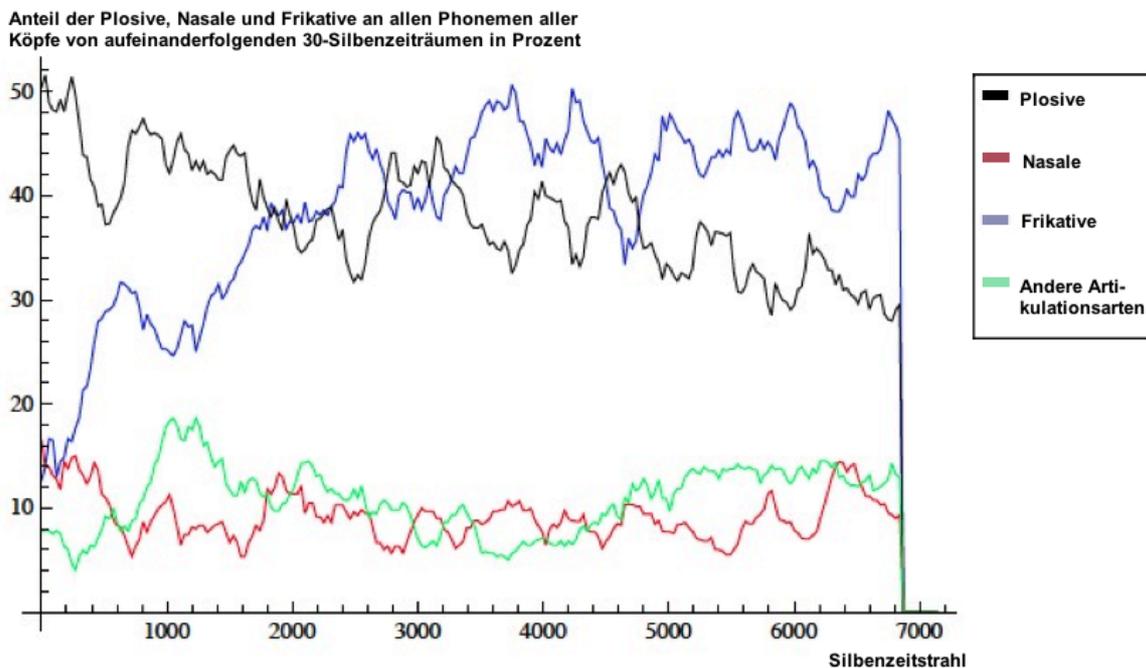
D.h. es erfolgt eine Gruppierung der Vokalphoneme nach angenommener phonetisch-auditiver Ähnlichkeit wie folgt: Die /a/-artigen Vokale /ɑ, ɐ, æ/ werden mit dem Vokal /a/ zusammengruppiert und somit ihr Vorkommen quantitativ als Vorkommen einer /a/-Gruppe gerechnet, analog werden die Phoneme /i, i/ mit dem Vokal /i/ zur /i/-Gruppe zusammengefasst und die Vokale /u,ʊ/ mit dem Vokal /u/ zur /u/-Gruppe. Diese drei Vokalgruppen sind ausgewählt worden, da sie in den ersten drei Vokalrängen dominieren (vgl. Abschnitt 7.3.9.7.). Die Abb. 74⁹⁹ zeigt die Verläufe der Anteile für die Phoneme aller Köpfe in je aufeinanderfolgenden Zeiträumen von 30 Silben der Artikulationsarten der Plosive (schwarz), Frikative (blau) und Nasale (rot)¹⁰⁰. Der grüne Verlauf entspricht der zusammenfassenden Darstellung der Anteile aller verbleibenden Artikulationsartenkategorien, die nicht weiter spezifiziert werden. Es zeigt sich, dass für die Phoneme der Köpfe zunächst die Plosive überwiegen, deren Anteil im Zeitverlauf abnimmt, während der der Frikative zunimmt¹⁰¹. Zudem ist zeitweise ein gegenläufiges Verhalten für die drei Artikulationsartenpaare der Frikative, Plosive und Nasale zu beobachten.

⁹⁹ Moving Average 10

¹⁰⁰ Phoneme von Kopf- oder Kodakonstituenten aus x Phonemen werden hinsichtlich der ihnen zugeordneten phonetischen Merkmale mit der Häufigkeit 1/x verrechnet. Dies wurde für alle entsprechenden Schaubilder dieses Abschnitts angewandt. Die Angabe der Anteile erfolgt in allen Abbildungen in Prozent.

¹⁰¹ Ein Vergleich der quantitativen Anteile von Plosiven und Frikativen im Deutschen anhand des Kieler Korpus (Kohler 1995: 222) mit insgesamt 23985 Wörtern (ebd.) könnte die Frage aufkommen lassen, ob die dargestellten Graphen aussagekräftig sind. Denn die Abb. 74 - 81 vermitteln den Eindruck, dass die Plosive summiert über den Gesamtzeitraum nicht unbedingt häufiger sind als die über den Gesamtzeitraum summierten Frikative oder dass evtl. sogar das umgekehrte Verhältnis gilt. Jedoch beträgt der Gesamtanteil der Plosive für die in Kohler (1995: 222) in Tabelle 3.a.) angegebenen Anteilswerte 36,61 % und 24,51 % für die Frikative, womit die Plosive erheblich häufiger als die Frikative sind. Ein Unterschied zu Kohlers (1995) Häufigkeitsverteilung (ebd.: 222) ist aber die im Kieler Korpus (ebd.) betrachtete Häufigkeit der Lautklassen in der Menge aller, auch wiederholten Silben in 23985 Wörtern des Deutschen. Somit wird dort die Häufigkeit von Lauten hinsichtlich der sog. token-Frequenz von Silben in dieser Wortmenge gezählt, während in der vorliegenden Untersuchung die nach Artikulationsartmerkmalen differenzierte Lauthäufigkeit hinsichtlich der sog. type-Frequenz von Silben betrachtet wird. Zudem darf man nicht vergessen, dass während der kindlichen Entwicklung andere quantitative Verteilung herrschen können,- v.a. in der kumulativen Aufsummierung,- als in der Zielsprache, (vgl. dazu auch die in Lindner (2009) festgestellten Differenzen zwischen der Sprache der untersuchten Kinder- und der Zielsprache hinsichtlich der Phonotaktik, Sonoritätshierarchie etc. des Deutschen). Es geht ja in dieser Arbeit überhaupt nicht um die Betrachtung einzelner Laute, sondern immer um ihren globalen Zusammenhang im System. Wenn Werte für einzelne Laute isoliert werden, bedeutet das nicht, dass sie in den Mittelpunkt gerückt werden (können).

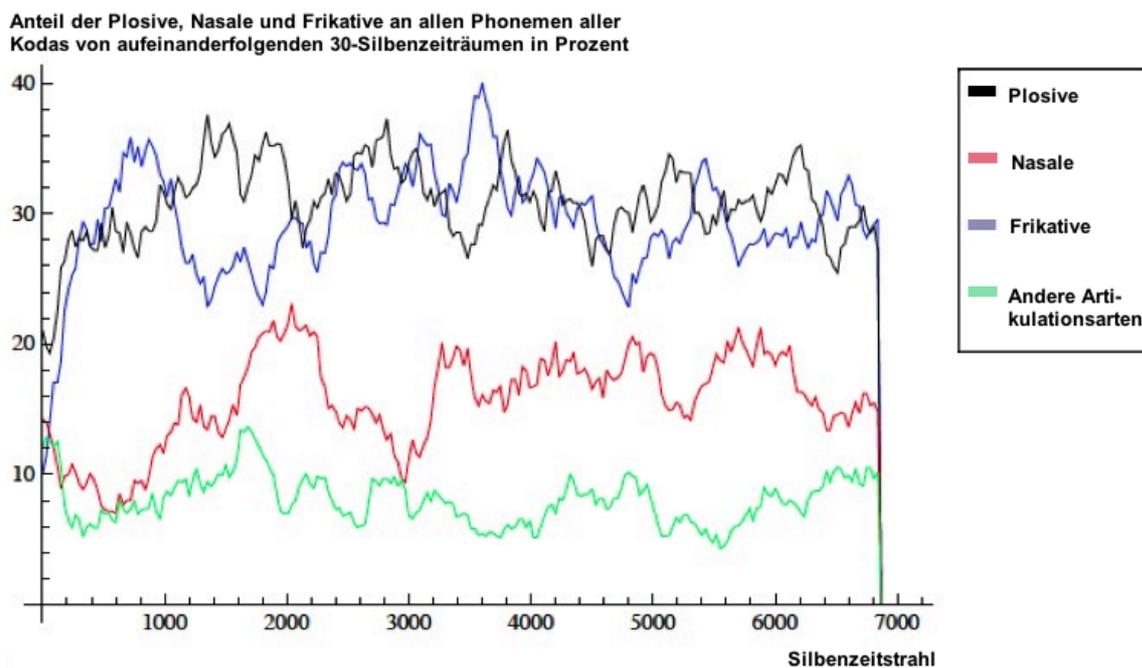
Abbildung 74: Differenzierung der Köpfe nach Artikulationsarten im Zeitverlauf



Die Abb. 75 stellt analog die Verteilung für die Kodas dar.

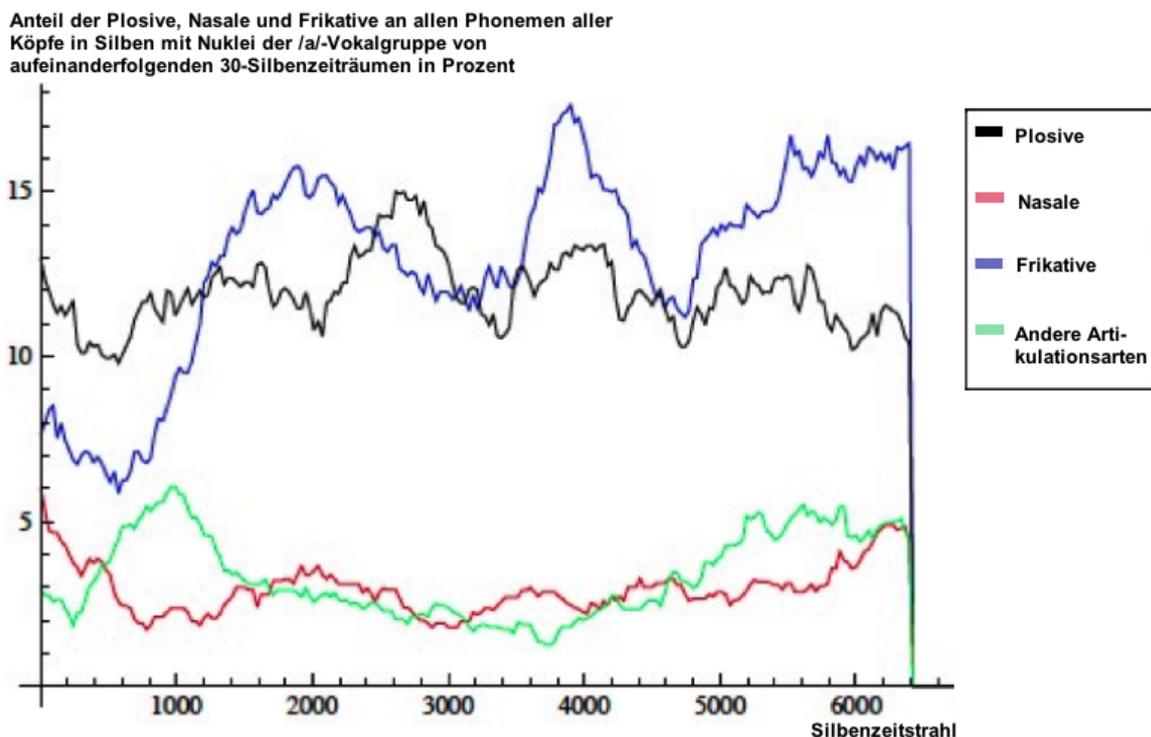
Auch hier ist eine zeitweise Gegenläufigkeit der Konstituentenartenpaare zu beobachten, aber nicht eine anfängliche Trennung der Frikativ- und Plosiv-Anteile wie für die Köpfe.

Abbildung 75: Differenzierung der Kodas nach Artikulationsarten im Zeitverlauf



Die Abb. 76 - Abb. 81¹⁰² zeigen die Differenzierung der Artikulationsartenanteile für die Köpfe und Kodas jeweils nach den /a/-, /ɪ/- und /ʊ/-Vokalgruppen: Die Abbildungen für die Köpfe weisen zum Beginn der Silbenentwicklung eine Trennung der Plosiv- und Frikativ-Anteile auf, bei der die Plosiv- über den Frikativ-Anteilen liegen. Die Trennung ist bei den Kodas nur für die /a/-Vokalgruppe vergleichbar stark und erfolgt für alle Vokalgruppen auf genau umgekehrte Weise, d.h. der Anteil der Frikative liegt über dem der Plosive. Während für die Köpfe für die /a/- und /ɪ/-Vokalgruppen gegen Ende der Silbenentwicklung wiederum eine stärkere Trennung der Plosiv- und Frikativ - Anteile erfolgt, so dass nun ihr Rangverhältnis umgekehrt zum Anfang der Entwicklung ist, ist solch eine Spreizung zum Ende des Silbenwachstums für keine der Vokalgruppen der Koda zu beobachten. Generell ist die Tendenz zu beobachten, dass einige phonetische Merkmalsgruppen bei den Köpfen sich gegenläufig entwickeln, was für die Kodas nur phasenweise gilt, wobei hier eher eine begrenzte Parallelität vorzufinden ist.

Abbildung 76: Differenzierung der Köpfe aller /a/-Silben nach Artikulationsarten im Zeitverlauf



¹⁰² jeweils Moving Average 25

Abbildung 77: Differenzierung der Kodas aller /a/-Silben nach Artikulationsarten im Zeitverlauf

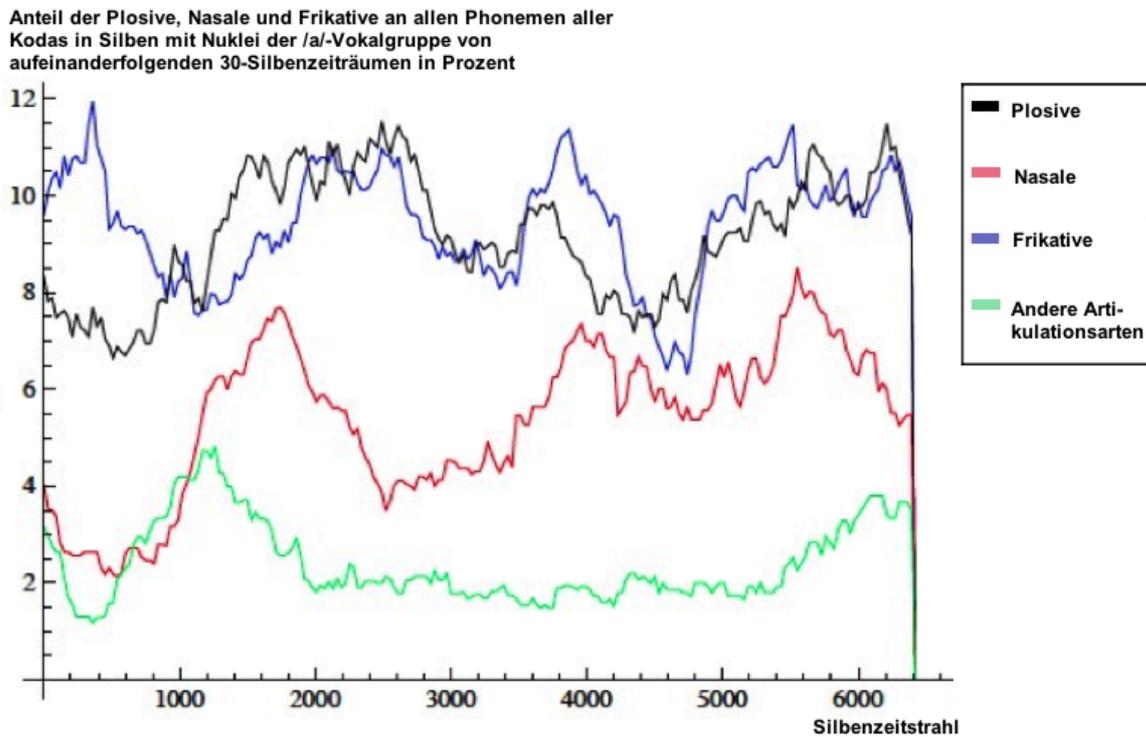


Abbildung 78: Differenzierung der Köpfe aller /ɪ/-Silben nach Artikulationsarten im Zeitverlauf

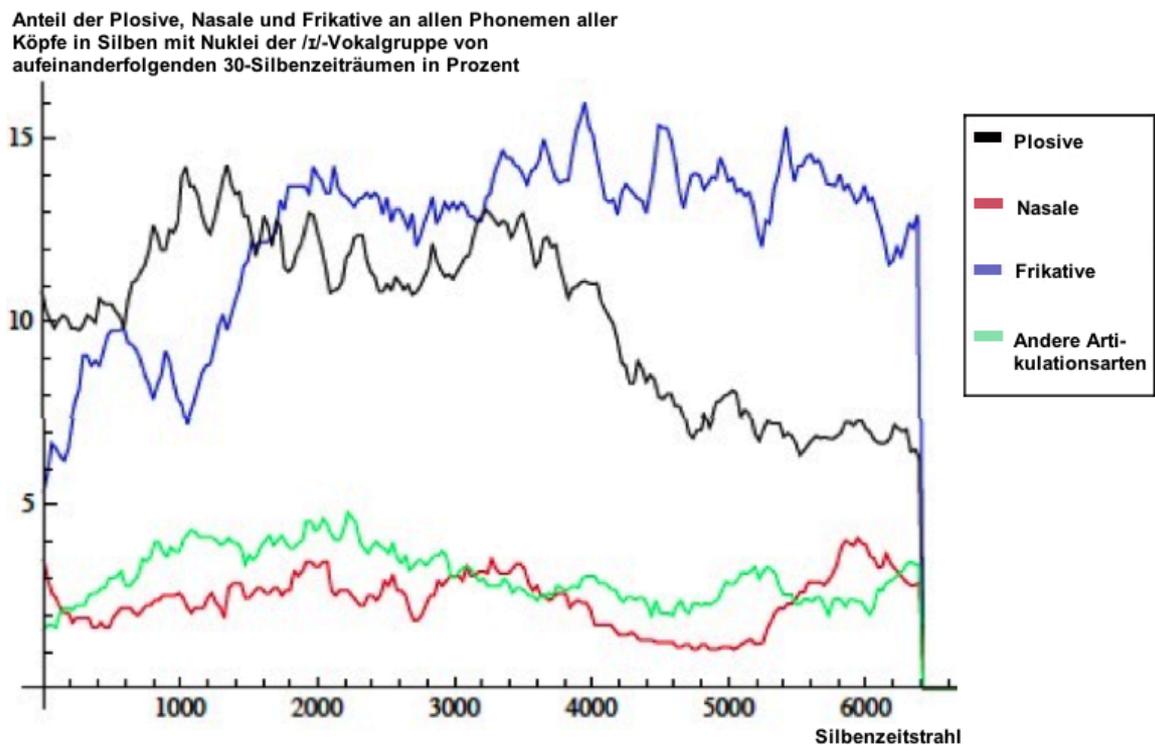


Abbildung 79: Differenzierung der Kodas aller /ɪ/-Silben nach Artikulationsarten im Zeitverlauf

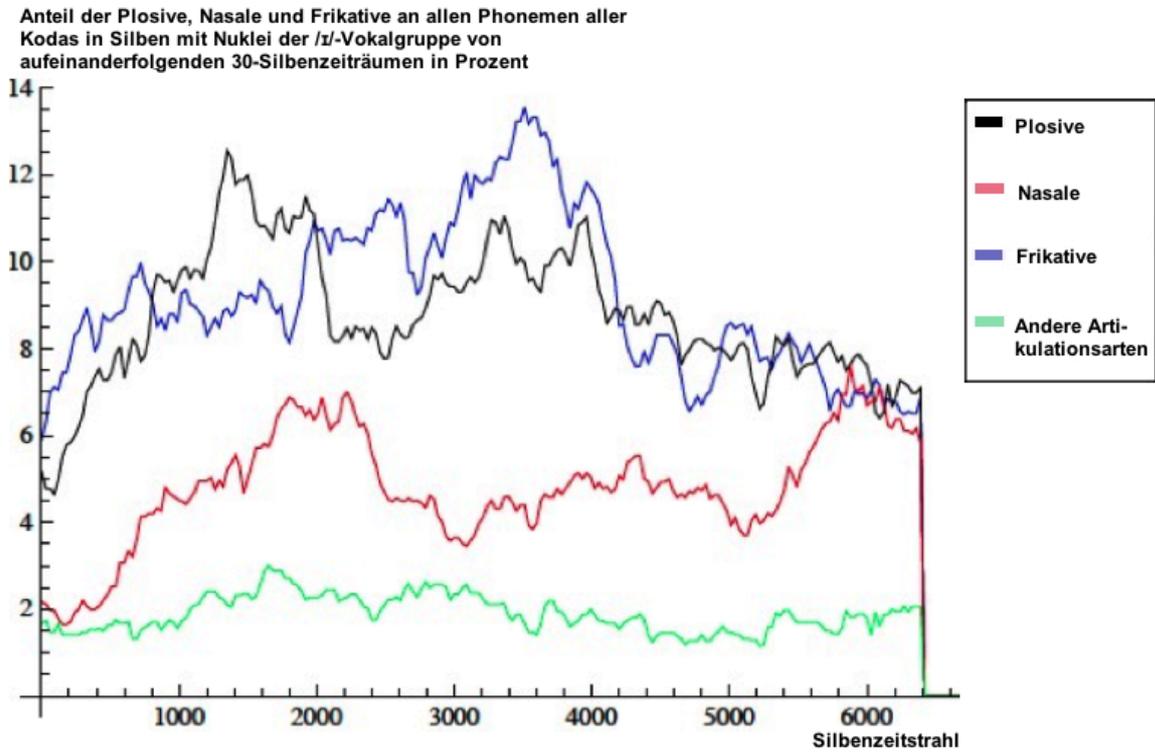


Abbildung 80: Differenzierung der Köpfe aller /u/-Silben nach Artikulationsarten im Zeitverlauf

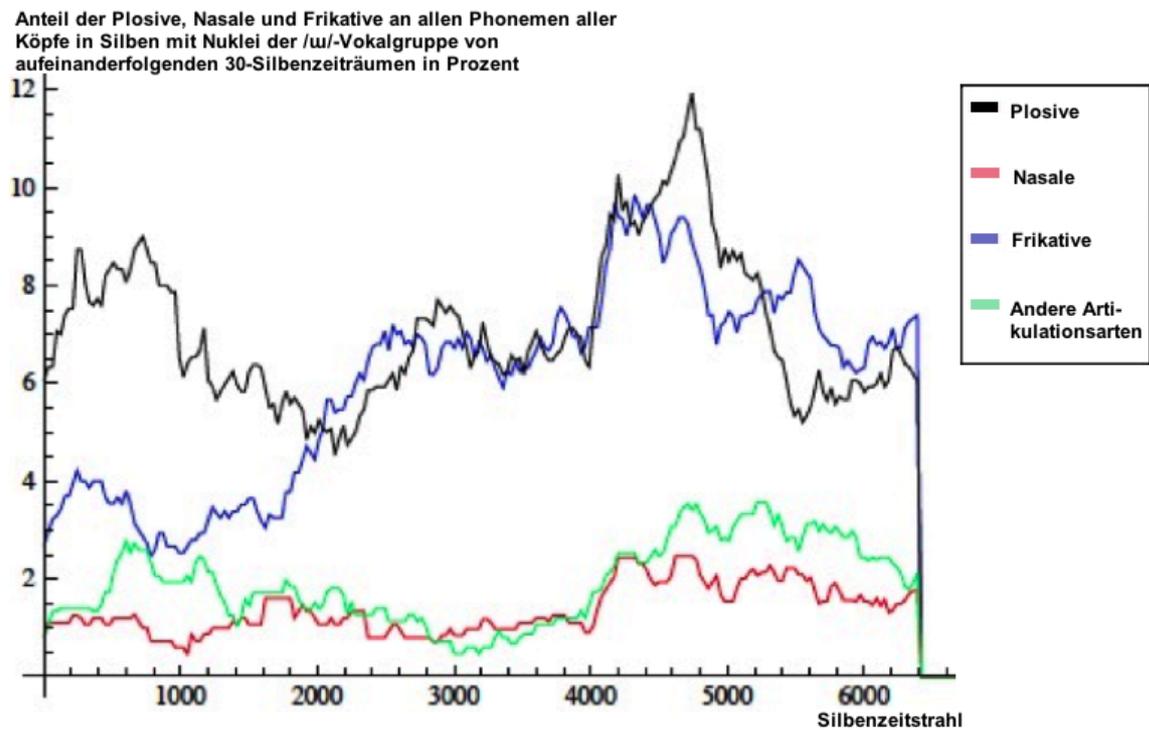
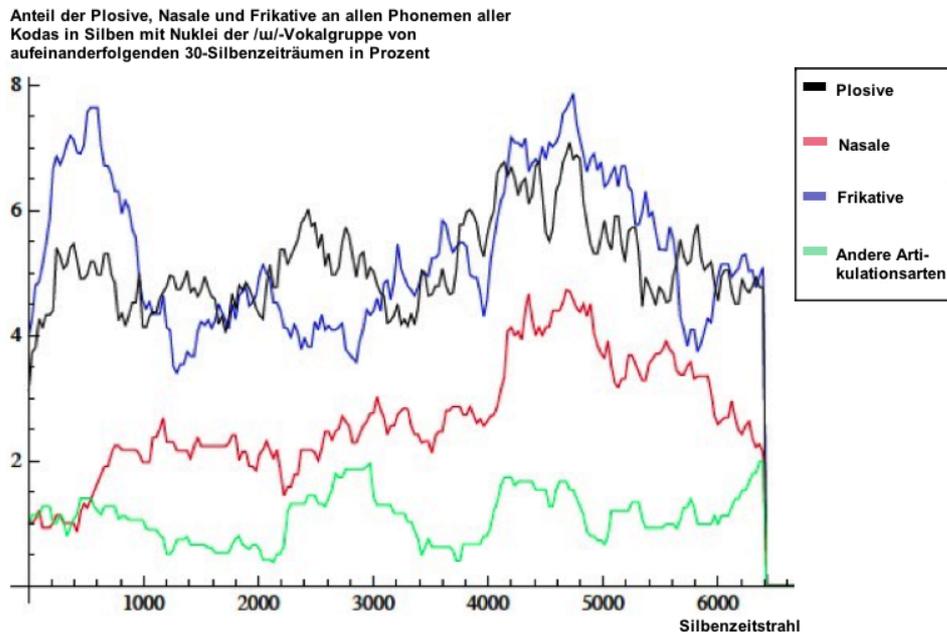


Abbildung 81: Differenzierung der Kodas aller /u/-Silben nach Artikulationsarten im Zeitverlauf



Die Abb. 82 und Abb. 83 stellen jeweils die Anteile der stimmhaften (schwarz) und -losen (rot) Phoneme aller Köpfe in jeweils aufeinanderfolgenden 30-Silbenzeiträumen (Moving Average 10) dar. Sowohl für den Kopf als auch die Koda ist ein gegenläufiges Verhalten der beiden Anteile zu beobachten. Für die Köpfe bewegen sich diese aber aufeinander zu, wobei die Stimmhaftigkeitsanteile ab- und die Stimmlosigkeitsanteile zunehmen, während sich diese beiden Anteile für die Kodas mehr auf gleichem Niveau im Zeitverlauf befinden.

Abbildung 82: Differenzierung der Köpfe nach (Nicht-)Stimmhaftigkeit im Zeitverlauf

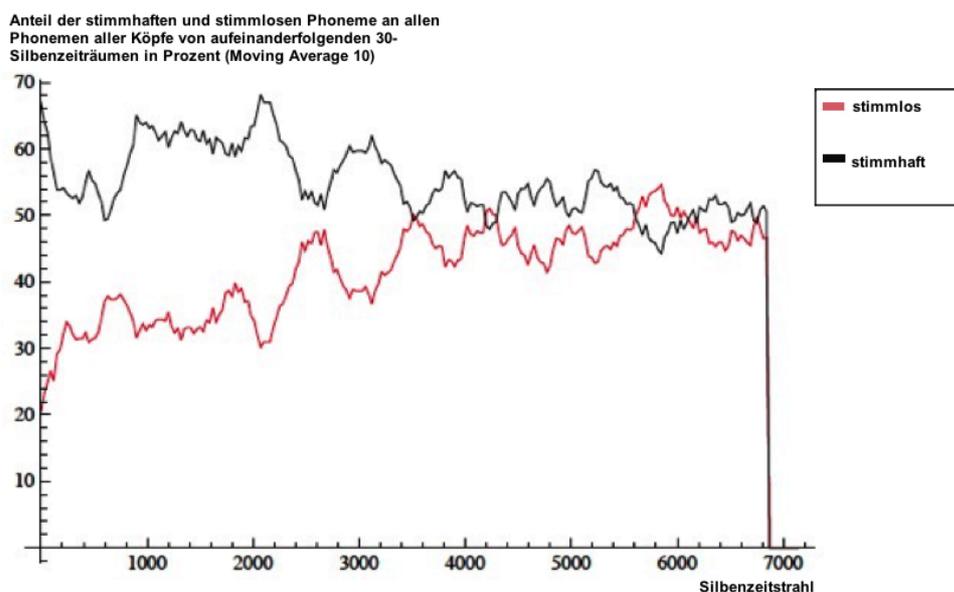
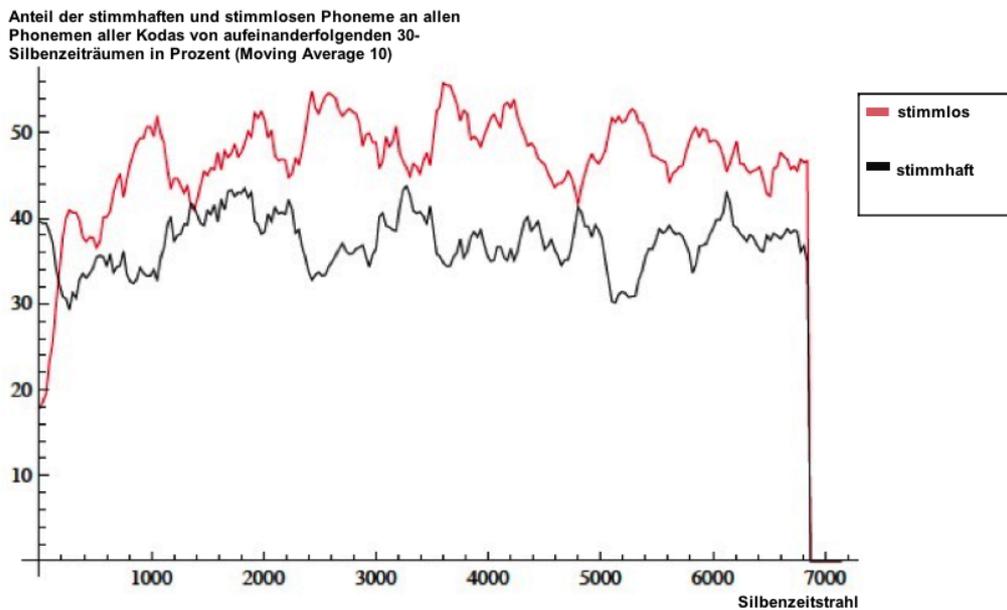


Abbildung 83: Differenzierung der Kodas nach (Nicht-)Stimmhaftigkeit im Zeitverlauf



Zudem liegen die Stimmhaftigkeitsanteile für die Köpfe über den Stimmlosigkeitsanteilen, während ein umgekehrtes Verhältnis für die Kodas gilt.

Dieses umgekehrte Verhältnis der beiden Anteile für die Köpfe und Kodas ist auch nahezu durchweg im Zeitverlauf bei Einschränkung auf die /a/-Vokalgruppe zu beobachten. Für die Köpfe gibt es bei ihrem Bezug auf die /a/-Vokalgruppe nur eine größere Durchbrechung dieses Verhältnisses um ca. die 3800-te Silbe herum.

Abbildung 84: Differenzierung der Köpfe aller /a/-Silben nach (Nicht-)Stimmhaftigkeit im Zeitverlauf

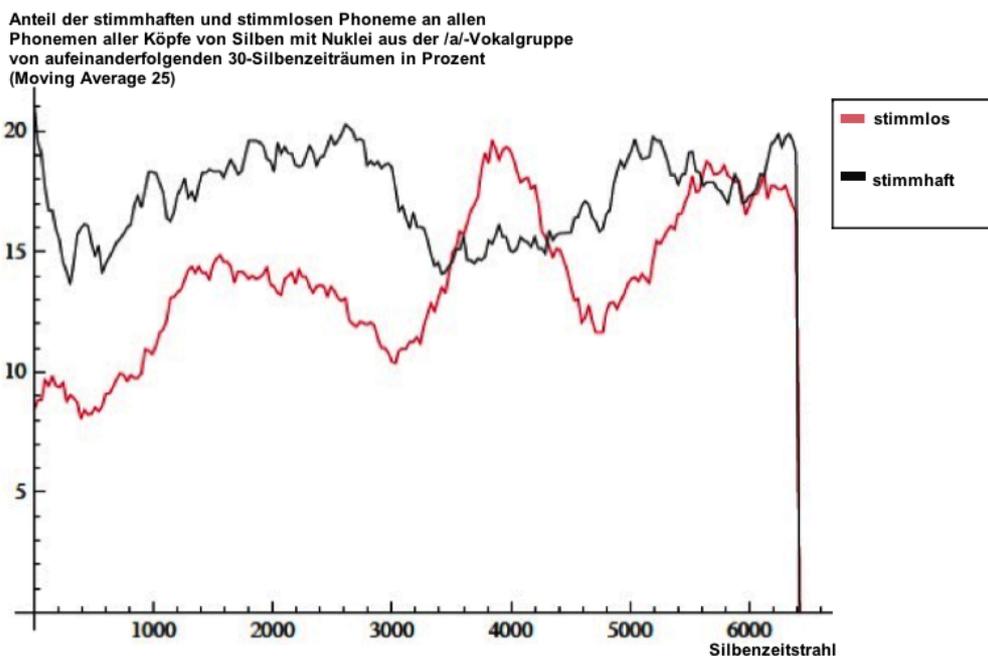
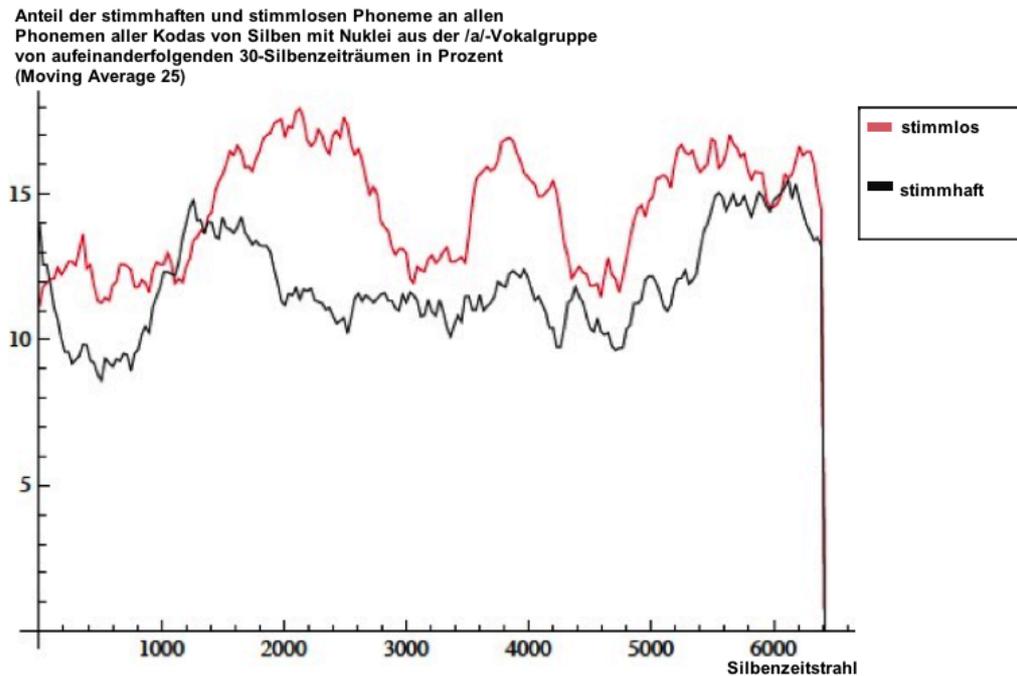


Abbildung 85: Differenzierung der Kodas aller /a/-Silben nach (Nicht-)Stimmhaftigkeit im Zeitverlauf



Sowohl für die Menge der Phoneme aller Köpfe des Annalena-Korpus als auch für alle der Kodas dieses Korpus sind die Alveolaren, Bilabialen und Velaren die drei häufigsten Lautgruppen. Für beide Konstituentenarten sind die Alveolaren die häufigste Lautgruppe, für die Köpfe folgen in der Reihenfolge absteigender Häufigkeit die Bilabialen und Velaren, die beide für die Kodas in einem genau umgekehrten Häufigkeitsverhältnis stehen:

Abbildung 86: Häufigkeitsverteilung der Artikulationsortmerkmale in allen Phonemen aller Köpfe

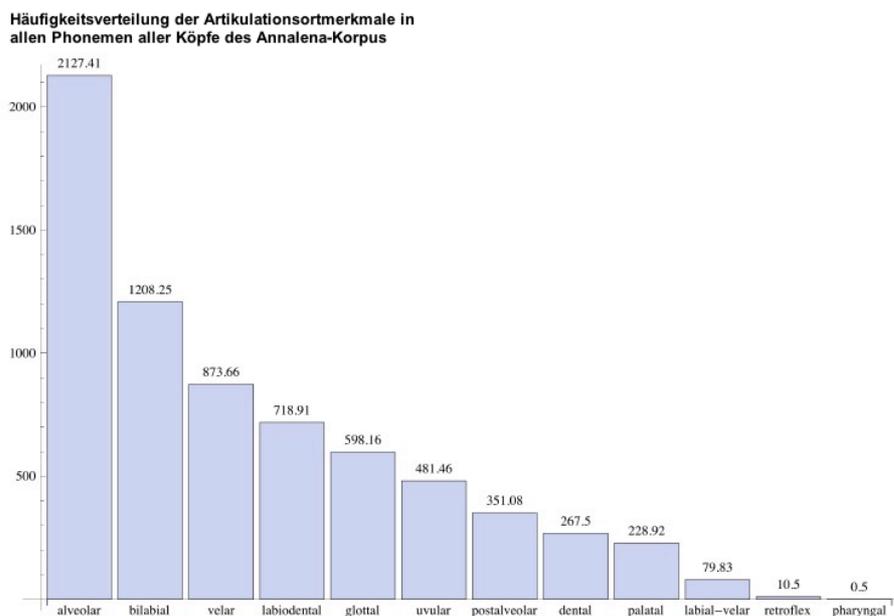
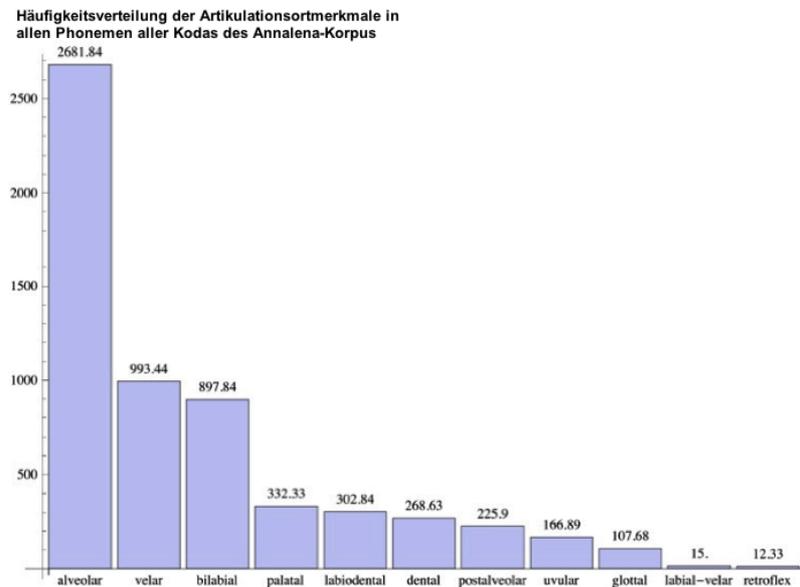


Abbildung 87: Häufigkeitsverteilung der Artikulationsortmerkmale in allen Phonemen aller Kodas



Wegen ihrer mengenmäßigen Dominanz sollen im folgenden nur die Anteile der Alveolaren, Labialen und Velaren im Zeitverlauf betrachtet werden. Die Abb. 88 und Abb. 89 zeigen die Anteilsverläufe dieser drei Ortsmerkmale für die Köpfe und nachfolgend die Kodas. Phasenweise ist eine Gegenläufigkeit für alle drei Konstituentenartenpaare zu beobachten. Bei den Kodas heben sich zudem die alveolaren Phoneme mehr von den Bilabialen und Velaren ab als dies bei den entsprechenden Verläufen für die Silbenköpfe der Fall ist.

Abbildung 88: Differenzierung der Köpfe nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf

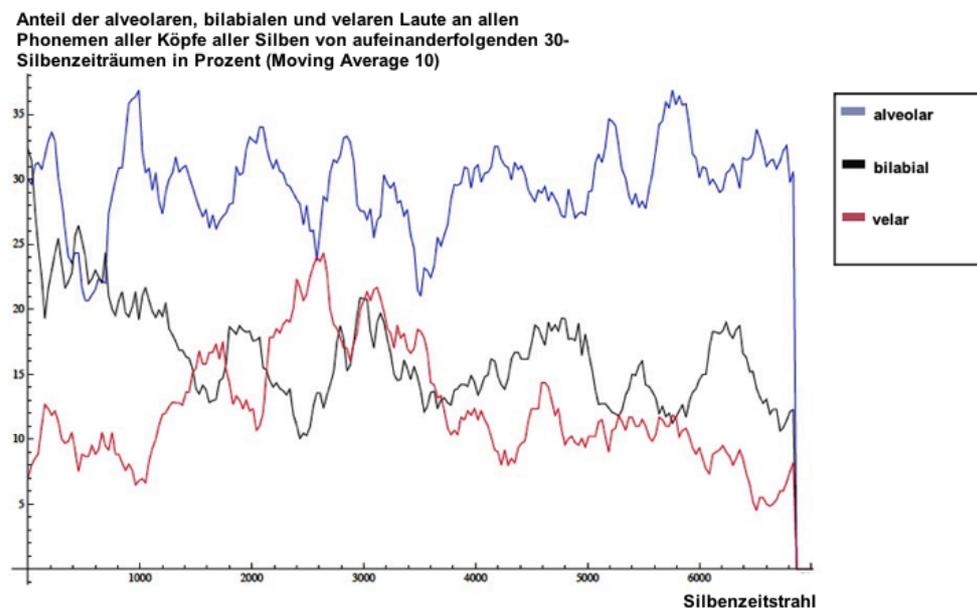
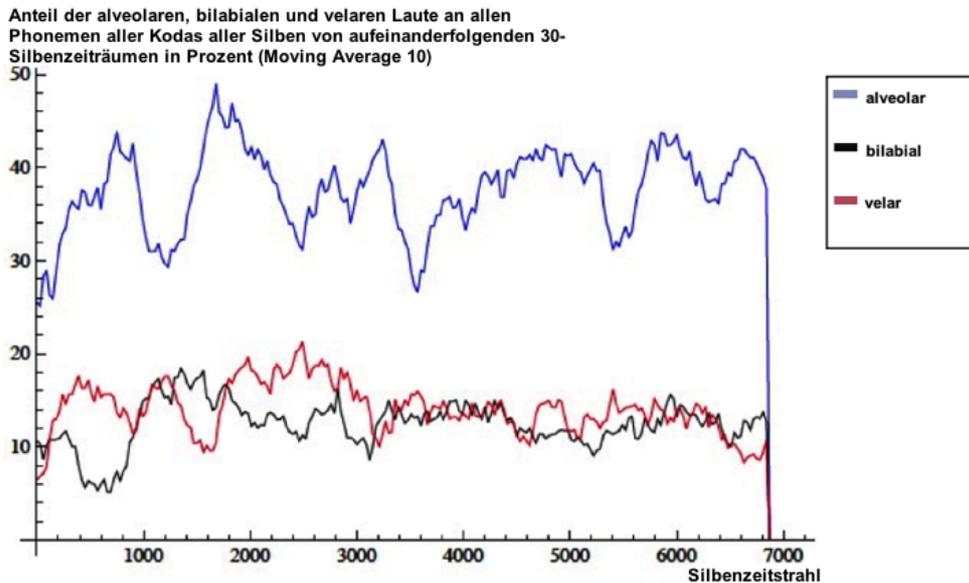


Abbildung 89: Differenzierung der Kodas nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf



Es sind zudem unterschiedliche Anteilsverläufe der Artikulationsorte (alveolar, bilabial und velar) für unterschiedliche Vokalgruppen zu beobachten. So ist für die Köpfe zu Beginn der Entwicklung eine absteigende Häufigkeitsrangfolge alveolar, bilabial, velar für die /a/- und /ɪ/-Kerngruppen zu beobachten, während die Anteilsrangfolge für die Merkmale alveolar und bilabial für die /ʊ/-Vokalgruppe anfänglich genau umgekehrt besteht:

Abbildung 90: Differenzierung der Köpfe aller /a/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf

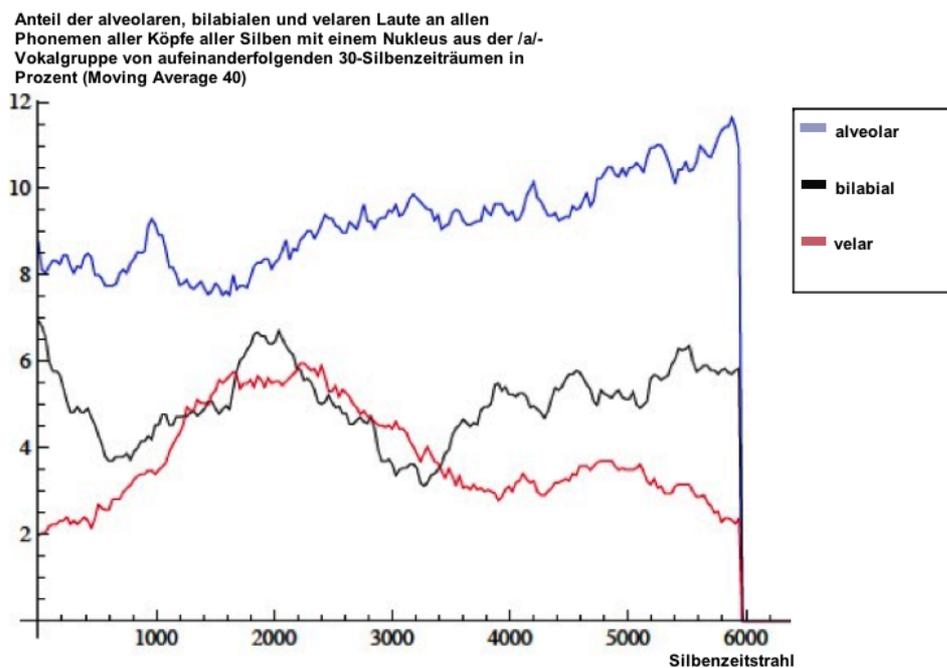


Abbildung 91: Differenzierung der Köpfe aller /i/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf

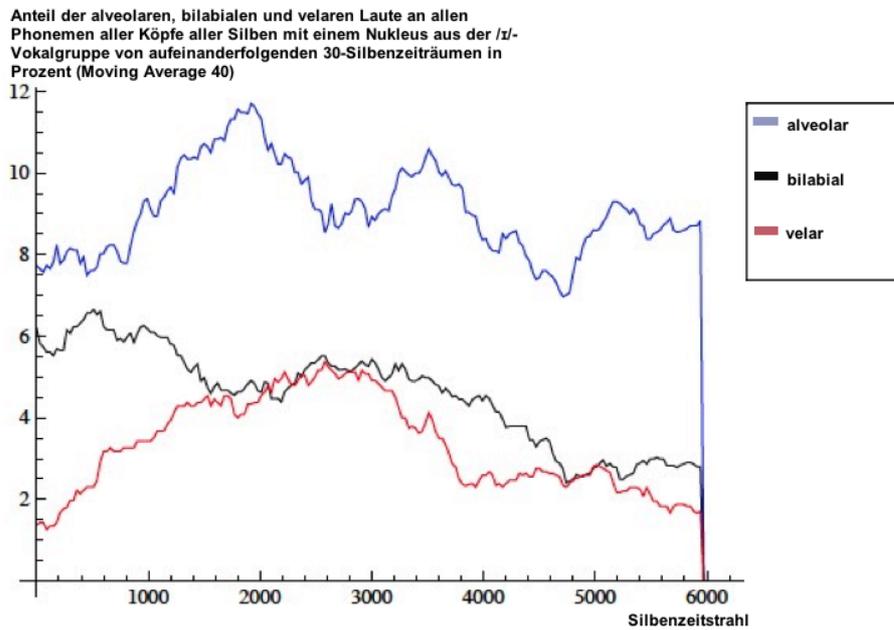
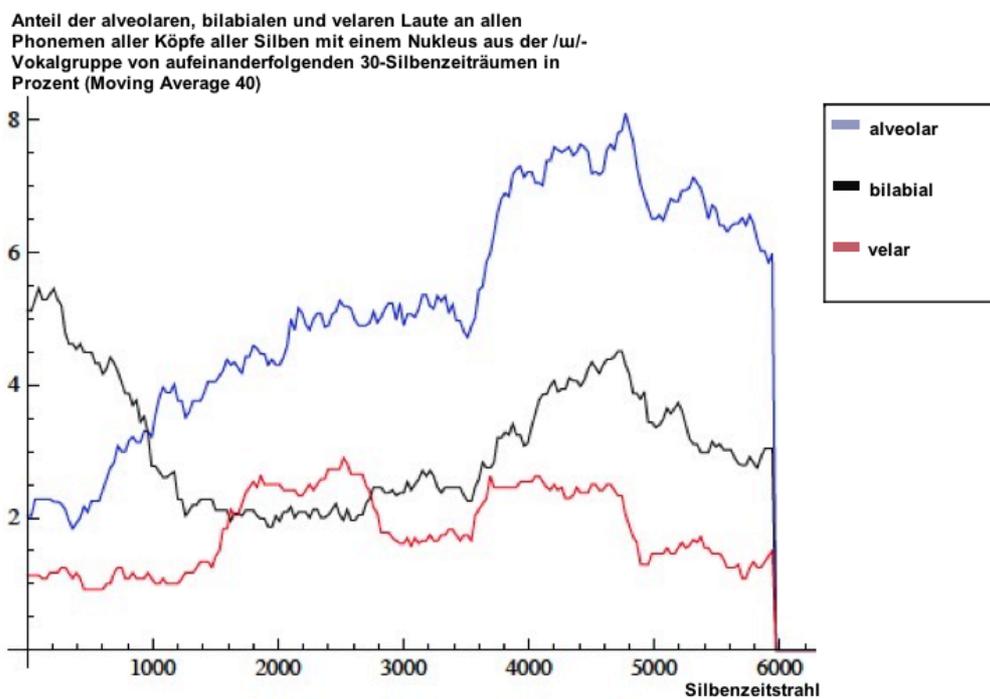


Abbildung 92: Differenzierung der Köpfe aller /u/-Silben nach den Artikulationsortmerkmalen bilabial, alveolar und velar im Zeitverlauf

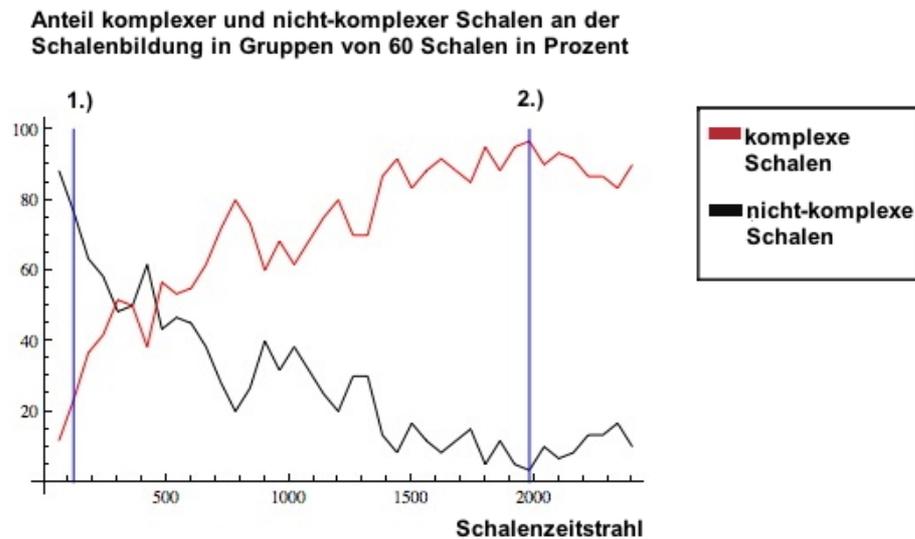


7.3.9.9. Verhältnis von bereits aufgetretenen komplexen Schalen zu nicht-komplexen Schalen oder bereits aufgetretenen C1-Komplexen zu einfachen C1-Köpfen

Die Schalen mit einem komplexen Kopf oder einer komplexen Koda, im Folgenden auch *komplexe Schalen* genannt, verzeichnen einen gegenläufigen Trend im Vergleich zu den nicht-komplexen Schalen. Der Prozentanteil der komplexen Schalen an Zeiträumen von jeweils 60 Schalen nimmt entlang des Schalenzeitstrahls¹⁰³ zu, während der der nicht komplexen abnimmt. In der Abb. 93 sind diese Anteilswerte jeweils bezogen auf eine Einteilung des Schalenzeitstrahls in direkt aufeinanderfolgende, überschneidungsfreie Zeiträume von je 60 Schalen jeweils für jeden dieser Zeiträume zu seinem Beginn abgetragen worden: Die mit "1.)" markierte blaue senkrechte Gerade in der Abb. 93 setzt die sich zeitlich entsprechenden Werte des Zeitpunktes der 120-ten Schale der beiden Anteilsgraphen zueinander in Beziehung: Zu diesem frühen Zeitpunkt hat der Anteilsverlauf der nicht-komplexen Schalen einen hohen Wert von 76,67% jeweils an Zeiträumen von 60 Schalen des Schalenzeitstrahls während der der komplexen Schalen einen niedrigen Wert von 23,33% aufweist. Die mit "2.)" markierte senkrechte blaue Gerade setzt die beiden Anteilsgraphen (komplexe und nicht-komplexe Schalen) wiederum zu einem späteren Zeitpunkt, der 1980-ten Silbe zueinander in Beziehung, zu dem sich das Verhältnis der beiden Anteile umgekehrt hat: Der Graph der nicht-komplexen Schalen weist zu diesem späteren Zeitpunkt einen niedrigen Anteilswert von 3,33% auf, während der der komplexen Schalen einen hohen Wert von 96,67% hat. Diese beiden den Markierungen "1.)" und "2.)" jeweils zugeordneten Anteilswertpaare verdeutlichen also beispielhaft die oben formulierte Beobachtung, dass der Prozentanteil der komplexen Schalen an Zeiträumen von jeweils 60 Schalen entlang des Schalenzeitstrahls zunimmt, während der der nicht komplexen abnimmt. Dies zeigt sich in der Abb. 93 durch einen ansteigenden Graphen der komplexen Schalen (rot) und einen demgegenüber abnehmenden Trend des blauen Graphen der nicht-komplexen Schalen.

¹⁰³ Analog zum *Silbenzeitstrahl* kann man vom *Schalenzeitstrahl* sprechen, weil die Menge aller unterschiedlichen Schalen eine Teilmenge der Silbenmenge des Korpus ist (s.o.). Der Schalenzeitstrahl wird beim Durchlaufen des Silbenzeitstrahls, d.h. der Menge aller unterschiedlichen Silben mit einer für sie angenommenen absoluten zeitlichen Ordnung, durch Extrahieren aller jeweils neuen Schalen im Zeitverlauf bestimmt.

Abbildung 93: Anteil (nicht-)komplexer Schalen an der Schalenbildung im Zeitverlauf



Die in der Abb. 93 dargestellte Realität des Entwicklungsprozesses von kindlichen Lautformen zeigt generell, dass am Anfang die Verwendung von komplexen Rändern viel seltener ist als die von nicht-komplexen. Um diesen Zusammenhang quantitativ näher zu betrachten, soll nun die Relation zwischen Einzelkonsonanten und den Kopfkomplexen untersucht werden, in denen diese Einzelkonsonanten den Anfang¹⁰⁴ bilden. Dazu wird mit *C₁-Silben* die Menge der Silben bezeichnet, die den Konsonanten C₁ als Silbenkopf aufweisen und mit *C₁-Komplexsilben*, die Menge der Silben, die einen komplexen Silbenkopf aufweisen, deren Anfang mit C₁ übereinstimmt. Wie ist das Verhältnis der Anteile der C₁-Silben zu dem der C₁-Komplexsilben im Zeitverlauf jeweils an ihren Korpusgesamtmengen¹⁰⁵? Die Abb. 94 zeigt die durchschnittlichen C₁-Komplexsilben-Anteile im Zeitverlauf für alle C₁ jeweils an ihrer Korpusgesamtmenge, die beim 10-, 20-, ..., 90-, 100%-igen C₁-Silben-Anteilsniveau erreicht werden: Ab Erreichen des 20%-Anteilsniveaus¹⁰⁶ der C₁-Silben überwiegt der Anteil der C₁-Silben den der C₁-Komplexsilben jeweils an der Korpusgesamtmenge. So liegt bei einem C₁-Silbenanteil von 20% ("1.")-Markierung in der Abb. 94) ein durchschnittlicher C₁-Komplexsilbenanteil von 18.707% vor, bei einem von 30% ("2.")-Markierung) einer von

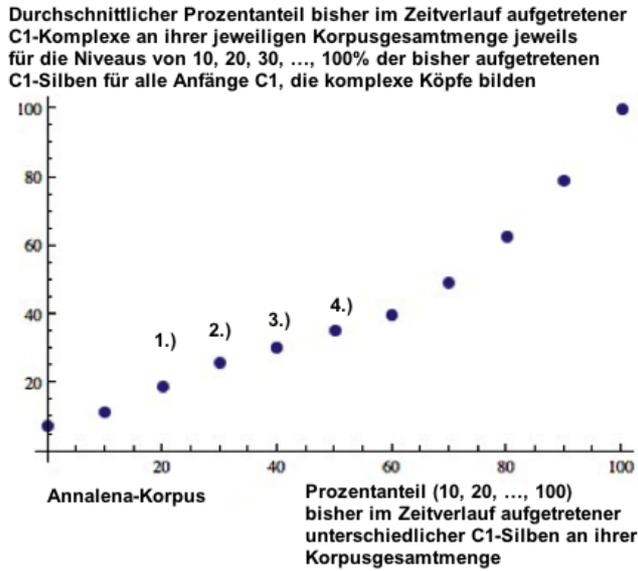
¹⁰⁴ Mit dem Begriff des *Anfangs* einer Silbe wird ihr nicht komplexer Kopf oder das erste Phonem ihres komplexen Kopfes bezeichnet.

¹⁰⁵ d.h. die entsprechende Menge, die im Gesamtkorpus vorliegt

¹⁰⁶ Das anfänglich anteilmäßige Überwiegen der Komplexe ist auf die Frikative /θ, v, x, ʒ, h/, die Nasale /m, n/, den Plosiv /ʔ/ und den Lateral /l/ zurückzuführen, die eine vergleichsweise geringe Gesamtanzahl von Komplexen ausbilden, weshalb ihr Anteil anfänglich so groß ist.

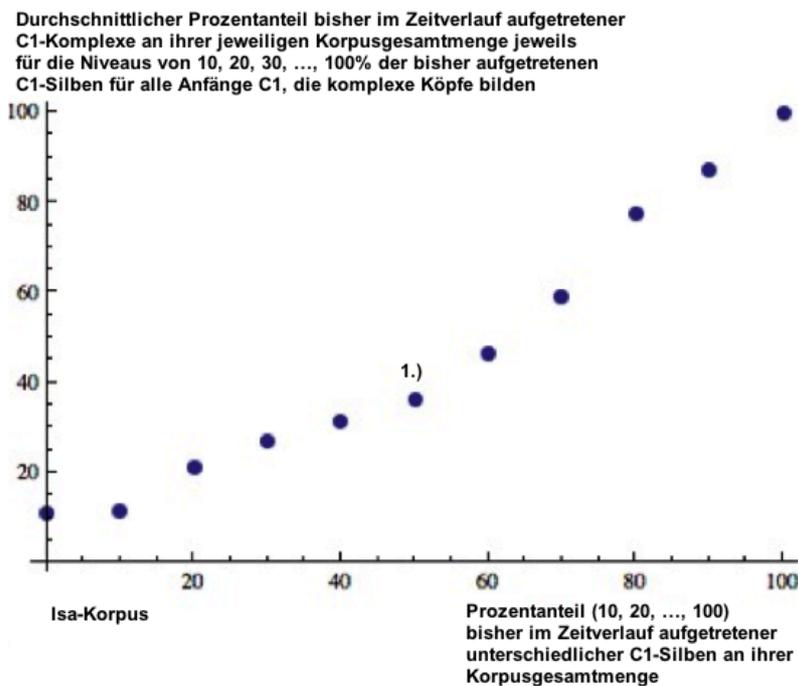
25.759%, bei 40% ("3.")-Markierung) einer von 30.171%, bei 50% ("4.")-Markierung) einer von 35.307%, etc.

Abbildung 94: Vergleich der durchschnittlichen C₁-Komplexsilbenanteile mit den entsprechenden durchschnittlichen C₁-Silbenanteilen im Zeitverlauf (Annalena-Korpus)



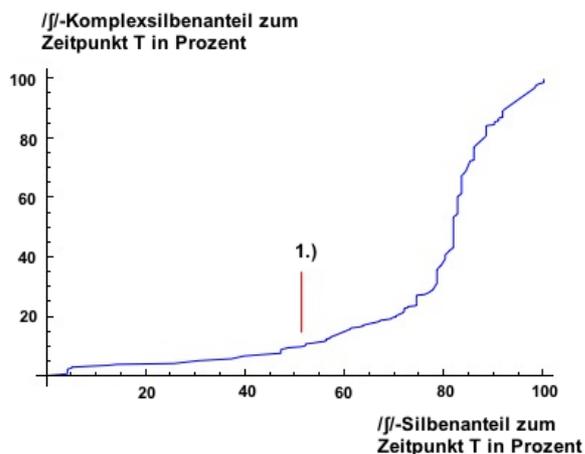
Die analoge Abbildung (Abb. 95) für das ISA-Korpus bestätigt diese Beobachtung:

Abbildung 95: Vergleich der durchschnittlichen C₁-Komplexsilbenanteile mit den entsprechenden durchschnittlichen C₁-Silbenanteilen im Zeitverlauf (Isa-Korpus)



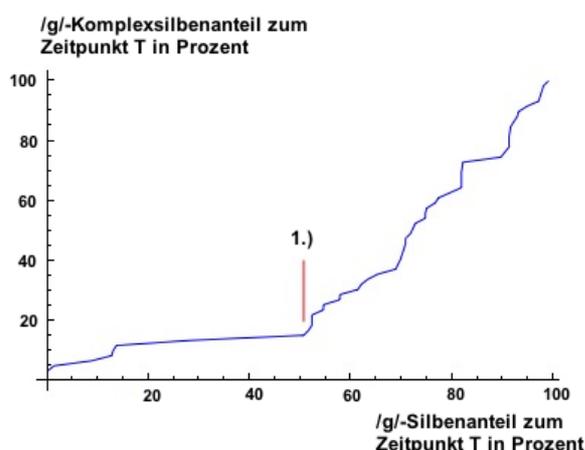
Für das ISA-Korpus ist etwa bei einem 50%-C₁-Silbenanteil erst ein C₁-Komplexsilbenanteil von 36.0249% erreicht. Diese quantitative Relation ist in der Abb. 95 durch "1.)" gekennzeichnet. Viel deutlicher als in der durchschnittlichen Betrachtung ist der starke zeitliche Vorlauf des Anteils der Einzelkonsonantensilben an ihrer jeweiligen Gesamtmenge für die Anfänge /f, g, d, b, ç, f, k, l, p, s, t, z/ (des Annalena-Korpus): So weist der Anfang /f/ beim Erreichen eines /f/-Silbenniveaus von 51% erst ein /f/-Komplexsilbenanteil von 10% auf ("1.)"-Markierung in Abb. 96).

Abbildung 96: Vergleich der /f/-Komplexsilbenanteile mit den entsprechenden /f/-Silbenanteilen im Zeitverlauf



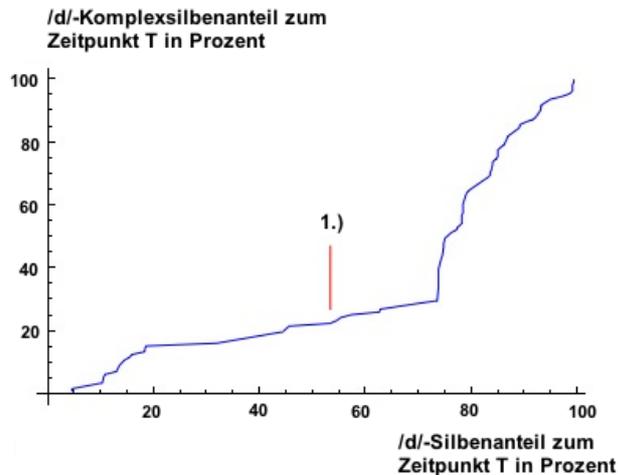
Ein ähnliches Bild zeigt sich für den Anfang /g/ mit einem /g/-Komplexsilbenanteil von 15,25% bei einem /g/-Silbenanteil von 50.64% ("1.)"-Markierung in Abb. 97).

Abbildung 97: Vergleich der /g/-Komplexsilbenanteile mit den entsprechenden /g/-Silbenanteilen im Zeitverlauf



oder /d/ mit einem /d/-Komplexsilbenanteil von 22,522% bei einem /d/-Silbenanteil von 53,285% ("1.)"-Markierung in Abb. 98).

Abbildung 98: Vergleich der /d/-Komplexsilbenanteile mit den entsprechenden /d/-Silbenanteilen im Zeitverlauf



Für die meisten anderen Konsonanten C_1 , die komplexe C_1 -Köpfe¹⁰⁷ bilden, gilt ein ähnlicher zeitlicher Vorlauf der C_1 -Komplexsilben gegenüber dem entsprechenden C_1 -Silbenanteil, für den hier der C_1 -Komplexsilbenanteil bei einem C_1 -Silbenanteilsniveau von ca. 50% angezeigt wird: /b/ (C_1 -Komplexsilbenanteil: 14.173%, C_1 -Silbenanteil: 50.547%), /f/ (Komplex: 16.67%, nicht-komplex: 53.261%), /k/ (Komplex: 19.075%, nicht-komplex: 51.301%), /p/ (Komplex: 18.519%, nicht-komplex: 53.953%), /s/ (Komplex: 18.072%, nicht-komplex: 50.0%), /t/ (Komplex: 16.0 %, nicht-komplex: 50.183%) und /z/ (Komplex: 27.586%, nicht-komplex: 52.83%). Es ist also ein zeitlich teilweise erheblicher Vorlauf der C_1 -Silbenanteile gegenüber den C_1 -Komplexsilbenanteilen zu beobachten. Das vorhergehende Beispiel der Ausbildung von C_1 -Komplexsilben zeigt, dass das Ausmaß des Auftretens von C_1 -Komplexsilben erst zeitlich nach der Ausschöpfung einer ausreichenden Anzahl von C_1 -Silben in größerem Umfang erfolgt.

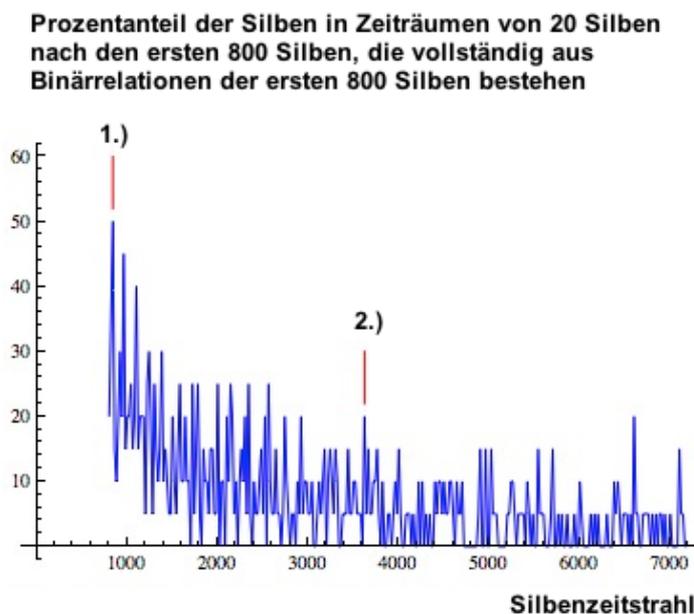
7.3.9.10. Anteil der Silben in gleich großen Zukunftsteilzeiträumen, die noch vollständig aus Binärrelationen des Vergangenheitszeitraums gebildet werden

Ausgehend vom Zeitraum der 0-800-ten Silbe, der in diesem Fall als Vergangenheit betrachtet wird (und auch mit *Vergangenheitszeitraum* benannt wird), wurde die Zukunft ab der 800-ten Silbe (auch *Zukunftszeitraum* genannt) in überschneidungsfreie, direkt aneinander anschließende Teilzeiträume von je 20 Silben eingeteilt. Für jeden 20-Silben Teilzeitraum dieser Einteilung des Zukunftszeitraum der 801-7190-ten Silbe wurde gefragt, welcher Anteil seiner Silben komplett aus Binärrelationen des Vergangenheitszeitraums

¹⁰⁷ C_1 -Kopf bezeichnet einen komplexen Silbenkopf mit C_1 als Anfang.

besteht. Dabei wird deutlich, dass die Bildung neuer Silben aus ausschließlich wieder-
 verwendeten Binärrelationen des Vergangenheitszeitraums in Zukunftsteilzeiträumen
 tendenziell (d.h. als zeitbezogene Durchschnitte über mehrere Anteilswerte betrachtet)
 umso größer ist, je näher diese Zukunftsteilzeiträume dem Vergangenheitszeitraum der 0-
 800-ten Silbe sind: Der mit "1.)" und einer über sich platzierten senkrechten roten Linie
 markierte Wert entspricht dem maximalen Anteilswert von 50% der Abb. 99 und gibt den
 Anteil der Anzahl der Silben des Zukunftsteilzeitraums 841-860 an der Anzahl von 20
 aller Silben dieses Teilzeitraums an, die vollständig aus Binärrelationen des Vergangen-
 heitszeitraums 1-800 bestehen. Der Beginn 841 dieses Zukunftsteilzeitraums 841-860 ist
 nur 41 Silben von dem Ende des Vergangenheitszeitraums entfernt. Der mit "2.)" und
 einer darüber platzierten senkrechten roten Linie markierte Wert von 20% ist hingegen
 erheblich kleiner als der zuvor aufgeführte Maximalwert ("1.)"-Kennzeichnung) und gibt
 den entsprechenden Anteilswert des Zukunftsteilzeitraums 3621-3640 wieder, dessen An-
 fang 3621 aber auch mit einem Abstand von 2821 Silben viel weiter vom Ende des Ver-
 gangenheitszeitraums entfernt ist als der Anfang des zuvor genannten Zukunftsteilzeit-
 raums, dem der Maximalwert von 50% zugeordnet ist.

Abbildung 99: Anteil der vollständig aus Binärrelationen der Vergangenheit (1.-800.-te
 Silbe) gebildeten Silben von Zukunftsteilzeiträumen

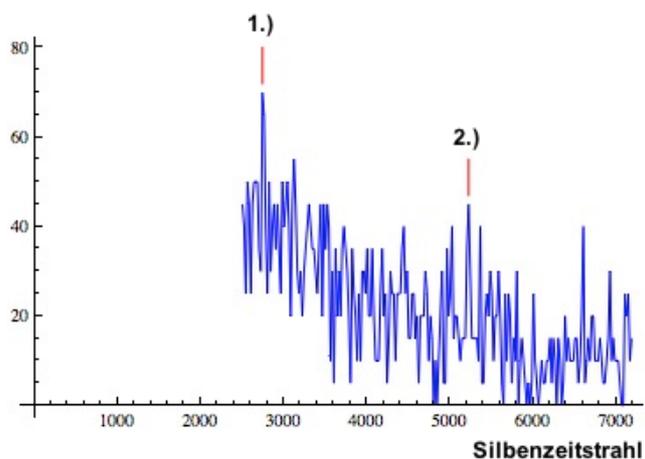


Diese Abfrage wurde für alle Vergangenheitszeiträume ausgeführt, beginnend mit dem
 Zeitraum der ersten 20 Silben (d.h. des Vergangenheitszeitraums der 1-20-ten Silbe), die

sich aus diesem ersten Vergangenheitszeitraum der 1-20-ten Silbe durch sukzessive Vergrößerung um jeweils weitere 20 Silben ergeben. Für alle diese Vergangenheitszeiträume konnte eine ähnliche Beobachtung wie zuvor für den 0-800-Vergangenheitszeitraum gemacht werden in dem Sinne, dass tendenziell (d.h. durchschnittlich über mehrere der Anteilswerte betrachtet) der Anteil der Anzahl der Silben in Zukunftsteilzeiträumen, die vollständig aus Binärrelationen des Vergangenheitszeitraums bestehen, an der Anzahl 20 aller Silben der Teilzeiträume umso größer ist, je näher die Zukunftsteilzeiträume dem Vergangenheitszeitraum sind. Die Abb. 100 ist zu der Abb. 99 analog, nur dass der Vergangenheitszeitraum sich nun von der 1-ten bis einschließlich der 2500-ten Silbe erstreckt:

Abbildung 100: Anteil der vollständig aus Binärrelationen der Vergangenheit (1.-2500-te Silbe) gebildeten Silben von Zukunftsteilzeiträumen

Prozentanteil der Silben in Zeiträumen von 20 Silben nach den ersten 2500 Silben, die vollständig aus Binärrelationen der ersten 2500 Silben bestehen



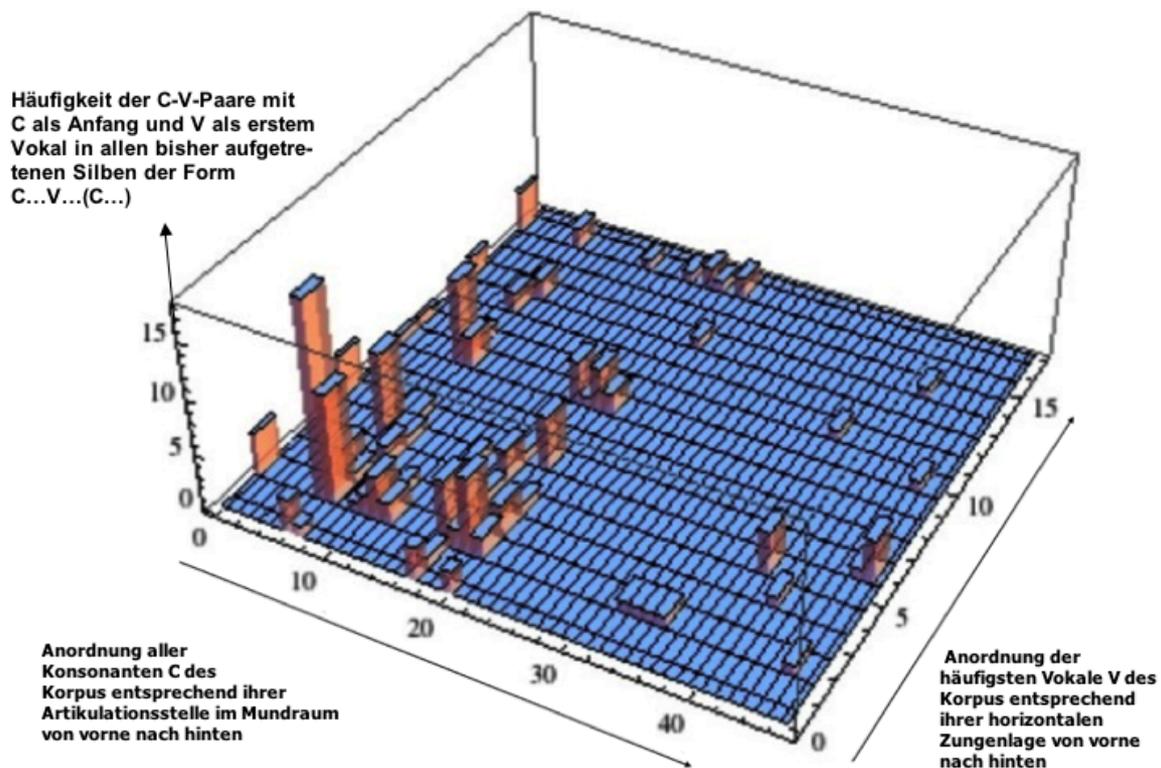
Der mit "1.)" und einer senkrechten roten Linie markierte Wert entspricht dem maximalen Anteilswert von 70% der Abb. 100 und gibt den Anteil der Anzahl der Silben des Zukunftsteilzeitraums 2741-2760 an der Anzahl von 20 aller Silben dieses Teilzeitraums an, die vollständig aus Binärrelationen des Vergangenheitszeitraums 1-800 bestehen. Der Beginn dieses Zukunftsteilzeitraums 2741-2760 ist nur 241 Silben von dem Ende des Vergangenheitsteilzeitraums entfernt. Der mit "2.)" und einer senkrechten roten Linie markierte Wert von 45% ist hingegen kleiner als der zuvor aufgeführte Maximalwert ("1.)"-Kennzeichnung) und gibt den entsprechenden Anteilswert des Zukunftsteilzeitraums 5221-5240 wieder, dessen Anfang 5221 aber auch mit einem Abstand von 2721

Silben viel weiter vom Ende des Vergangenheitszeitraums entfernt ist als der Anfang des zuvor genannten Zukunftsteilzeitraums, dem der Maximalwert von 70% zugeordnet ist.

7.3.9.11. Relation zwischen dem Silbenanfang und dem ersten Phonem eines Silbenkerns hinsichtlich ihrer phonetischen Lage (Artikulationsort, horizontale Zungenlage)

Es ist eine anfänglich starke Konzentration in vorderen Bereichen des Mundraums zu beobachten. Die Abb. 101 zeigt für einen Anfangszeitraum der Entwicklung die Häufigkeit des Auftretens konkreter C_1 - V_1 -Paare aller bereits aufgetretenen unterschiedlichen Silben, die in der Form $C_1(C_2\dots)V_1(V_2\dots)(C'_1\dots)$ geschrieben werden können.

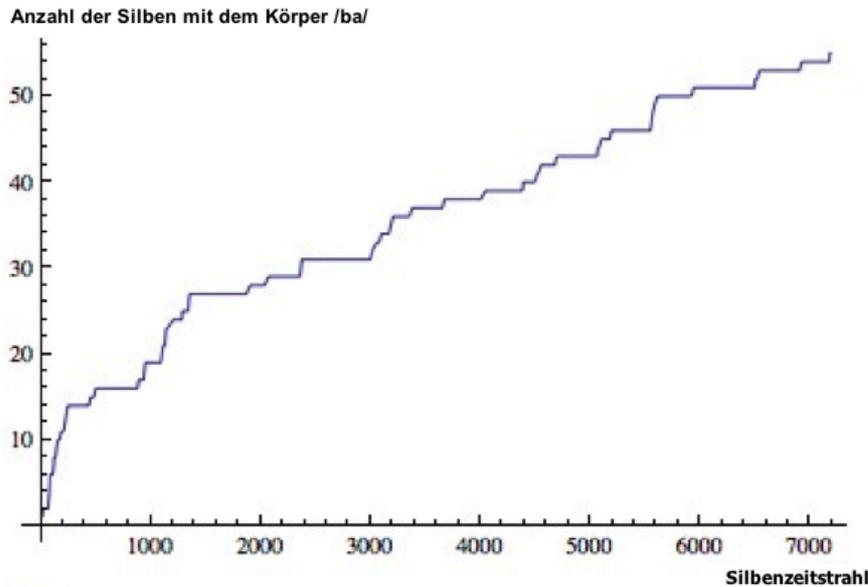
Abbildung 101: Häufigkeitsverteilung von Artikulationsort und horizontaler Zungenlage aller C-V-Paare der Anfangsilbenmenge



Es ist eine starke Konzentration der C_1 - V_1 -Paare in vorderen Bereichen des Mundraums zu sehen. Da diese Häufung besonders hinsichtlich vorderer Konsonanten ausgeprägt ist, zeigen sich für zwei Gruppen vorderer Konsonanten angeordnete "Bahnen", die von den

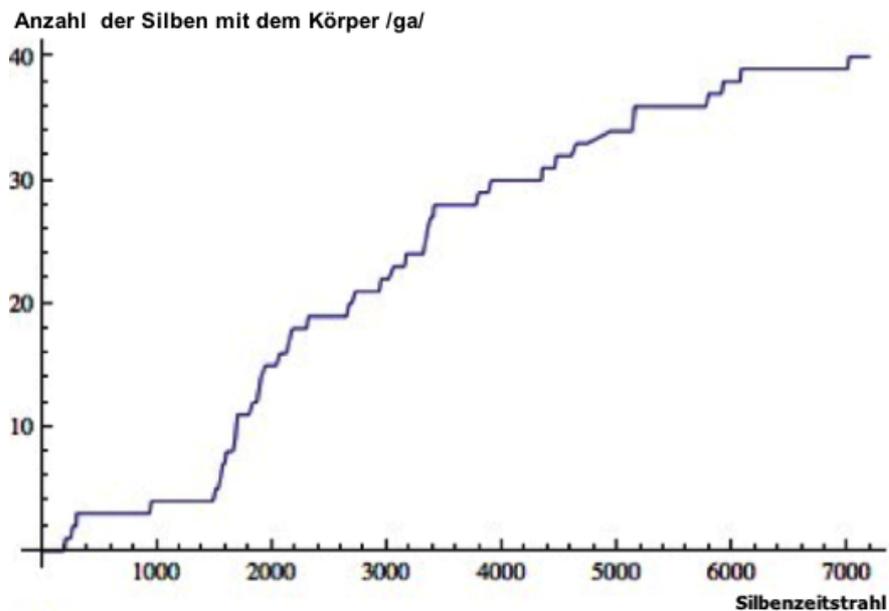
vorderen über die mittleren zu den hinteren Vokalen verlaufen. Beispielsweise gehört der Silbenkörper /ba/ zu diesem vorderen, anfänglich schon stark ausgeprägten Konsonanten-Vokal-Bereich:

Abbildung 102: Kumulative Häufigkeit des Körpers /ba/



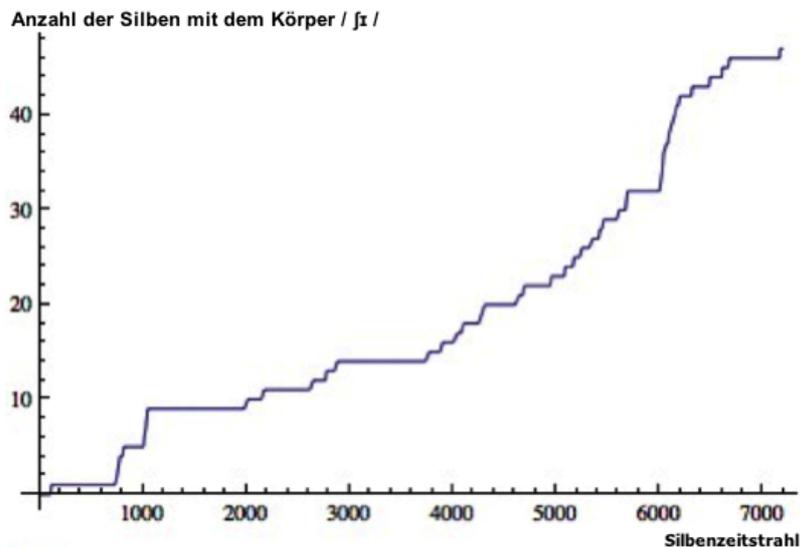
Der Kopf /ga/ ist hingegen anfänglich bisher kaum aufgetreten:

Abbildung 103: Kumulative Häufigkeit des Körpers /ga/



Aber nicht alle vorderen Konsonant-Silbenkern-Bereiche müssen auch häufig sein, wie das Beispiel des Kopfes /fi/ zeigt:

Abbildung 104: Kumulative Häufigkeit des Körpers /fi/

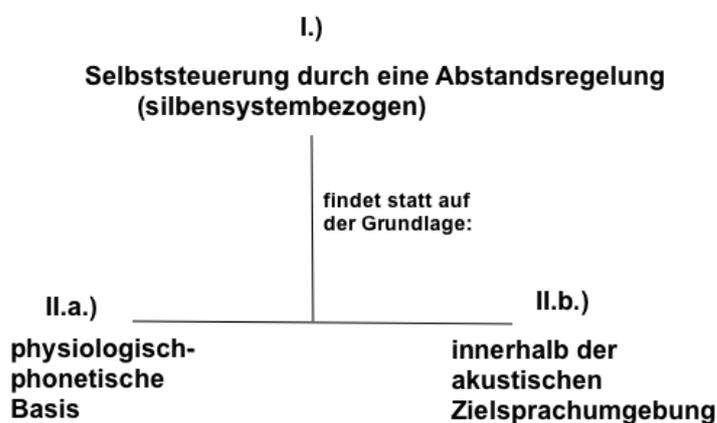


7.3.10. Diskussion der Beobachtungen und Ergebnisse der Verteilungsuntersuchungen in den Korpora

In diesem Abschnitt soll eine abschließende Zusammenfassung, Bewertung und Interpretation der in den Abschnitten 7.3.5. - 7.3.9. beschriebenen Beobachtungen zur Verteilung der Silbenkonstituenten in den Korpora in Form von 14 Tendenzen vorgenommen werden.

1.) Die Selbststeuerung des Silbenwachstums durch eine Abstandsregelung:

Abbildung 105: Die Selbststeuerung des Silbenwachstums durch eine Abstandsregelung



Auch wenn letztlich alle Beobachtungen der Korpusuntersuchungen als das Ergebnis des durch eine Abstandsregelung selbst gesteuerten Silbenwachstums verstanden werden, so sollen unter diesem Prinzip einige Beobachtungen aufgeführt werden, die sich als eine

während der Abstandsregelung wirkende ökonomische oder differenzierende Tendenz oder als ein Verhältnis der beiden begreifen lassen, die die Selbststeuerung des Silbenwachstums begrenzen und die Abstandsregelung steuern. Der Begriff "Abstand" ist nicht primär im phonetischen Sinne räumlich zu verstehen, auch wenn eine räumliche Vorstellung auf Grund der physiologischen Basis der sprachlichen Lauterzeugung beteiligt ist. Vielmehr bezieht sich dieser Begriff auf die Annahme einer in jeder Phase der Lautentwicklung geltende innere Beziehung zwischen allen ihren Elementen, die auf der einen Seite eine zu starke Ähnlichkeit zwischen einzelnen Elementen verhindert, jedoch andererseits die maximale Ausnutzung von Lautbildungsmöglichkeiten begünstigt. Da das gesamte Lautsystem über variierende Schnittmengenbildungen ständig in Veränderung begriffen ist, verschieben sich auch die "Abstände" in diesem Prozess in der Form einer Bewegung zwischen zwei Gegensätzen: Werden kritische Schnittmengengrenzen überschritten oder zu große Systemdichten erreicht, verlagert sich die Schnittmengenbildung. Somit vermeidet der Silbenbildungsprozess, dass Silben ihre Unterscheidbarkeit verlieren. Dagegen verlangen die ökonomische Ausnutzung begrenzter physiologischer Ressourcen und Begrenzungen innerhalb von Silbensystemkonstellationen die maximale Wiederverwendung von Silbenbestandteilen durch variierende Kombinationen. Da eine in dieser Weise begriffene Abstandsregelung ein komplexer Prozess ist, können die im Folgenden erläuterten Teilprinzipien nur als Annäherung an die angestrebte Formulierung der Abstandsregelung begriffen werden:

1.a.) Tendenz der maximalen Ökonomie und minimalen Differenzierung:

Als ein Hinweis auf eine Tendenz der maximalen Ökonomie und minimalen Differenzierung wird die bis auf einen Anfangszeitraum mehrheitlich beobachtete Wiederverwendung (Abschnitt 7.3.8.4.) verstanden.

1.b.) Zunehmende Streuung:

Auf eine zunehmende Streuung der Silbenmenge deuten die im letzten zeitlichen Drittel mehrheitlich nur in einer einzigen Silbe vorkommenden Schalen hin (Abschnitt 7.3.9.6., Abb. 66) sowie der ansteigende Trend des Kopf-Graphen in Abb. 29, weil sich die Silbenmenge jeweils auf eine größere Zahl von Bestandteilen (Schalen, Köpfe) verteilt.

1.c.) Zunehmende Wiederverwendung:

Der zum Ende des Silbenzeitstrahls in Abschnitt 7.3.9.5. (Abb. 65) gemessene ansteigende Wert bedeutet die zunehmende Wiederverwendung von Silbenkonstituenten im Zeitverlauf.

1.d.) Optimaler Ausgleich/Wechsel zwischen Ökonomie + Differenzierung:

Die in Abschnitt 7.3.9.5. (Abb. 65) beobachteten Wellenbewegungen zwischen lokalen Maxima und lokalen Minima könnten als Anzeichen eines optimalen Ausgleichs zwischen der Ökonomie und Differenzierung angesehen werden, indem Anstiege eine Verstärkung der ökonomischen und eine Verringerung der Differenzierungstendenz bedeuten, während umgekehrt ein Abstieg als eine Verstärkung der Differenzierungstendenz und eine Verringerung der Ökonomisierungstendenz aufgefasst werden könnte.

1.e.) Kerndifferenzierung mittels Schalenmengendifferenzierung:

Die in Abschnitt 7.3.9.4 beobachtete Verteilung der Silbenkerne auf die Menge aller Schalen soll folgendermaßen interpretiert werden: Ohne die Unterschiede zwischen einzelnen Lautformen zu berücksichtigen, soll für die in Abschnitt 7.3.9.4. gemachten Beobachtungen angenommen werden, dass große Übereinstimmungen zwischen den Schalenmengen aller Silbenkerne in einem temporären Silbensystem als Anzeichen für geringe Abstände im System angesehen werden können und tendenziell als Verringerung der Eindeutigkeit und Verständlichkeit, während umgekehrt ein konstanter oder in einer gewissen Bandbreite schwankender Prozentsatz fehlender Übereinstimmung die Eindeutigkeit und Verständlichkeit optimiert. Die erzielten Ergebnisse könnten also als Wirken einer Abstandsregelung angesehen werden, da die starke Differenzierung der Schalenmengen vermutlich zur Unterscheidbarkeit innerhalb des Silbensystems beiträgt. Die (fast) vollständigen Übereinstimmung für die Silbenkerne der Gruppe B der Selteneren wird nicht als Widerspruch zur angenommenen Abstandsregelung gesehen, da durch die geringen Silbengesamthäufigkeiten in der Gruppe B schon dadurch eine Differenzierung erreicht wird.

2.) Unterschiedliches Verhalten verschiedener Arten von Silbenbestandteilen:

Eine Vielzahl von Beobachtungen deutet auf die Unterschiedlichkeit von Silbenbestandteilen hin, besonders auf die der drei verschiedenen Silbenbestandteilklassen (bzw. Sil-

benkonstituentenarten) der Köpfe, Kerne und Kodos. Diese wird durch die in dem Modell dieser Arbeit angenommenen unterschiedlichen Funktionen der Silbenkonstituentenarten verständlich: Die Silbenkerne sind Hauptsignalträger der im Silbenbildungsprozess produzierten Silben, während die Ränder die Eigenschaften der Silbenkerne sind. Eine weitere Differenzierung der Ränder versteht die Köpfe als Haupteigenschaften der Silbenkerne und die Kodos als deren Nebeneigenschaften. Es wurden u.a. die folgenden Unterschiede für die Silbenkonstituentenarten beobachtet: Die Anzahl jeweils aller unterschiedlichen Köpfe, Kodos und Kerne in gleichgroßen Silbenmengen verläuft auf unterschiedlichem Niveau: Es sind phasenweise paarweise Gegenläufigkeiten festzustellen, der Graph für die Kerne befindet sich, abgesehen vom Anfang auf niedrigstem Niveau und der der Köpfe zeigt einen ansteigenden Trend (Abschnitt 7.3.8.3.). Auch bei einer phonetisch basierten und vokalbezogenen Differenzierung zeigen sich Unterschiede für die Anteile der einzelnen Silbenkonstituentenarten entlang des Silbenzeitstrahls (Abschnitt 7.3.9.8.). Die folgenden 4 Differenzen beziehen sich auf den besonderen Unterschied der Nuklei gegenüber den Rändern: Erstens hat die Kernklasse mit 206 (Annalena) bzw. 73 (Isa) die geringste Anzahl unterschiedlicher Elemente gegenüber den beiden Ränderklassen (Abschnitte 7.2.1. und 7.2.2.). Zweitens zeigt die Anzahl unterschiedlicher Kerne anfänglich den größten Zuwachs gegenüber den Rändern, bevor sich in der Folge dieses Verhältnis umkehrt (Abschnitt 7.3.9.3.: Abb. 48 und Abb. 54). Drittens zeigt die Klasse der Kerne hinsichtlich der Anzahl ihrer bereits aufgetretenen unterschiedlichen Elemente bezogen auf den Silbenzeitstrahl ein nicht-lineares Wachstum, während das der Schalen (d.h. der Kombination von je einem Element der beiden Ränderklassen (Köpfe, Kodos)) annähernd als linear beschrieben werden kann. Viertens kann eine unterschiedliche Silbengesamthäufigkeit des jeweils häufigsten Elements der drei Silbenkonstituentenarten (Köpfe, Kerne, Kodos) festgestellt werden (Abschnitt 7.3.5.3.): Beim Vergleich der Verteilungen der Kerne mit denen der anderen Silbenbestandteilsarten (d.h. der Köpfe, Kodos oder Schalen) wird aber deutlich, dass sich die Kerne von diesen unterscheiden, da ihre häufigsten "Elemente" eine weitaus größere Silbengesamthäufigkeit als jeweils die häufigsten "Elemente" der anderen Silbenbestandteilsarten haben. So hat der häufigste Kern eine Silbengesamthäufigkeit von mehr als 800 unterschiedlichen Silben, der häufigste Kopf oder die häufigste Koda hingegen nur eine Silbengesamthäufigkeit, die etwas größer als 420 ist. Der zuletzt genannte Unterschied bezieht sich im Gegensatz zu den bisher genannten nicht auf die Konstituentenarten der Silbe, sondern ihre Binärrelationsarten (Abschnitt 7.3.5.4.): Es wurde beobachtet, dass bei beiden Kindern die Reime (d.h. die Kern-

Koda-Kombinationen) am wenigsten zahlreich ausgeprägt sind. Hingegen sind bei Annalena die Schalen etwas häufiger als die Körper, während bei ISA die Körper erheblich zahlreicher sind. Die eingangs erwähnte unterschiedlichen Funktion für die Koda gegenüber den Köpfen kann möglicherweise als Grund für die beobachtete stärkere "Kombination" von Kernen mit Köpfen als mit Koda und damit dem Vorliegen von mehr Körpern als Reimen angesehen werden.

3.) Innovationsdominanz des häufigsten Silbenkerns hinsichtlich der Schalenmenge:

Der am häufigsten in unterschiedlichen Silben vorkommende Silbenkern greift anteilmäßig am wenigsten auf Schalen zurück, die bereits in Silben mit anderen Silbenkernen vorgekommen sind (vgl. den Begriff des *Schalenverhältnisses* in Abschnitt 7.3.6.2.). Der häufigste Silbenkern ist also „innovativer“ hinsichtlich der Bildung neuer Schalen als seltenere Nuklei.

4.) Integration statt klarer Trennung von Prozessbestandteilen:

Unter der Integration gegenüber der Trennung von Bestandteilen des Silbenbildungsprozesses werden Aktivitäten von Bestandteilen des Silbenbildungsprozesses verstanden, die sich nicht voneinander trennen lassen, sondern durchmischt auftreten.

So deutet das in Abschnitt 7.3.8.2 global (d.h. auf den Gesamtzeitraum des Annalena-Korpus bezogene) beobachtete phasenweise gleichmäßige Wachstum einzelner Silbenkerne darauf hin, dass ihre Aktivitäten miteinander kombiniert werden, weil in dieser auf den Gesamtzeitraum bezogenen Sicht i.A. keine prägnanten Ruhepausen der einzelnen Kerne beobachtet werden können. Die ineinandergreifende Aktivität der Silbenkerne könnte evtl. so erklärt werden, dass die Aktivitäten von Silbenkernen als Voraussetzung und Folge die Aktivitäten eines oder mehrerer anderer Silbenkerne y_1, y_2, \dots, y_n ($n \geq 2$) erfordern. Es wird vermutet, dass die Reihenfolge des zeitlichen Auftretens einer bestimmten Schale s bei verschiedenen Silbenkernen bedeutungsvoll ist und deshalb ihre Wiederverwendung bei y ihr zeitlich vorhergehendes erstmaliges Vorkommen mit x in einer Silbe erfordert. Im Grunde genommen ist dies die Verallgemeinerung der Beobachtung von Davis und MacNeilage (1995 in MacNeilage 1998: 504f.), dass es durch quantitative Häufung bestimmte Präferenzen für die Kombination von Köpfen mit Nuklei gibt, die aber in MacNeilages (1998: 503f., 2008: 88ff.) Interpretation mit den bereits in Abschnitt 2.2. erwähnten grundlegenden Lebensäußerungen (z.B. die Kieferoszillationen beim Kauen oder Lippenbewegungen beim Saugen (ebd.)) übereinstimmen, während für

die hier vorgenommenen Überlegungen der Zustand des Silbensystems zu einem vorliegenden Zeitpunkt eine entscheidende Rolle spielt. Auf diese voneinander abhängige Aktivität einzelner Silbenkerne deutet auch die in Abschnitt 7.3.9.7. durchgeführte Phaseneinteilung hin, die zeigt, dass der Silbenbildungsprozess als Aktivität zweier phasenweise durch das dort festgelegte Phasenkriterium miteinander in Beziehung stehender Silbenkerne aufgefasst werden kann. Die in Abschnitt 7.3.9.1. durchgeführte Untersuchung des Kombinationsverhaltens konkreter Köpfe mit allen mit ihnen in einer Silbe auftretenden Kodas zeigt, dass es auch eine zeitliche Durchmischung des Silbenbildungsgeschehens gibt, indem ein konkreter Kopf aus dem Blickwinkel seines zeitlichen Geschehens spät auf Kodas zur "Kombination" in Silben zurückgreifen kann, die evtl. bei einem anderen konkreten Kopf sehr früh aufgetreten sind. Die Beobachtung der paarweisen Gegenläufigkeit der Anzahl jeweils unterschiedlicher Köpfe, Kodas und Kerne in gleich großen, direkt aufeinanderfolgenden und überschneidungsfreien Silbenmengen im Zeitverlauf kann möglicherweise als Anzeichen für Interaktionen zwischen je zwei dieser Silbenbestandteilklassen gewertet werden. Damit verliefen die Aktivitäten der einzelnen Klassen nicht unabhängig voneinander, sondern bedingten sich gegenseitig. Diese Annahme muss allerdings noch durch weitere Untersuchungen geprüft werden.

5.) Tendenz der ungleichen Verteilung:

Die Tendenz der ungleichen Verteilung bezieht sich auf die Aufteilung der gesamten Silben-, Schalen- oder Lautformenmenge (kurz: *Gesamtmenge* als Bezeichnung für eine dieser Mengen) jeweils auf eine Menge einer bestimmten Art von Bestandteilen ihrer Elemente¹⁰⁸ (kurz: *Bestandteilmenge*) und besagt, dass der überwiegende Teil aller Elemente der Gesamtmenge hauptsächlich von einer kleinen Anzahl von Elementen der Bestandteilmenge gebildet wird, d.h. wenige Elemente einer Bestandteilart dominieren quantitativ-anteilmäßig den Prozess der Bildung von Elementen der Gesamtmenge, und alle übrigen sind sehr selten bzw. sie kommen in sehr wenigen unterschiedlichen Elementen der Gesamtmenge vor. Diese Tendenz der ungleichen Verteilung wird zum einen in Häufigkeitsverteilungen von Gesamtmengenelementen für den Gesamtzeitraum sichtbar (s.o.), aber auch in dem Anteil, den bestimmte Gruppen der jeweils häufigsten Silbenkerne jeweils im Zeitverlauf an den neu gebildeten Silben haben (Abschnitt 7.3.8.5.). Für auf das Gesamtkorpus bezogene Untersuchungen lassen sich diese Verteilungen auch so aus-

¹⁰⁸ (d.h. Silben- (Kopf, Koda, Kerne)) (Abschnitt 7.3.5.2. und 7.3.5.3.), Schalen- (Kopf, Koda) (Abschnitt 7.3.5.3.) oder Lautformenbestandteilen (Silben) (Abschnitt 7.3.5.5.)

drücken, dass es tendenziell umso mehr Bestandteile einer bestimmten Art im gesamten betrachteten Korpus gibt, je kleiner die für jeden einzelnen Bestandteil dieser Art geforderte Gesamthäufigkeit in der betrachteten Gesamtmenge ist. Die Vergleichbarkeit der Art dieser ungleichen Verteilungen für die Silben in der Lautformenmenge wie für die Silbenbestandteile in der Silbenmenge wird zum Anlass für die folgende Vermutung genommen: Es wird vermutet, dass die (mehrsilbigen) Lautformen ebenfalls einem organischen Wachstum durch die "Kombination" der Silben folgen, wie dies für die "Kombination" der Konstituenten in den Silben beobachtet wurde. Zudem wird es neben einem zum organischen Silbenwachstum (bezogen auf die Konstituenten) analogen Lautformenwachstum (bezogen auf die Silben) auch für das Lautformenwachstum spezielle Prinzipien geben. Die Daten der hier untersuchten kindlichen Korpora sind allerdings nicht ausreichend umfangreich, um eine genügende Wiederholung von Silben in unterschiedlichen Lautformen zu gewährleisten.

6.) Tendenz des Silbenkonstituentenbezugs:

Die Verteilungen der Silbengesamthäufigkeiten einzelner konkreter Phoneme auf die einzelnen Silbenpositionen hängen i.A. von der konstituentenbezogenen Silbenposition ab, die sich u.a. in unterschiedlichen Silbengesamthäufigkeiten für dasselbe Phonem für unterschiedliche Positionen oder unterschiedliche Positionierungen des Silbengesamthäufigkeitsmaximums für unterschiedliche Phoneme zeigen kann (Abschnitt 7.3.5.1.).

7.) Tendenzuelle Zeit-Quantitätsrelation:

Es wurde für verschiedene Arten von Silbenbestandteilen beobachtet, dass die "Elemente" einer solchen Art tendenziell umso früher zum ersten Mal in einer festgelegten Menge von Silben auftreten (*Zeitpunkt des ersten Auftretens*) müssen, je größer ihre Gesamthäufigkeit in dieser Silbenmenge ist (Abschnitt 7.3.7.1.). So gilt beispielsweise für die Menge aller Schalen in der Gesamtmenge aller Silben als deren Bestandteilen, dass eine Schale tendenziell umso früher im Verlauf des Gesamtprozesses der Silbenbildung zum ersten Mal auftreten muss, in je mehr unterschiedlichen Silben sie in der Gesamtsilbenmenge des Korpus vorkommt (Abb. 17,18).

8.) Tendenz der Kontinuität (Gleichgewicht) und der Diskontinuität (Ungleichgewicht):

In Abschnitt (7.3.8.2.) wurden auffällige Wechsel der Wachstumsgeschwindigkeit in Graphen der kumulativen Silbenhäufigkeit einiger der häufigsten Silbenkerne beobachtet (Abb. 23, 24), die den Übergang bilden zwischen zwei Phasen unterschiedlicher durchschnittlicher Wachstumsgeschwindigkeit. Für die Interpretation dieser Beobachtung scheint die Verwendung des Begriffspaares des (Un-)Gleichgewichts passend, weil mit einem Ungleichgewicht oft stärker zu beobachtende Veränderungen einhergehen als mit einem Gleichgewichtszustand. Erstens kann der Wechsel einen Übergang von einer Phase stärkeren Wachstums zu einer Phase geringeren Wachstums wie für den Nukleus /a/ bedeuten: Zu einem Ungleichgewicht könnte es z.B. deshalb kommen, weil ein regelmäßiges "Kombinations"-geschehen von Silbenkonstituenten des Silbenkerns /a/ mit anderen "Kombinations"-partnern nicht mehr in dem bisherigen Ausmaß möglich ist. Weil der Bedarf an neuen Signalen unvermindert anhält, kommt es zu internen Systemverschiebungen. Beide Begriffe des (Un-)Gleichgewichts und der Systemverschiebungen müssen noch durch eine weitere Analyse genauer gefasst werden. Zweitens kann der Wechsel einem Übergang zu einer Phase stärkeren Wachstums entsprechen: So beginnt der im Deutschen vorkommende Vokal /ʊ/ zu einem Zeitpunkt lange nach dem Phonemerwerbsbeginn überhaupt erst merklich zu wachsen. Oft sind solche erst spät in Erscheinung tretenden Laute wie /ʊ/ mit der ungefähr gleichzeitigen Abnahme oder dem völligen Verschwinden anderer nicht im Deutschen vertretener Phoneme wie dem Vokal /u/ verbunden (vgl. auch das Paar aus nativem /ð/ und nicht-nativem /f/ (Abschnitt 7.3.8.2.)). Bei einer solchen zeitliche Korrelation eines nativen und nicht nativen Lautes wird häufig davon gesprochen, dass der native Laut für den nicht-nativen in entsprechenden Lautformen erst zu diesem späten Zeitpunkt "ersetzt" wird, weil er physiologisch schwieriger als der nicht native sei (vgl. Stampe (1979) oder Schultze (1880)). So stellt beispielsweise Schultze bereits 1880 sein von ihm der Kindersprache "zu Grunde [...] [gelegtes] Gesetz" auf, das besagt, "dass die Sprachlaute im Kindermunde in einer Reihe hervorgebracht werden, die von den mit der geringsten physiologischen Anstrengung zu Stande kommenden Lauten allmählich übergeht zu den mit größerer, und endet bei der mit größter physiologischer Anstrengung zu Stande gebrachten Sprachlauten." (Schultze 1880: 27). Dieses verbindet er mit seinem "Lautverschiebungs- oder Verstümmelungs- oder Verwandlungsgesetz der Kindersprache [...]: Für den dem Kinde noch unaussprechbaren

Laut (Vokal oder Konsonanten) setzt dasselbe den diesem schwierigen Laute nächstverwandten, mit geringerer physiologischer Schwierigkeit sprechbaren Laut, und wenn es auch diesen noch nicht zu beherrschen vermag, so lässt es ihn einfach ganz und gar weg." (Schultze 1880: 37). Diese Sichtweise der Kindersprache fällt in die Klasse der in dieser Arbeit als Diskurstheorien bezeichneten Spracherwerbtheorien (vgl. Abschnitt 2.1.) und wird als unzureichend angesehen. Vielmehr bedarf es der Bestimmung der inneren Systemverhältnisse wie sie in dieser Arbeit u.a. in Form des Abstandsmaßes angestrebt wurden. Damit ist nicht eine Vernachlässigung der physiologischen Seite der Sprache verbunden, sondern eine Betonung des Silbensystems, welches aber auf der Grundlage u.a. der Physiologie sich entwickelt (vgl. Abschnitt 2.2.).

9.) Ausnutzung von Silbenbildungspotentialen:

Es wird angenommen, dass es von dem aktuellen Zustand eines Silbensystems ausgehend jeweils determinierte Möglichkeiten der Silbenbildung gibt, die durch die Bildung von Silben ausgeschöpft werden müssen, bevor neue Silbenbildungsmöglichkeiten bzw. -potentiale erschlossen werden. Es wird also die Annahme vertreten, dass das Silbensystem zu jedem Zeitpunkt an bestimmte artikulatorische, perzeptive sowie sich aus dem Zustand des Systems selbst ergebende Voraussetzungen gebunden ist, die genutzt und ausgeschöpft werden, bevor neue Bereiche durch eine innovative Anpassung erschlossen werden. Diese Voraussetzungen oder Möglichkeiten sollen mit dem Begriff des *Silbenbildungspotentials* bzw. der *Silbenbildungsressource* bezeichnet werden. So werden beispielsweise die nicht komplexen Ränder als Ressourcen im zuvor definierten Sinne betrachtet, die zunächst ausgeschöpft werden, bevor andere Silbenbildungsressourcen wie die komplexen Ränder erschlossen werden (Abschnitt 7.3.9.9.). Das Konstrukt der Silbenbildungspotentiale bzw. -ressourcen bedingt u.a. die Gerichtetheit des Prozesses, weil es von system- und damit zeitbezogenen Möglichkeiten der Silbenbildung ausgeht. Der Begriff der Silbenbildungspotentiale wird durch die Modellannahme begründet, dass die Silbenbildung zu jedem Zeitpunkt des Silbenbildungsprozesses in Abhängigkeit vom Zustand des Silbensystems erfolgt.

10.) Gerichtetheit des Silbenbildungsprozesses:

Unter der *Gerichtetheit* des Silbenbildungsprozesses eines Kindes wird eine Nicht-Austauschbarkeit von Silben(bestandteils)mengen unterschiedlicher Zeiträume miteinander verstanden. Anstatt von Nicht-Austauschbarkeit könnte man auch allgemei-

ner davon sprechen, dass je zwei überschneidungsfreie gleichgroße Silbenteilzeiträume kaum oder keine Schnittmengen hinsichtlich bestimmter Bestandteile miteinander gemeinsam haben. Auf die Gerichtetheit des Silbenbildungsprozesses deuten u.a. die Beobachtungen in den Abschnitten 7.3.9.2., 7.3.9.9., 7.3.9.10. hin, weil sie Reihenfolgen von klar umrissenen Silben(bestandteils)mengen, die einseitige Determinierung von Paaren von Teilzeiträumen oder geringe zeitliche Schnittmengen von Teilzeiträumen nahelegen.

11.) Individualität der Silbenentwicklung für unterschiedliche Kinder:

Es wird davon ausgegangen, dass die Entwicklung des Silbensystems für jedes einzelne Kind individuell, d.h. unterschiedlich für verschiedene Kinder verläuft und sich diese Individualität in unterschiedlich ausgeprägten Merkmalen der Verteilungen der Silbenkonstituenten in unterschiedlichen Silbenkorpora zeigt wie beispielsweise unterschiedlichen Ränge gleicher Nuklei für die beiden Kinder Isa und Annalena (Abschnitt 7.3.5.2. (Abb. 5, 6)). Die Individualität der Entwicklung ist dabei durch die angenommene Art der Systemsteuerung begründet.

12.) Starke Konzentration im Raum:

Es wird vermutet, dass auf Grund des begrenzten phonetischen und phonologischen Raumes Konzentrationen bei der Lautbildung stattfinden, die sich in der Form von Wiederholungen oder Wechseln zwischen wenigen Konstituenten abspielen. Die in Abschnitt 7.3.9.7. im Zusammenhang mit der silbenkernpaarbasierten Phaseneinteilung gemachte Beobachtungen der häufigen Bildung von Phasen für Paare aus Kernen einer der vier /a/-, /i/-, /u/- oder /ε/- Silbenkerngruppen deutet möglicherweise auch auf eine zeitliche Konzentration auf auditive Vokalkontraste hin und damit ein Beispiel der zuvor beschriebenen Konzentration im (phonetischen) Raum.

13.) Der begrenzte Raume und die Existenz einer Silbenhöchstgrenze:

Für beide untersuchten Korpora (Annalena, Isa) konnte in Abschnitt 7.3.8.1. eine Zunahme der Silbenproduktivität (d.h. Anzahl der pro Sitzung gebildeten neuen Silben) vom Sprachbeginn an beobachtet werden, aber nur bei dem Kind Annalena das Erreichen eines Maximums der Silbenproduktivität und einer daran anschließenden Abnahme dieser Größe. Vermutlich ist der Verlauf der Silbenproduktivität für das Kind Annalena Ausdruck des begrenzten Raums, der eine nur in Grenzen variable Silbenhöchstgrenze (vgl.

Abschnitt 2.2.) erzwingt, die wegen der Notwendigkeit der Eindeutigkeit der Silben nicht überschritten werden darf. Vor Erreichen der Silbenhöchstgrenze gibt es Ausschöpfungserscheinungen, die sich evtl. in der Art der Verteilung der "Kombinationspartner" von Konstituenten zeigt, und zu einer langsameren Kombination und damit einer Verringerung der Silbenproduktivität führen: Zunächst sind viele Bereiche des phonetischen Raums und des phonologischen Pendants verfügbar, was die Steigerung der Silbenproduktivität ermöglicht. Nimmt jedoch die Anzahl der bereits gebildeten unterschiedlichen Silben zu, so nehmen die Abstände zwischen den Silben im phonetischen Raum ab und es verringert sich die Anzahl der Möglichkeiten, neue, noch verbliebene Räume des Raumes zu besetzen oder durch neue "Kombinationen" von in Silben bereits aufgetretenen Konstituenten neue Silben zu bilden. Ein begrenzter Raum kann nicht unbegrenzt mit Elementen besetzt werden, ohne dass ihre Unterscheidbarkeit gefährdet wäre. So erzwingt die Begrenztheit des Raumes, auf den bei der Silbenbildung zurückgegriffen wird, und seine zunehmende Nutzung und die damit verbundene abnehmende Zahl an Silbenbildungsmöglichkeiten eine der Silbenproduktivitätssteigerung folgende Verringerung der Zuwachsgeschwindigkeit neuer Silben und in der weiteren Folge vermutlich eine innere Umstrukturierung des Silbensystems und damit verbunden des Phonemsystems, die sich phänomenal u.a. in der Form von „Verschmelzungen“ von Silben und Wörtern vollzieht.

14.) Silbenproduktivitätssteigerung und das Schalen- und Konstituentenwachstum:

In Abschnitt 7.3.9.3 wurde eine auffällige Zunahme der durchschnittlich pro Sitzung produzierten neuen Silben beobachtet für die zwei mögliche Gründe diskutiert werden sollen: Erstens soll analog zur Definition der physikalischen Arbeit angenommen werden, dass die produzierten Silben ein Produkt aus Kraft und Weg sind. Unter der weiteren Annahme, dass der physiologische Aufwand des Sprechens innerhalb einer Bandbreite konstant ist, ließe sich für die ersten Silben eine relativ größere in ihnen verkörperte Energie annehmen als in den Silben der folgenden Phasen. Zweitens wird eine Verringerung beim Wachstum neuer Kerne als Auslöser für ein mit der Silbenproduktivitätssteigerung einhergehendes vermehrtes Schalenwachstum vermutet. Zunächst konzentriert sich der Prozess auf die Kerne als Hauptsignalträger, dann haben sie einen Umfang erreicht, der ihre Unterscheidbarkeit im Raum gefährdet und eine vermehrte Kerndifferenzierung über die Ränder erfordert. Deshalb ist vermutlich ein verstärktes Schalenwachstum und insbesondere eine stärkere "Kombination" der bei einzelnen, konkreten Kernen bereits aufgetrete-

nen Köpfe und Koda miteinander zu Schalen in Silben jeweils mit diesen Kernen zu beobachten.

8. Ausblick

Voraussetzung für eine weitergehende Untersuchung des Phonemerwerbs sind detaillierte Korpora mit einer absoluten Ordnung der Lautformen. Aufbauend auf detaillierten Korpora mit einer verlässlichen Zeitkorrelation können die in dieser Arbeit entwickelten Konzepte erweitert werden¹⁰⁹. Das in dieser Arbeit entwickelte Schnittmengenkonzept kann auf alle Bereiche der Linguistik angewendet werden, und es sollte auch bei der Entwicklung einer Alternative zur IPA Verwendung finden. Auf der Lautebene der Sprache muss eine quantitative Modellrechnung für die Silben- und Phonementwicklung aufgestellt werden, die deren naturwissenschaftliche Analyse ermöglicht. Vor dem Hintergrund des angenommenen engen Zusammenhangs zwischen kausalem Motor und damit dem Lexikonumfang und dem Silbensystem sollten historische Lautveränderungen zu den technologischen und sozialen Veränderungen der jeweiligen Zeit in Beziehung gesetzt werden und ihre Auswirkungen auf das Lexikon (und damit das Silben- und Phonemsystem) überprüft werden. Wegen des engen Zusammenhangs zwischen der Ausbildung eines Denksystems (kognitive Funktion der Sprache, kausaler Motor) und der ersten Sprache (Muttersprache) sollte der Einfluss des frühen Lernens einer Fremdsprache auf die Kognition überprüft werden. Außerdem sollte das Verhältnis von Sprachnormen und individuellen Sprachsystemen neu überdacht werden. Mit einem besseren Verständnis der Verteilungslogik von individuellen Silbensystemen sind auch neue Möglichkeiten der Komprimierung von Sprachdaten denkbar.

Obwohl auch andere Ansätze zum L1-Phonemerwerb teilweise eine evolutionäre Herangehensweise anstreben, sind sie entweder den Diskurstheorien oder Theorien zuzordnen (vgl. Kapitel 2, 4, 5, 6), die eine phonetische Steuerung der Sprachentwicklung und -veränderung annehmen. Während in den Diskurstheorien ein Angleich an die normative Zielsprache als treibende Kraft oder in den Theorien der phonetischen Steuerung ein Optimierungsstreben für den L1-Phonemerwerb und den Spracherwerb angenommen wird, sieht der in dieser Arbeit vertretene Ansatz die kognitive Funktion (kausaler Motor) als treibende Kraft der Sprachentwicklung des Kindes (und der Menschheit) an. Dieser evolutionäre Ansatz verzichtet auf alle "black boxes" von angeborenen oder aus dem Diskurs

¹⁰⁹ siehe dazu auch die Anmerkungen im Abschnitt 7.1.

oder der Physiologie der Sprachorgane sich ergebenden Regelsysteme und legt den Schwerpunkt auf das individuelle eigenständige Sprachsystem des individuellen Sprachteilnehmers und insbesondere des Kindes, während bei den Vergleichstheorien (siehe Kapitel 2, 4, 5, 6) eine Norm- oder Durchschnittsprache im Vordergrund steht.

9. Die Digitalisierung, Auswertung und Notation der Korpora

Für die Notation, Digitalisierung und Auswertung der Korpora und die Visualisierungen der auf der Grundlage der Korpora berechneten Daten wurden folgende Hilfsmittel verwendet: Für die Eingabe des in Lindner (2009: 455ff.) als "Kardinaltabelle" (ebd.) angegebenen ISA-Korpus in ein Textdokument wurde die XSampa-Notation (Wells 2012) genutzt. Eine Tabelle der in von Wells (2012) vorgeschlagenen XSampa-IPA-Entsprechungen ist auch unter

<http://www.diku.dk/hjemmesider/studerende/thorinn/xsamchart.gif>
abrufbar.

In dieser Arbeit werden für die Bezeichnung der Laute die IPA-Symbole des Internationalen Phonetischen Alphabets (International Phonetic Association 2007) in der für 1996 aktualisierten Version (ebd.) verwendet wie sie zu Beginn des Handbuchs der *International Phonetic Association* (ebd.) u.a. in Tabellenform aufgeführt sind und erklärt werden (ebd.: 3ff.). Bei der Digitalisierung wurde für die Eingabe des Annalena-Korpus aus Elsen (1991: 136ff.) die Eingabemaske des über die Adresse

<http://childes.psy.cmu.edu/phon/>

verfügbaren Programms "Phon" des *PhonBank*-Projekts des CHILDES-Projekts (MacWhinney 2000) genutzt. Nach der vollständigen Eingabe wurden die Daten aus dem "Phon"-Programm als csv-Dateien exportiert und in ein Objekt der Programmiersprache Python (z.B. Beazley 2006) eingelesen. Das ISA-Korpus (Lindner 2009 455ff.) wurde mittels XSampa-Zeichen in eine Textdatei eingegeben und im Anschluss in ein Python-Objekt eingelesen. Die Konvertierung der XSampa- in IPA-Unicode-Zeichen erfolgte mit Hilfe der Python-Funktion "xsampa.py" (Kleiweg 2007). Im Anschluss an die Digitalisierung und Speicherung der Korpora in Python-Objekten erfolgte ihre Auswertung durch in der Programmiersprache Python geschriebene Funktionen. Die in dieser Arbeit dargestellten Graphiken wurden mit Hilfe des Programms Mathematica^{®110} erstellt, wobei die

¹¹⁰ Mathematica[®] ist ein eingetragenes Warenzeichen der Firma Wolfram Research, Inc.

für die Graphiken notwendigen Daten im Vorhinein mit den in Python programmierten Funktionen aus den Korpora abgefragt bzw. berechnet wurden.

10. Literatur

Auer, P. 1994. "Einige Argumente gegen die Silbe als universale prosodische Hauptkategorie". In: K.-H. Ramers & H. Vater & H. Wode (Hrsg.). *Universale phonologische Strukturen und Prozesse*. Tübingen: Niemeyer, 55–78.

Baese-Berk, M. & M. Goldrick. 2009. "Mechanisms of interaction in speech production." *Language and Cognitive Processes* 24: 527-554.

Barsalou, L. W. 1999. "Perceptual symbols systems". *Behavioral and Brain Sciences* 22: 577-660.

Bates, Elizabeth A. & Elman & Jeffrey L. & Johnson, Mark & Karmiloff-Smith, Annette & Parisi, Domenico & Plunket, Kim 1997. *Rethinking innateness. A connectionist perspective on development*. Cambridge/MA/London: MIT Press.

Bavin, Edith L. (Hrsg.). 2009. *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press.

Beazley, D. 2006. *Python. Essential reference*. Indiana, USA: Sams Publishing.

Bell, Alan & Joan Bybee Hooper. 1978. "Issues and evidence in syllabic phonology". In: Bell, Alan & Joan Bybee Hooper (Hrsg.). *Syllables and segments*. Amsterdam: North-Holland Publishing Company, 3-22.

Benkí, J. R. 2003. "Analysis of English Nonsense Syllable Recognition in Noise". *Phonetica* 60: 129-157.

Bertoncini, J. & J. Mehler 1981. "Syllables as units in infant speech perception". *Infant behavior and development* 4: 247-260.

Best, C. T. 1995. "Learning to perceive the sound pattern of English". In: C. Rovee-Collier & L. P. Lipsitt (Hrsg.). *Advances in infancy research*. Norwood, New Jersey: Ablex, 217-304.

Blake, J. 2000. *Routes to Child Language. Evolutionary and Developmental Precursors*. New York: Cambridge University Press.

Blevins, J. 1995. "The syllable in phonological theory". In: J. A. Goldsmith (Hrsg.). *The handbook of phonological theory*. Cambridge: Blackwell, 206-244.

Bloom, Paul 1996. *Language acquisition. Core readings*. Cambridge/MA: MIT Press.

- Borghi, Anna M., Fabian Chersi, Serge Thill und Tom Ziemke 2010. "Sentence processing: Linking language to motor chains". *Frontiers in Neurorobotics*. Bd. 4. Artikel 4. 1-9.
- Bosch, A. R. K. 2011. "Syllable-internal structure". In: M. van Oostendorp & K. Rice & B. Hume & C. Ewen (Hrsg.). *The Blackwell Companion to Phonology*. London: Wiley-Blackwell, 781-798.
- Browman, C. P. & L. Goldstein 1986. "Towards an articulatory phonology". *Phonology Yearbook* 3: 219-252.
- Browman, C. & L. Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée* 5: 25-34.
- Bugge, Franz 1997. *Die Entwicklungspsychologie Jean Piagets*. Stuttgart, Berlin, Köln: Verlag W. Kohlhammer.
- Bybee, Joan L. 2004. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan L. 2006. "Language change and universals". In: Mairal, Ricardo & Juana Gil (Hrsg.). *Linguistic Universals*. Cambridge: Cambridge University Press, 179-194.
- Bybee, Joan L. 2008. "Formal universals as emergent phenomena: The origins of Structure Preservation". In: Good, Jeff (Hrsg.). *Linguistic universals and language change*. New York: Oxford University Press, 108-121.
- Charles-Luce, J., & Luce, P. A. 1990. "Similarity neighborhoods of words in young children's lexicon". *Journal of Child Language* 17: 205-215.
- Cholin, J. & W. J. M. Levelt & N. O. Schiller 2006. "Effects of syllable frequency in speech production". *Cognition* 99: 205-235.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Christiansen, M. & S. Kirby. 2003. "Language Evolution: The hardest problem in science?" In: Christiansen, M. & S. Kirby (Hrsg.). *Language Evolution*. New York: Oxford University Press.
- Croft, W. 2001. *Radical construction grammar. Syntactic theory in typological perspective*. New York: Oxford University Press.

- Clark, Eve V. 2009. *First language acquisition*. Cambridge: Cambridge University Press.
- Clements, G.N. 1990. "The role of the sonority cycle in core syllabification". In J. Kingston & M. Beckman (Hrsg.). *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech*. New York: Cambridge University Press, 283-333.
- Content, A. & R. K. Kearns und U. H. Frauenfelder. 2001. "Boundaries versus onsets in syllabic segmentation". *Journal of Memory and Language* 45: 177-199.
- Curtin, Suzanne & Dan Hufnagle. 2009. "Speech perception". In: Bavin (Hrsg.), 107-123.
- Dabrowska, Ewa. 2004. *Language, mind and brain. Some psychological and neurological constraints on theories of grammar*. Edingburgh: Edingburgh University Press.
- Davis, Barbara L. & Krisztina Zajdó (Hrsg.). 2008. *The syllable in speech production*. New York/London: Lawrence Erlbaum Associates.
- Davis, S. 2006. "Syllabic constituents". In: K. Brown (Hrsg.). *Encyclopedia of language and linguistics*. Band 12. Oxford: Elsevier, 326-328.
- de Boer, B. 2001. *The origins of vowel systems*. New York: Oxford University Press.
- Delattre, P. C. & A. M. Liberman & F. S. Cooper. 1955. "Acoustic loci and transitional cues for consonants". *Journal of the Acoustical Society of America* 27: 769-773.
- Dunabeitia, J. D. & J. Cholin & J. Corral & M. Perea & M. Carreiras. 2010. "SYLLABARIUM: An online application for deriving complete statistics for Basque and Spanish orthographic syllables." *Behavior Research Methods* 42, 118-125.
- Dresher, Elan B. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- Elsen, Hilke. 1991. *Erstspracherwerb. Der Erwerb des deutschen Lautsystems*. Wiesbaden: Deutscher Universitäts Verlag.
- Elsen, Hilke 1994. "Phonological constraints and overextensions". *First Language* 14: 305-315.
- Elsen, Hilke. 1995. "Der Aufbau von Wortfeldern". *Lexicology* 1: 219-242.
- Elsen, Hilke 1999. *Ansätze zu einer funktionalistisch-kognitiven Grammatik. Konsequenzen aus Regularitäten des Erstspracherwerbs*. Tübingen: Max Niemeyer Verlag.

Ferguson, C. A. & Farwell, C. B. 1975. "Words and sounds in early language acquisition". *Language* 51: 419-439.

Fitch, Tecumseh W. 2010. *The evolution of language*. Cambridge.

Fowler, C. A. & E. Saltzman 1993. "Coordination and coarticulation in speech production". *Language and Speech* 36: 171-195.

Gafos, A. & S. Benus 2006. "Dynamics of phonological cognition". *Cognitive Science* 30: 905–943.

Geeraerts, Dirk. 2010. *Theories of lexical semantics*. New York: Oxford University Press.

Gerdes, Adele. 2008. *Spracherwerb und neuronale Netze. Die konnektionistische Wende*. Marburg: Tectum Verlag.

Goldsmith, John A. 1979. *Autosegmental Phonology*. New York/London: Garland Publishing.

Goldstein, L. & H. Nam & E. Saltzman & I. Chitoran 2008. "Coupled oscillator planning model of speech timing and syllable structure". In: *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers*, Beijing, April 18-20, 2008 <http://sail.usc.edu/~lgoldste/ArtPhon/Papers/Week%2012/LG_final.pdf> (Zugriff: 3. Oktober 2012)

Goldstein 2012. "Dynamical stability in speech production and sound change" <http://www.traces.uab.cat/soundchange/abstracts_talks/Goldstein_abstract.pdf> (Zugriff: 02.10.2012)

Gopnik, A., & Schulz, L. E. (2007). *Causal learning: Psychology, philosophy, and computation*. New York: Oxford University Press.

Harris, Zellig 1968. *Mathematical structure of language*. New York, London, Sydney, Toronto: John Wiley and Sons.

Inkelas, Sharon & Yvan Rose. 2008. "Positional Neutralization: A Case Study from Child Language". *Language* 83: 707-736.

International Phonetic Association. 2007. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala. [Wiederabdruck in: R. Jakobson. 1962: *Selected writings I: Phonological studies*. Den Haag: Mouton 328- 401].

Jakobson, R. & L. R. Waugh 1986. *Die Lautgestalt der Sprache*. Berlin/New York: Walter de Gruyter.

Johnson, Wyn & Paula Reimers. 2010. *Patterns in child phonology*. Edinburgh: Edinburgh University Press.

Johnson, Keith. 1997. "Speech perception without speaker normalization". In: Keith Johnson & John W. Mullenix (Hrsg.). *Talker Variability in Speech Processing*. San Diego: Academic Press, 145–166.

Johnson, Keith & John W. Mullenix. 1997. "Complex representations used in speech processing". In: Keith Johnson & John W. Mullenix (Hrsg.). *Talker Variability in Speech Processing*. San Diego: Academic Press, 1-8.

Jusczyk, Peter W. 2000. *The discovery of spoken language*. Cambridge/MA/London: MIT Press.

Kempelen, Wolfgang von. 1791. *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*. Wien: J. V. Degen. [Facsimile Neudruck 1970 der Ausgabe Wien 1791 mit einer Einleitung von Herbert E. Brekle und Wolfgang Wildgen. Stuttgart: Friedrich Frommann.]

Kleiweg, P. 2007. *Translate X-SAMPA to Unicode*.
<<http://www.let.rug.nl/kleiweg/L04/devel/python/xsampa.py>> (Zugriff: 23. September. 2012)

Koda, Keiko. 2007. "Phonology and literacy". In: Pennington, Martha C. (Hrsg.). *Phonology in context*. Basingstoke/New York: Palgrave MacMillan, 219-244.

Kohler, Klaus J. 1995. *Einführung in die Phonetik des Deutschen*. Erich Schmidt Verlag.

Kuhl, P. K. 2000. "A new view of language acquisition". *Proceedings of the National Academy of Sciences USA* 97: 11850–11857.

Ladefoged, P. & M. H. Draper & D. Whitteridge. 1958. "Syllables and stress". *Miscellanea Phonetica* 3: 1-14.

- Ladefoged, Peter & Keith Johnson. 2011. *A course in phonetics*. Wadsworth: Cengage Learning.
- Leopold, Werner F. 1947. *Speech development of a bilingual child. A linguists record*. Northwestern University Press. [Nachdruck (1970) von AMS Press, New York].
- Levelt, W. J. M. & A. Roelofs & A. S. Meyer 1999. "A theory of lexical access in speech production". *Behavioral and Brain Sciences* 22: 1-75.
- Levelt, C. & N. Schiller & W. Levelt. 2000. "The acquisition of syllable types". *Language Acquisition* 8: 237–64.
- Liberman, A.M. & P. C. Delattre & F. S. Cooper & L. J. Gerstman. 1954. "The role of consonant-vowel transitions in the perception of the stop and nasal consonants". *Psychological Monographs* 68: 1-13.
- Liberman, A. M. & F. S. Cooper & D. P. Shankweiler & M. G. Studdert-Kennedy. 1967. "Perception of the speech code". *Psychological Review* 74: 431-461.
- Liberman, A. M. & I. G. Mattingly. 1985. "The motor theory of speech perception revised". *Cognition* 21: 1-36.
- Lindblom, B. 1983. "Economy of Speech Gestures". In: P. MacNeilage (Hrsg.) *The Production of Speech*. New York: Springer-Verlag, 217-245.
- Lindblom, B. 1990. "Explaining phonetic variation: A sketch of the H&H theory". In: W. Hardcastle & A. Marchal (Hrsg.). *Speech production and speech modeling*. Dordrecht/Boston/London: Kluwer Academic Publishers, 403-439.
- Lindblom, B. 1999. "Emergent Phonology". *25th Annual Meeting of the Berkeley Linguistics Society*. University of California, Berkeley.
<http://www.ling.su.se/fon/perilus/1999_14.pdf> (Zugriff: 3. Oktober 2012)
- Lindblom, Björn 2000. "Developmental Origins of Adult Phonology: The Interplay between Phonetic Emergents and the Evolutionary Adaptations of Sound Patterns". *Phonetica* 57, 297–314.
- Lindner, Udo 2009. *Silbenstrukturen im frühen LI-Erwerb*. Dissertation. Christian-Albrechts-Universität zu Kiel. Marburg: Tectum Verlag.
- Locke, J. L. 1983. *Phonological acquisition and change*. New York: Academic Press.
- Luce, P. A. 1986. *Neighborhoods of Words in the Mental Lexicon*. Dissertation. Indiana University.

- Lust, Barbara C. 2006. *Child language acquisition and growth*. Cambridge: Cambridge University Press.
- MacNeilage, P. F. 1998. "The frame/content theory of evolution of speech production". *Behavioral and Brain Sciences* 21: 499-546.
- MacNeilage, P. F. 2008. *The origin of speech*. New York: Oxford University Press.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. 3. Auflage. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mairal, Ricardo & Juana Gil. 2006. "A first look at universals". in: Mairal, Ricardo & Juana Gil (Hrsg.). *Linguistic Universals*. Cambridge: Cambridge University Press, 1-45.
- Marques de Sá & Joaquim P. 2001. *Pattern recognition. Concepts, methods and applications*. Berlin/Heidelberg: Springer.
- Meibauer, Jörg & Monika Rothweiler 1999. "Das Lexikon im Spracherwerb - ein Überblick". In: Meibauer, Jörg & Monika Rothweiler (Hrsg.). *Das Lexikon im Spracherwerb*. Tübingen/Basel: A. Francke Verlag, 9-31.
- Meinschaefer, Judith 1998. *Silbe und Sonorität in Sprache und Gehirn*. Dissertation. Ruhr-Universität Bochum.
- Meltzoff, A. N. 2007. "Infants' causal learning. Intervention, observation, imitation". In: A. Gopnik & L. Schulz (Hrsg.), 37-47.
- Menn, L. & M. Vihman. 2011. "Features in child phonology inherent, emergent, or artefacts of analysis In: Clements, Nick G. & Rachid Ridouane. (Hrsg.). *Where Do Phonological Features Come From? Cognitive, physical and developmental bases of distinctive speech categories*. John Benjamins Publishing Company, 261-301.
- Metsala, J. L. & A. C. Walley 1998. "Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability". In: J. L. Metsala & L. C. Ehri (Hrsg.). *Word recognition in beginning literacy*. Mahwah, NJ: Erlbaum, 89-120.
- Montada, Leo 1970. *Die Lernpsychologie Jean Piagets*. Stuttgart: Ernst Klett Verlag.
- Morton, K. & M. Tatham. 2006. *Speech production and perception*. Basingstoke: Palgrave Macmillan.
- Moskowitz, A. I. 1970. "The acqllsition of phonology". *Working Paper 34*. Language-Behavior Research Laboratory, University of California, Berkeley.

Müller, H. M. 2003. "Neurobiologische Grundlagen der Sprachfähigkeit". In: Deutsch, W. & T. Herrmann & G. Rickheit (Hrsg.). *Psycholinguistik. Ein Internationales Handbuch*. Berlin, 57-80.

Nais, C. 2010. "Origin of language debate". In: Stainton, Robert J. & Alex Barber (Hrsg.). *Concise encyclopedia of philosophy of language and linguistics*. Oxford: Elsevier, 551-556.

Ohala, John J. & Haruko Kawasaki-Fukumori 1997. "Alternatives to the sonority hierarchy for explaining segmental sequential constraints". In: Eliasson, Stig & E. H. Jahr (Hrsg.). *Language and its ecology. Essays in memory of Einar Haugen*. Berlin/New York: Mouton de Gruyter, 343-365.

Ohala, John J. 2008. "The emergent syllable". In: Davis & Zajdó (Hrsg.), 179-186.

Oller, D. Kimbrough & Ulrike Griebel. 2008. "Contextual Flexibility in Infant Vocal Development and the Earliest Steps". In: Oller, D. Kimbrough & Ulrike Griebel (Hrsg.). *Evolution of communicative flexibility. Complexity, creativity, and adaptability in human and animal communication*. Cambridge/MA/London: MIT Press, 141-168.

Ott, R. Lyman 2010. *An introduction to statistical methods and data analysis*. 6. Auflage. Belmont, USA: Brooks/Cole.

Piaget, Jean 1972. *Sprechen und Denken des Kindes*. Düsseldorf: Pädagogischer Verlag Schwann.

Piaget, Jean 1974. *Abriß der genetischen Epistemologie*. Olten, Freiburg: Walter Verlag.

Piaget, Jean 1976. *Psychologie der Intelligenz. Das Wesen der Intelligenz. Die Intelligenz und die sensomotorischen Funktionen. Die Entwicklung des Denkens*. Mit einer Einführung von Hans Aebli. Kindler Taschenbücher. München: Kindler Verlag.

Pierrehumbert, Janet B. 2003a. "Probabilistic Phonology: Discrimination and Robustness" In: Bod, Rens & Jennifer Hay & Stefanie Jannedy (Hrsg.). *Probabilistic Linguistics*. Cambridge/MA/London: MIT Press, 177-228.

Pierrehumbert, Janet B. 2003b. "Phonetic diversity, statistical learning, and acquisition of phonology". *Language and speech* 46: 115-145.

Pike K. & Pike, E. G. 1947. "Immediate constituents in Mazatec Syllables". *International Journal of American Linguistics* 13: 78-91.

Piske, Thorsten 2001. *Artikulatorische Muster im frühen Laut- und Lexikonerwerb*. Tübingen: Gunter Narr Verlag.

Port, R. F. & A. P. Leary 2005. "Against formal phonology". *Language* 81: 927-964.

Poupier, M. 2011. "The atoms of phonological representations". In: van Oostendorp, M. & C. J. Ewen & E. Hume & K. Rice (Hrsg.). *The Blackwell Companion to Phonology*: Blackwell Publishing, 107-129.

Poupier, M. & Beňuš, Š. 2011. "On the phonetic status of syllabic consonants: Evidence from Slovak". *Journal of Laboratory Phonology* 2: 243-273.

Rothweiler, Monika. 1999. "Neue Ergebnisse zum *fast mapping* bei sprachnormalen und bei sprachentwicklungsgestörten Kindern". In: Meibauer, Jörg & Monika Rothweiler (Hrsg.). *Das Lexikon im Spracherwerb*. Tübingen/Basel: A. Francke Verlag, 252-276.

Saji, N. & M. Imai & H. Saalbach & Y. Zhang & H. Shu & K. Okada. 2011. "Word Learning Does Not End At Fast-Mapping: Evolution of Verb Meanings Through Reorganization of An Entire Semantic Domain". *Cognition* 118: 45-61.

Sanford, Anthony J. 2006. "Semantics in Psychology" in: Brown, Keith (Hg.), *Encyclopedia of Language and Linguistics*. Elsevier, 152-158.

Scarborough, R. A. 2004. *Coarticulation and the structure of the lexicon*. Unveröffentlichte Dissertation, UCLA.

Schiller, Niels O. 2008. "Syllables in psycholinguistic theory: now you see them, now you don't". In: Davis & Zajdó (Hrsg.), 155-176.

Schultze, Fritz. 1880. *Die Sprache des Kindes: Eine Anregung zur Erforschung des Gegenstandes*. (Darwinistische Schriften, Folge 1, Band 10). Leipzig: Ernst Günther's Verlag.

Schweitzer, Antje & Bernd Möbius. 2004. "Exemplar-based production of prosody: evidence from segment and syllable durations". In: B. Bel & I. Marlien. (Hrsg.). *Speech Prosody 2004, International Conference*. Nara, Japan, 459-462.

Shattuck-Hufnagel, Stefanie & Dennis H. Klatt. 1979. "The limited use of distinctive features and markedness in speech production: Evidence from speech error data". *Journal of Verbal Learning and Verbal Behavior* 18: 41-55.

Sievers, E. 1893. *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der Indogermanischen Sprachen*. 4. Auflage. Leipzig: Breitkopf & Härtel.

Smith, Neilson V. 2010. *Acquiring phonology. A cross-generational case-study*. Cambridge: Cambridge University Press.

Smith, Neilson V. 1973. *The acquisition of phonology*. Cambridge: Cambridge University Press.

Sommerville, J. A. 2007. "Detecting causal structure: The role of interventions in infants' understanding of psychological and physical causal relations". In: A. Gopnik & L. Schulz (Hrsg.), 48–57.

Spitzer, Manfred 2000. *Geist im Netz. Modelle für Lernen, Denken und Handeln*. Berlin/Heidelberg: Spektrum Akademischer Verlag.

Stampe, David 1979. *A dissertation on natural phonology*. New York/London: Garland Publishing.

Stetson, R. H. 1951. *Motor phonetics. A study of speech movements in action*. Amsterdam: North Holland Publishing Company.

Stemberger, J. P. 1992. "A connectionist view of child phonology: Phonological processing without phonological processes". In: Ferguson, C. A. & L. Menn & C. Stoel-Gammon (Hrsg.). *Phonological development. Models, research, implications*. Tiomium, Maryland: York Press, 165-189.

Stoel-Gammon, Carol. 2010. "Relationships between lexical and phonological development in young children". *Journal of Child Language*: 1-34.

Studdert-Kennedy, M. & Goodell, E.W. (1992), "Gestures, features and segments in early child speech". In: *Haskins Laboratories Status Report in Speech Research 1992*, SR-111/112, 89-102. <http://www.haskins.yale.edu/sr/SR111/SR111_06.pdf> (Zugriff: 3. Oktober 2012)

Taylor, A. H. & R. Miller & R. D. Gray 2012. "New Caledonian Crows reason about hidden causal agents". *PNAS Early Edition*. <<http://www.pnas.org/content/early/2012/09/10/1208724109.full.pdf+html>> (Zugriff: 27. September 2012)

Thiessen, Erik 2009. "Statistical learning". In: Bavin (Hrsg.), 35-50.

Tomasello, M. 2006. *Die kulturelle Entwicklung des menschlichen Denkens. Zur Evolution der Kognition*. Frankfurt am Main: Suhrkamp. [Übersetzung von Jürgen Schröder der englischen Originalausgabe: Tomasello, M. 1999. *The cultural origin of human cognition*. Cambridge/MA/London: Harvard University Press.]

Trubetzkoy, N. S. 1962. *Grundzüge der Phonologie*. 3. Auflage. Göttingen: Vandenhoeck & Ruprecht.

Vater, Heinz. 1992. "Zum Silben-Nukleus im Deutschen". In: Eisenberg, Peter & Karl Heinz Ramers & Heinz Vater. (Hrsg.). *Silbenphonologie des Deutschen*. Tübingen: Gunter Narr, 100-133.

Vennemann, Theo 1978. "Universal syllabic phonology". *Theoretical Linguistics* 5: 175-215.

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change. With special reference to German, Germanic, Italian, and Latin*. Berlin/New York/Amsterdam: Mouton de Gruyter.

Vennemann, Theo & David Restle. 2001. "Silbenstruktur" In: Haspelmath, Martin & Ekkehard König & Wulf Oesterreicher & Wolfgang Raible. (Eds.). *Language Typology and Language Universals, II, Handbücher zur Sprach- und Kommunikationswissenschaft* 20. Berlin/New York: Mouton de Gruyter, 1310-1336.

Vihman, M. M. & R. A. DePaolis & T. Keren-Portnoy 2009. "A dynamic systems approach to babbling and words". In E. L. Bavin (Hrsg.), 163-182.

Vihman, M. M. 2009. "Word learning and the origins of phonological system". In: Foster-Cohen, S. (Hrsg.). *Advances in language acquisition*. Basingstoke: Macmillan, 15-39.

Vihman, M. M. & W. Croft. 2007. "Phonological development Toward a 'radical' templatic phonology". *Linguistics* 45: 683-725.

Vitevitch, M. S. 2002. "The influence of phonological similarity neighborhoods on speech production." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 735-747.

Wedel, Andrew B. 2006. "Exemplar models, evolution and language change". *The Linguistic Review* 23: 247-274.

Wells, J. C. 2012. *Computer-coding the IPA: a proposed extension of SAMPA*. University College London. <<http://www.phon.ucl.ac.uk/home/sampa/ipasam-x.pdf>> (Zugriff: 23.09.2012)

Wester, M. 2003. "Syllable classification using articulatory-acoustic features," In: *Proceedings of Eurospeech*. Geneva, 233-236.

White, John A. 2002. "Action Potential". In: Ramachandran, V. (Hrsg.). *Encyclopedia of the Human Brain*. Bd. 1. Amsterdam: Academic Press, 1-12.

Wiese, Richard. 1988. *Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen*. Tübingen: Max Niemeyer Verlag.

Woodward, J. 2007. "Interventionist theories of causation in psychological perspective". In: A. Gopnik & L. Schulz (Hrsg.), 19-36.

Wurzel, Wolfgang Ulrich 2001. *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie Verlag.