

Dreidimensionale Phonologie und der Erwerb von Protowörtern

Inauguraldissertation

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Neuere Philologien (10)
der Johann Wolfgang Goethe-Universität
zu Frankfurt am Main

vorgelegt von

Sascha Dümig

aus Offenbach am Main

Einreichungsjahr: 2017

Erscheinungsjahr: 2018

1. Gutachter: Prof. Dr. Helen Leuninger

2. Gutachter: Prof. Dr. Dr. Dr. Robert Sader

Tag der Promotion: 21. Juni 2018

Vorwort

Es ist gängiger Brauch, Doktorarbeiten mit einem Feuerwerk an Lobhudelei für Institute, Professoren und Mitarbeiter zu beginnen. Dies erstaunt sehr, wenn man sich die gegenwärtige Realität an deutschen Universitäten vor Augen hält. Da ich selbst das zweifelhafte Vergnügen genießen durfte, im Rahmen meines Studiums und der Tätigkeit als wissenschaftlicher Mitarbeiter über mehrere Jahre wesentliche Einblicke in den Universitätsbetrieb zu erhalten, stellt sich einer einfachen Aneinanderreihung von Danksagungen mein Gewissen als Wissenschaftler, vornehmlich aber als Mensch, entgegen. An einer Institution, wo diejenigen systematisch aussortiert werden, die sich abseits vom eigenen Prestige noch um Werte der wahren Persönlichkeitsbildung, abseits des Impact-Faktors für Publikationen noch um die Wahrheit von Forschung, abseits der eigenen Karriere noch um Gerechtigkeit sorgen, an einer solchen Institution kann allgemeiner Dank nur der eines Knechts für die Ketten seines Herren sein. Ich persönlich verdanke auf meinem akademischen Weg vielen nichts, sehr wenigen dafür umso mehr. Ich möchte es insofern mit Tolstois Worten halten:

„Kümmere dich nicht um den Beifall von Leuten, die du nicht kennst oder die du verachtest.“
Immer noch bin ich aber, trotz der aktuellen Erscheinung der Universität, ungebrochen im Glauben an ihre Idee, so wie Karl Jaspers sie formulierte:

„Die Universität, frei im freien Staat, steht überstaatlich und übernational am Ort der Wahrheit, die in ihrer für unser Zeitdasein allein möglichen Gestalt alle Menschen verbinden würde, nicht in der einen nun wieder dogmatischen Gewußtheit, sondern in der Kommunikation des Vielfältigen durch unvollendbare Bewegungen. Wo Wahrheit wirkt, gibt es Pluralität der geistigen Mächte.“

Frau Prof. Dr. Helen Leuninger, meine Doktormutter und langjährige einzige Mentorin, lebte immer diese Form der Wahrheit und ermöglichte mir in der wissenschaftlichen Freiheit und dem persönlichen Vertrauen, die sie mir über alle Jahre gab, das Eigene zu entwickeln und zu verwirklichen. Für diese unendlich wichtige Möglichkeit der Orientierung im eigenen Denken und ihr gelebtes Vorbildsein als Wissenschaftlerin und Mensch gilt ihr all mein Dank.

Den in dieser Arbeit beobachteten Kindern danke ich, dass sie, wenn auch noch nicht bewusst, es mir mit Freude ermöglicht haben, einen kleinen Beitrag zum wissenschaftlichen Diskurs zu leisten. Ich wünsche ihnen und ihren Eltern für die Zukunft von Herzen alles erdenklich Gute.

Inhaltsverzeichnis

Einleitung.....	1
1 Phonologische Repräsentationen.....	5
1.1 Zur Realität phonologischer Repräsentationen.....	5
1.2.1 Modulare Sprachverarbeitung.....	13
1.2.2 Modularität im Spracherwerb.....	18
1.2.3 Die Irrelevanz konnektionistischer Modelle.....	22
1.3 Die Struktur phonologischer Repräsentationen.....	26
1.3.1 Phonologie als sprachliches Submodul mit spezifisch repräsentationaler Verarbeitung.....	26
1.3.2 Prosodische Phonologie vs. Segmentale Phonologie.....	30
1.3.3 Dreidimensionale Phonologie.....	34
1.4 Sonorität.....	48
1.4.1 Definitive Annäherungen.....	48
1.4.2 Sonorität und phonologische Komplexität nach Clements.....	56
1.5 Sprachproduktionsmodelle und 3D-Phonologie.....	65
1.5.1 Das Sprachproduktionsmodell von Garrett.....	65
1.5.2 Die Berechnung von Silbenpositionen und segmentaler Information.....	67
1.6 Sprachplanung und -verarbeitung, Gedächtnissysteme und die Parallele Architektur.....	77
1.7 Reduplikation.....	86
1.7.1 Precedence-Based Phonology.....	86
1.7.2 Distributed Reduplication.....	88
2 Der Erwerb phonologischer Strukturen.....	92
2.1 Der Erwerb in mentalistischen Ansätzen.....	92
2.1.1 Structure-elaborating constituency model.....	94
2.1.2 Strong Identity Hypothesis	98
2.2 Der Erwerb auf einzelnen Repräsentationsebenen.....	103
2.2.1 Der Erwerb von Betonungsmustern.....	103
2.2.2 Erwerb der Silbenstruktur.....	114
2.2.2.1 Demisilbentheorie und der Erwerb von Silbenpositionen.....	123
2.2.3 Sonorität im Erwerb.....	126
2.2.4 Reduplikation im Erwerb.....	129
2.2.4.1 Precedence-Based Phonology im Spracherwerb.....	133
3 Protowörter.....	139
3.1 Phasen der Protowortforschung.....	140
3.1.1 Prästrukturalistische Phase.....	140
3.1.2 Strukturalistische Phase.....	141
3.1.3 Nachstrukturalistische Phase	142
3.2 Semantische und lexikalische Aspekte der Protowortproduktion.....	154
3.2.1 Erste Klassifikationen in der Spracherwerbsforschung.....	154
3.2.2 Das Problem der semantischen Analyse allgemein und im Spracherwerb.....	158
3.2.3 Fodor und die Komponentialsemantik.....	163
3.2.4 Ingrams Analysen: Erste Annäherungen an die strukturelle Elaboriertheit früher Wortbedeutungen.....	168
3.2.5 Die konzeptuelle Struktur von Protowörtern und ihre Integration in die Parallele Architektur.....	173
3.2.6 Zeigegesten und ihr Zusammenhang mit dem frühen Worterwerb.....	186
3.3 Protowörter und morphosyntaktische Strukturen: Der Fall P.....	189
3.3.1 P.s phonologische Entwicklung.....	191
3.3.2 P.s morphosyntaktische Entwicklung.....	194
3.3.3 Der Fall P. unter modelltheoretischer Perspektive.....	200
3.4 Protowortproduktion: Eine vorläufige modelltheoretische Synthese.....	202
3.4.1 Protowortproduktion im erweiterten Garrett-Modell.....	202

3.4.2 Die Protowortproduktion als Indikator informationeller Integrationsstufen.....	206
4 Die empirische Untersuchung der Protowortproduktion.....	211
4.1 Methode.....	211
4.1.1 Zielsetzungen und Ausgangshypothesen.....	211
4.1.2 Methodische Begründung des Untersuchungsverfahrens: Einzelfallstudien.....	212
4.1.3 Methodische Begründung des Untersuchungsverfahrens: Spontansprachuntersuchungen	218
4.2 Durchführung.....	220
4.2.1 Datenerhebung.....	220
4.2.2 Datenbearbeitung.....	220
4.2.3 Spezifische Fragestellungen und Hypothesen.....	222
4.3 Ergebnisse.....	224
4.3.1 Qualitative Auswertung der Sprachdaten von Kind J.....	224
4.3.2 Quantitative Auswertung der Sprachdaten von Kind J.....	237
4.4 Diskussion.....	246
5 Schlussbetrachtungen.....	255
6 Literaturverzeichnis.....	260
Anhang A.....	284

Abbildungsverzeichnis

Abbildung 1: Grundriss der Parallelen Architektur (Jackendoff 2002).....	17
Abbildung 2: Lateralisierung prosodischer Funktionen (nach Sidtis & Van Lancker Sidtis 2003: 98).	32
Abbildung 3: Degemination, Vokalepenthese und Vokalkürzung im Türkischen.....	37
Abbildung 4: Maximale Silbenstruktur nach Wiese (1996).....	37
Abbildung 5: Zugrundeliegende Form von 'Mein Bruder hat nicht gegeben' /wci/.....	40
Abbildung 6: Silbifizierungsvariante 1 /ur y-uci wma/.....	40
Abbildung 7: Silbifizierungsvariante 2 /ur i-wci wma/.....	40
Abbildung 8: Splitting nach Goldsmith (1990).....	41
Abbildung 9: Prosodische Hierarchie.....	42
Abbildung 10: Beispiel einer dreidimensionalen Struktur.....	44
Abbildung 11: Dreidimensionale Repräsentation im Südlichen Paiute (ohne Segmentebene).....	44
Abbildung 12: Teiltheorien der Phonologie.....	47
Abbildung 13: Beispiele für Sonoritätshierarchien.....	48
Abbildung 14: Silbenstruktur von VATER.....	50
Abbildung 15: Silbenstruktur mit Sonoritätskontur nach Neef (2004).....	51
Abbildung 16: Der Erwerb von komplexen Silbenonsets über Affrikata.....	52
Abbildung 17: Kompensatorische Längung über die X-Ebene.....	53
Abbildung 18: Diphtongisierung nach Hayes und Selkirk.....	54
Abbildung 19: Sonoritätszyklus des Wortes template (nach Clements 1990).....	57
Abbildung 20: Entwicklung des protogermanischen Diphtongs im Altenglischen.....	61
Abbildung 21: Struktur- und Demisilbenangabe für Langvokale und Diphtonge in der Two-Root Theory.....	62
Abbildung 22: Grundaufbau einer phonologischen Merkmalsgeometrie.....	63
Abbildung 23: Merkmalsgeometrie für Konsonanten nach Clements & Hume (1995).....	64
Abbildung 24: Das Sprachproduktionsmodell von Garrett (1975, 1980).....	66
Abbildung 25: Erweitertes Garrett-Modell mit Input- und Output-Lexikon.....	70
Abbildung 26: Dreistufig konsekutive Top-down-Berechnung von phonologischer Information mit Fehlplanungstypen.....	75
Abbildung 27: Schalenmodell des phonologischen Merkmalausbaus.....	76
Abbildung 28: Phonologisches Arbeitsgedächtnis nach Baddeley & Hitch (1974).....	78
Abbildung 29: TOTE-Einheiten nach Miller et al. (1960).....	79
Abbildung 30: MEKIV nach Hussy (1993).....	80
Abbildung 31: Monotone Zahlenlistung (oben) vs. Chunking im Kurzzeitgedächtnis (unten).....	81
Abbildung 32: Mögliche Rückkopplungsschleife nach Anwendung eines Operators.....	81
Abbildung 33: Die Parallele Architektur nach Jackendoff (2002).....	83
Abbildung 34: Erwerbshierarchie phonologischer Merkmale nach Rice & Avery (1995).....	94
Abbildung 35: Ortsmerkmale bei Plosiven.....	95
Abbildung 36: Kontrastentwicklung des Merkmals Place.....	96
Abbildung 37: Phonologisches Segmentinventar bei Ausbau des Artikulationsartmerkmals.....	97
Abbildung 38: Phonologisches Segmentinventar bei Ausbau des Ortskontrastes.....	97
Abbildung 39: Das Subset Principle im traditionellen Modell.....	98
Abbildung 40: Traditionelle Interpretation des Subset Principles (links) vs. reinterpretierte Version.	99
Abbildung 41: Die Struktur des minimalen Wortes.....	104
Abbildung 42: Prosodisches Parsing-Modell nach Archibald (1995).....	106
Abbildung 43: Erwerbsreihenfolge und Optionen im Silbenstrukturausbau des Deutschen und Englischen.....	121
Abbildung 44: Erwerbsreihenfolge von Silbenpositionen mit Kennzeichnung von Ressourcenbeanspruchung.....	124
Abbildung 45: Aufbau der Lexikons nach Levelt (1998).....	152

Abbildung 46: Verknüpfung von Subsystemen in der Produktion von frühen Protowörtern.....	157
Abbildung 47: Beispielrepräsentation für Stein nach Ingram (1971: 890).....	169
Abbildung 48: Repräsentationen von wa-wa nach Ingram (1971: 896).	169
Abbildung 49: Repräsentation von be-be nach Ingram (1971: 897).....	170
Abbildung 50: Repräsentation von up nach Ingram (1971: 906).....	170
Abbildung 51: Veränderung von Wortbedeutungen nach Carter (1975).....	174
Abbildung 52: Die Parallele Architektur erweitert um die räumliche Struktur (SpS), Modalitäten und interne Registrierungskomponenten.....	182
Abbildung 53: Frequenzangaben initialer Demisilben in P.s Produktion.....	192
Abbildung 54: Frequenzangaben finaler Demisilben in P.s Produktion.....	192
Abbildung 55: Verknüpfung von Subsystemen in der späten Protowortproduktion von P.....	201
Abbildung 56: Protowortproduktion im erweiterten Garrett-Modell.....	204
Abbildung 57: Etablierte Produktionspfade von Protowörtern der 1. Integrationsstufe.....	207
Abbildung 58: Etablierte Produktionspfade von Protowörtern der 2. Integrationsstufe.....	207
Abbildung 59: Etablierte Produktionspfade von Protowörtern mit morphosyntaktischer Information.	208
Abbildung 60: Schema des Zusammenhangs von Theorie und Evidenz.....	213
Abbildung 61: Schema des Zusammenhangs von Theorie und Evidenz mit lokaler Modifikation Li des kognitiven Systems (rechts).....	214
Abbildung 62: Silbenstruktur des Protoworts [ʔaw] in Sitzung 3.....	225
Abbildung 63: Ambisilbische Struktur des Protoworts [ʔawa].....	226
Abbildung 64: Struktur einer Oberflächenverschränkung nach Hohenberger (2002: 215).....	229
Abbildung 65: Speicherung von Chunks auf Basis einer Simulatorstruktur (X = Einheit auf X- Ebene, O = Onset, N = Nukleus).....	231
Abbildung 66: Partielle Reduplikation des Vokals in [ʔawa].....	232
Abbildung 67: Freischaltungs- und Transferprozesse von Proto- zu Zielwörtern.....	236
Abbildung 68: Tokens Protowörter und Zielwörter.....	238
Abbildung 69: Entwicklung W-NW-Ratio.....	238
Abbildung 70: Produktion von CV- und CVC-Silben in Proto- und Zielwörtern.....	239
Abbildung 71: Produktion von Reduplikationen in Protowörtern und Zielwörtern.....	240
Abbildung 72: Absolute Häufigkeit initialer und finaler Demisilben in Proto- und Zielwörtern...	240
Abbildung 73: Häufigkeit initialer Demisilben in Protowörtern.....	241
Abbildung 74: Häufigkeit finaler Demisilben in Protowörtern.....	242
Abbildung 75: Häufigkeit initialer Demisilben in Zielwörtern.....	243
Abbildung 76: Häufigkeit finaler Demisilben in Zielwörtern.....	243
Abbildung 77: Produktion semantischer Klassifikationstypen in Proto- und Zielwörtern.....	244
Abbildung 78: Verwendung von Zeigegesten in interaktionalen Zielwörtern.....	245
Abbildung 79: Sublexikon-Modell des Japanischen.....	254

Tabellenverzeichnis

Tabelle 1: Oberklassenmerkmale und Sonoritätsgrad nach Clements (1990).....	57
Tabelle 2: Dispersionswerte (D) und Komplexitätsgrade (C) von Demisilben nach Clements (1990).	59
Tabelle 3: Wesentliche Ergebnisse der Versprecherforschung zusammengefasst nach Poulisse (1999).....	67
Tabelle 4: Metrisches Erwerbsmodell nach Maßgabe von Prominenz und Berechnungskapazität.	110
Tabelle 5: Merkmalswerb im Onset nach Fikkert (1994).....	117
Tabelle 6: Silbentypen im Datenkorpus von Weir (nach Ingram 1978).....	118
Tabelle 7: Silbentypen im Erwerb von Jennika (nach Ingram 1978).....	119
Tabelle 8: Silbentypen im Erwerb von Anthony (nach Ingram 1978).....	119
Tabelle 9: Zusammenfassung von Annalenas Erwerb der Silbenstruktur.....	121
Tabelle 10: Systemanforderung nach Maßgabe von Positionen und Vokalkomplexität.	123
Tabelle 11: Silbentypen in verschiedenen Sprachen (nach Levelt & Van de Vijver 2004).....	125
Tabelle 12: Hypothesen im Erwerb (nach Raimy 2009).....	135
Tabelle 13: Typenanalyse von Reduplikationen in Stern & Stern (1928).....	136
Tabelle 14: Produzierte PCFs und Zielwörter (modifiziert nach Dore et al. 1976).....	145
Tabelle 15: Entwicklung der Wort-Nichtwort-Ratio in der Studie von Robb et al. (1994).....	146
Tabelle 16: Protowörter von Kind Will.....	148
Tabelle 17: Zusammenfassung der Produktionsmuster in Stoel-Gammon & Cooper (1984).....	149
Tabelle 18: Silbenmerkmale nach Ferguson et al (1973).....	150
Tabelle 19: Frühes domänenspezifisches Kernwissen nach Pauen & Vonderlin (2007: 22).....	160
Tabelle 20: Semantische Paraphasien und Merkmalsanalyse (nach Peuser 2000: 120).....	167
Tabelle 21: Übersicht über Davids produzierte sensomotorische Morpheme und ihrer kommunikativen Verwendung im Alter von 1;0 bis 1;4 (Foster 1990: 39).	175
Tabelle 22: Merkmalsanalyse der acts of meaning.....	185
Tabelle 23: P.s Wortschatz im Alter von 5;5 (Dümig & Leuninger: 86-87).	191
Tabelle 24: Phasen des Demisilbenerwerbs (nach Dümig & Leuninger 2013: 93).....	193
Tabelle 25: Klassifikationsschema der Protowortproduktion.....	208
Tabelle 26: Klassifikation von dokumentierten Protowörtern nach Maßgabe der Parallelen Architektur.....	211
Tabelle 27: Gegenüberstellung der Vor- und Nachteile von Checkliste-/Fragebogen-Methode und Spontansprachuntersuchung nach Kauschke (2000: 77-78).....	218
Tabelle 28: Wörter von J. in Sitzung 1.....	224
Tabelle 29: Kontrastentwicklung in J.s Produktionen.....	237

Einleitung

Ein heimlicher Burgfrieden besteht in der aktuellen Spracherwerbsforschung hinsichtlich der Frage, ab wann frühe Wortproduktionen phonologisch analysiert werden können. Maßgeblich wirkt hier Jakobsons Entwurf eines Zwei-Phasen-Erwerbsmodells des kindlichen Lautsystems bis heute nach (Jakobson 1941). Die Einteilung in eine phonetische Phase, in der vorsprachliche Lalllaute produziert werden, und einer eigentlichen Sprachstufe, in der es „um die Auslese der Sprachlaute und um den Aufbau des Phonemsystems geht“ (ebd.: 30), ist aktuellen Theorien zum Phonologieerwerb in zweifacher Form immanent.

Phonetisch orientierte Ansätze gehen davon aus, dass die ersten Produktionen im frühen Spracherwerb auf zugrundeliegende artikulatorische Muster (engl.: templates), meist in Form von unanalysierten Worteinheiten, zurückzuführen sind, welche die Grundlage für spätere Abstraktionsprozesse und Segmentierungen im Sinne einer Phonologisierung bilden (vgl. z.B. Piske 2001; Vihman 1996; Vihman & Croft 2007). Diese Ansätze betonen den kontinuierlichen Übergang von der sogenannten Lallphase zu ersten Zielwortproduktionen (vgl. Vihman 1985), ohne jedoch Jakobsons diskontinuierliches Zwei-Phasen-Modell grundsätzlich aufzugeben. Diesem Ansatz nach emergieren phonologische Muster sukzessiv aus phonetischen Formen.

Eine solche Diskontinuität bildet auch die Hintergrundfolie für psycholinguistische Studien zur frühen Sprachproduktion, welche dem Forschungsparadigma der Generativen Grammatik folgen (vgl. z.B. Fikkert 1994; Kehoe & Stoel-Gammon 2001; Kehoe & Lleó 2003). In diesen Studien steht die linguistisch-phonologische Analyse der Wortproduktionen im Vordergrund, wobei die Daten meist erst ab der Zielwortproduktion erhoben werden, ab einem Zeitpunkt also, zu dem die produzierten Phone der Kinder den Annahmen des Prager Strukturalismus nach auf paradigmatischer Ebene erkennbare phonemische Oppositionen bilden. Maßstab der Beurteilung ist dementsprechend das Phonem der Ziel- bzw. Erwachsenensprache und Jakobson zufolge wird hier die Zielwortproduktion als die eigentliche Sprachstufe verstanden, auf der die „real phonology“ (Smith 2010) beginnt. Vorherige Produktionen werden gemäß dieser Forschungslogik als rein phonetische Formen meist ignoriert. Dies ist ein Grund dafür, dass systematisch verwendete frühe Produktionen ohne overten lautlichen Bezug zu Zielwörtern, sogenannte Protowörter, zwar in phonetisch orientierten Ansätzen untersucht wurden (vgl. z.B. Menn 1983; Vihman 1996) und die Frequenz ihrer Produktion mit der Erwerbsrate von Zielwörtern verglichen wurde (vgl. z.B. Stoel-Gammon & Cooper 1984),

dies aber ohne Zugrundelegung einer ausgearbeiteten phonologischen Theorie. Zusätzlich zu der Tatsache, dass die Generative Phonologie in ihrer Bewertung früher Wortproduktionen im Spracherwerb weiterhin grundlegend an der Nabelschnur des Prager Strukturalismus hängt, legen reduktionistische Konzepte aus anderen Forschungsrichtungen ebenfalls nahe, dass eine auf Systematizität phonologischer Repräsentationen ausgelegte Untersuchung früher Wortproduktionen ein fruchtloses Unterfangen sei. So z.B. das aus den Neurowissenschaften kommende konzeptuelle Mantra des plastischen Gehirns und mit ihm die Anschauung einer Genese, an deren Ende erst, wenn denn überhaupt, eingrenzbare Systemeigenschaften im Sinne eines kognitiven Moduls untersuchbar seien. Wie bei der Theorie artikulatorischer Muster versteckt sich hinter dieser Anschauung die Annahme, dass eine kategorial operierende Perzeption und Produktion kognitiver Systeme das Resultat eines Filter- und Übungsprozesses sei, der entsprechend von einer anfänglichen physikalisch bestimmbar „Fuzziness“ zu stabilen Zuständen (im Sinne von reziprok und reentrant verschalteten neuronalen Netzwerken) voranschreitet. Solchen Unbestimmtheitsannahmen, die einer Untersuchung von Protowörtern als Produktionen eines kognitiven Phonologiemoduls entgegenstehen, soll in der vorliegenden Arbeit entgegengewirkt werden und erstmalig eine phonologische Definition der Klasse der Protowörter innerhalb eines psycholinguistischen Paradigmas vorgenommen werden. Ein solch ambitioniertes Unternehmen, bei dem ein breites Spektrum inner- und interdisziplinärer Forschung herangezogen und teilweise reanalysiert werden muss, verlangt eine ausführliche Fundierung.

In **Kapitel 1** gilt es deshalb zuerst eine formale und inhaltliche Bestimmung phonologischer Repräsentationen und eine Standortbestimmung derselben im Gesamt der Sprachverarbeitung vorzunehmen. Zu diesem Zwecke erfolgt zuvorderst eine wissenschaftstheoretische Argumentation für die Realität und Notwendigkeit phonologischer Repräsentationen in Rekurs auf Chomskys Mentalismus (vgl. Chomsky 1980), Fodors Computational Theory of Mind (vgl. Fodor 1987) und Bromberger & Halles Konzeption einer Ontologie der Phonologie (Bromberger & Halle 2000). Ist diese in Abgrenzung zum strukturalistischen Paradigma vorgenommen, wird weiterhin die Verarbeitung kognitiver Repräsentationen innerhalb eines modular strukturierten kognitiven Systems beleuchtet und das Konzept der Modularität auf seine Gültigkeit für den frühen Spracherwerb hinterfragt. Ist das „Wie“ von phonologischen Repräsentationen in Bezug auf ihre interne Verarbeitung geklärt, soll das „Was“ genauer bestimmt werden. D.h., welche strukturellen Formate der Phonologie können für eine Untersuchung des frühen Spracherwerbs als wegweisend gelten? Hier ist das Konzept

der dreidimensionalen Phonologie maßgeblich, in das die Berechnung von Komplexitätsgraden nach Silbenposition und Sonorität nach Clements (1990) sowie neuere Ansätze zur reduplikativen Prozessen optimal integriert werden können.

Mit dem so entfalteten Hintergrund dieser (psycho)-linguistischen Prolegomena sollen in **Kapitel 2** aktuelle Ansätze zum Erwerb bzw. Ausbau phonologischer Strukturen dargestellt, analysiert und letztlich kritisch komplementiert bis revidiert werden. Besonders fokussiert werden soll hier die Strong Identity Hypothesis von Hale & Reiss (2008) mit ihrer Reaktualisierung der klassischen Unterscheidung zwischen Kompetenz und Performanzsystemen. Im Lichte dieser Unterscheidung wird gezeigt, dass viele Analysen gängiger Erwerbsmodelle in ihrer theoretischen Fundierung fragmentarisch bleiben und unter Zugrundelegung einer arbeitsgedächtnisbasierten Produktion dreidimensionaler Strukturen adäquater vorgenommen werden können. Dieser Ansatz, der grundsätzlich von einer sukzessiven, auf mehreren Strukturebenen stattfindenden, Integration schon vorhandener Sprachkenntnis (im Sinne von Kompetenz) in die Prozesse der Sprachproduktion (im Sinne von Performanz) ausgeht, steht eo ipso top-down-basierten und zweidimensional konzipierten Erwerbsmodellen entgegen. Hier werden vor allem die Untersuchungen und Ergebnisse von Fikkert (1994) sowie Demuth und Fee (1995) beleuchtet.

Das analytische Instrumentarium des hier vertretenen Ansatzes soll im Zuge dieser Arbeit auf Protowörter angewendet werden. Bei der Bestimmung dieser relevanten Datenklasse in **Kapitel 3** soll die definitorische Vagheit bisheriger Ansätze herausgestellt werden. Hier werden in einem wissenschaftshistorischen Abriss drei grundsätzliche Phasen klassifikatorischer Zuordnung von Protowörtern aus der Menge bestehender Studien heraufdestilliert. Vor allem oben benannte Problemstellung des strukturalistischen Zwei-Phasen-Erwerbsmodell und seiner andauernden Residuen gilt es hier detailliert zu erfassen, so dass ein neues Erklärungsmodell, das auf den vorherigen Ausführungen beruht, mit seinen distinkten Komponenten und Prozessebenen erstellt werden kann, aus dem sich in einer optimal engen Formulierung optimal falsifizierbare Hypothesen herleiten lassen, die in **Kapitel 4** durch eine Spontansprachuntersuchung überprüft werden sollen. Methodisch wird sowohl eine linguistisch-qualitative als auch eine statistisch-quantitative Analyse der Sprachdaten des Kindes J. vorgenommen, die über den Zeitraum eines Jahres für 30 Minuten wöchentlich erhoben wurden. Nach der Diskussion der gewonnenen Ergebnisse und ihren möglichen Konsequenzen für eine Reinterpretation der Bedeutung von Protowörtern innerhalb von Spracherwerb und Sprachproduktion und psycholinguistischer Analysen

allgemein schließt die vorliegende Arbeit mit definitorischen Überlegungen und möglichen Schlussfolgerungen für Forschung und Therapie in **Kapitel 5** ab.

„One of the major concerns of laboratory phonology is that of determining the nature of the transition between discrete phonological structure (conventionally, “phonology”) and its expression in nondiscrete physical or psychoacoustic parameters (conventionally, “phonetics”). [...] a complete and fully adequate theory of the phonetics/phonology interface remains to be worked out. A new synthesis of the methodology of phonology and phonetics, integrating results from the physical, biological and cognitive sciences, is required if we are to make significant progress in this area.” (Clements 1990: 283)

1 Phonologische Repräsentationen

1.1 Zur Realität phonologischer Repräsentationen

Kann aus der strikten Sammlung und Klassifikation von sprachlichen Äußerungen individueller Sprecher nicht nur eine bloße Beschreibung von Daten gegeben, sondern auch etwas über die objektiven Regularitäten einer Sprache ausgesagt werden? Die Verneinung, dass ein solch korpuslinguistischer Forschungsansatz hinreichend ist, um die Kreativität, das Erzeugungspotential menschlicher Sprachfähigkeit, zu erfassen, war und ist grundlegend in die Theorie der Generativen Grammatik eingeschrieben. Der wesentliche Untersuchungsgegenstand von Sprache als System ist hier die Kenntnis, über die ein jeweiliger Sprecher implizit verfügt, die als Kompetenz bezeichnet wird. Nur die Realität derselben verbürgt eine adäquate Erklärung dessen, was Humboldt als den unendlichen Gebrauch von endlichen Mitteln bezeichnet hat (Humboldt 1998 (1836): 221). Das so grob umrissene Forschungsprogramm der Generativen Grammatik steht im wissenschaftlichen Diskurs nicht isoliert, obige Eingangsfrage könnte breiter formuliert folgendermaßen lauten:

Kann aus der strikten Sammlung und Klassifikation physikalischer Erscheinungen etwas über die ihnen zugrundeliegenden Gesetzmäßigkeiten ermittelt werden?

In einer solchen Umformulierung wird deutlich, weshalb in der Generativen Grammatik der sogenannte Galileische Stil eine zentrale Rolle einnimmt. Im Wesentlichen geht dessen Bestimmung auf Edmund Husserl zurück. Dieser betont in *Erfahrung und Urteil* (1939) den Unterschied zwischen Erfahrungsurteilen und wissenschaftlichen Urteilen:

„„Erfahrungsurteile“, deutlicher gesprochen Urteile, die nur aus den ursprünglichen Erzeugungen in kategorialen Akten gewonnen sind, rein auf Grund der Erfahrung, der sinnlichen und der in ihr fundierten Erfahrung geistigen Seins, sind keine endgültigen Urteile, sind keine Urteile der Wissenschaft im prägnanten Sinne – eben der Wissenschaft, die unter Ideen der Endgültigkeit arbeitet.“ (ebd.: 41-42)

Eigentlich wissenschaftliche, d.h. endgültige Urteile, erlangt man nach Husserl und in Rückbesinnung auf Galileo durch die Methode (und als solche will Husserl sie auch nur verstanden wissen; vgl. ebd.: 43) der Idealisierung von sinnlichen Erfahrungen. Diese leistet nichts Geringeres, „als eine ins Unendliche erweiterte Voraussicht des in der Erfahrung zu Erwartenden“ (ebd.: 42). Noam Chomsky hat als Begründer der Generativen Grammatik die Rückbesinnung auf Galileo und die Anwendung seiner Methode in den Humanwissenschaften unentwegt stark gemacht:

„What was striking about Galileo, and was considered very offensive at the time, was that he dismissed a lot of data; he was willing to say „Look, if the data refute the theory, the data are probably wrong“... But the Galilean style ... is the recognition that it is the abstract systems that you are constructing that are really the truth [...].“ (2002: 98 vgl. auch Chomsky 1981: 16)

In diesem Zusammenhang sieht Chomsky die Frage nach der Realität sprachlicher Repräsentationen nicht unterschieden von der Frage nach der Realität empirisch nicht direkt messbarer, aber für eine adäquate theoretische Erklärung notwendiger Modelle in der Physik. An einem sehr schönen Beispiel, der physikalischen Untersuchung thermonuklearer Reaktionen innerhalb der Sonne, illustriert er die Äquivalenz von physikalischer und linguistischer Methodik (Chomsky 1981a: 191-193). Steht durch bestimmte Voraussetzungen nur indirekte Evidenz zur Verfügung, hier aufgrund verfügbarer Techniken die Untersuchung ausgesendeten Lichts ausschließlich an den äußeren Schichten der Sonne, so kann eine erklärende Theorie mit Annahme bestimmter Entitäten und Ereignissen, die zur Entstehung der Sonnenwärme führen, nur in Hinblick auf die gegebene Evidenz physikalische Realität beanspruchen. Selbst bei Bestätigung der erklärenden Theorie durch eine direktere Methode der Untersuchung ändert sich der Status der postulierten physikalischen Realität der angenommenen Entitäten und Ereignisse nicht, da in den empirischen Wissenschaften keine Evidenz in dem Sinne absolut ist, als dass durch einen solchen Standard physikalische Realität verbürgt werden könne. Ähnlich verhält es sich nun bei der Untersuchung des Sprachvermögens:

„Wir beobachten, was Menschen sagen und tun, wie sie sich sprachlich und nicht sprachlich mit anderen auseinandersetzen; wobei die Situationen häufig so beschaffen sind, dass dieses Verhalten (hoffentlich) einige Evidenz über die operativen Mechanismen liefert. Wir versuchen dann, so gut wir können, eine Theorie von einiger Reichweite und empirischem Gehalt über diese Mechanismen zu entwickeln, und beurteilen unsere Theorie nach ihrem Erfolg bei der Erklärung ausgewählter Phänomene. Würden wir zum Nachweis aufgefordert, dass die in der Theorie angenommenen Konstrukte „psychologische Realität“ haben, können wir nichts weiter tun, als die Evidenz und die Erklärungen, welche diese Konstrukte enthalten, zu wiederholen.“ (ebd.: 193)

Für die Untersuchung sprachlicher Daten gilt also, dass ein Modell des menschlichen Sprachvermögens sich nicht in einer reinen Sammlung derselben erschöpfen kann und darf. Wird auf der Ebene empirischer Datensets verblieben, so sind grundsätzlich zwei Stoßrichtungen der wissenschaftlichen Untersuchung gegeben. Einerseits eine bloß quantitative Erfassung, in der entsprechend der Faktor Frequenz das maßgebliche Kriterium

darstellt, andererseits ein taxonomisches Vorgehen, welches die Regularitäten der Sprache auf der Ebene der Daten selbst verortet und mittels der Klassifikation von Distributionen und Oppositionen von sprachlichen Einheiten als externale Objekte eine strukturelle Beschreibung der jeweiligen Sprache vornimmt. Beide Richtungen, erstere in Form des Behaviorismus von Skinner (1957), die zweite in Gestalt des amerikanischen Strukturalismus in der Tradition Bloomfields (1933), sind von Chomsky in ihren wesentlichen Annahmen ad absurdum geführt worden. An dieser Stelle kann keine umfassende Darstellung der Kritikpunkte erfolgen, allerdings soll ein philosophisches Kernargument Chomskys illustrieren, was die Demarkationslinie zwischen physikalistischen und mentalistischen Ansätzen darstellt. Eine Eigenschaft, die Sprache wesentlich charakterisiert, ist ihr kreatives Potential. Chomsky (1975a) verweist darauf, dass jedem Sprachlerner nur eine umschriebene Menge von Sätzen der jeweiligen Zielsprache zugänglich ist, er allerdings nicht auf die Produktion derselben beschränkt bleibt:

„A speaker of a language has observed a certain limited set of utterances in his language. On the basis of this finite linguistic experience he can produce an indefinite number of new utterances which are immediately acceptable to other members of his speech community. He can also distinguish a certain set of 'grammatical' utterances, among utterances that he has never heard and might never produce. He thus projects his past linguistic experience to include certain new strings while excluding others.“ (ebd.: 61)

Die entscheidende Folgerung von Chomsky ist nun, dass das kreative Potential auf einen, als kognitiv zu charakterisierenden, generativen Erzeugungsmechanismus im Geist des Sprechers zurückzuführen sei. Im Sinne der Computertheorie des Geistes¹ ist dieser Erzeugungsmechanismus im Gehirn implementiert und durch Zustände desselben realisiert. Es gilt hier allerdings keine notwendige Eins-zu-eins-Beziehung zwischen einem mentalen Zustand und seiner physischen Implementierung. Analog zum Funktionalismus Hilary Putnams' (z.B. Putnam 1960) ist die multiple Realisierbarkeit mentaler bzw. sprachlicher Repräsentationen wesentlich:

„For the purpose of linguistic investigation, it is immaterial whether the mechanism inside the speaker's head is in reality a net-work of electronic relays, a mechanical system of cardboard flip flops and rubber bands, or, for that matter, a group of homunculi industriously at work in a tiny office. All of these possibilities, and others, are on a par for the linguist as physical

¹ Es gilt an dieser Stelle zu betonen, dass Chomsky selbst die Analogie von Geist und Gehirn mit der Software und Hardware eines Computers immer sehr vorsichtig gebraucht: „The metaphors are harmless if not taken too seriously, but it should be borne in mind that the proposed analogues are much more obscure than the original they are supposed to clarify. The hardware-software distinction raises all sorts of problems that do not arise in the case of an organic object. [...] So the metaphor should not be pressed beyond the point where it may be helpful.“ (Chomsky 1996: 12-13)

realizations of this mechanism, so long as each is isomorphic to the representation of linguistic structure given by the theory of the language. The critical distinction is, then, between an abstract, formal characterization of linguistic structure - the theory itself - and a physical system of some kind which instances this structure.“ (Katz 1964: 129)

Erst durch diese strikte Redefinition von Sprache als externem Geschehen zu einem mentalen Kenntnissystem² wurden also kognitive Verarbeitungsprozeduren und mit ihnen der Unterschied zwischen repräsentationalen Ebenen möglich.³ Solchermaßen konnte der Ableitungsprozess von Oberflächenstrukturen aus Tiefenstrukturen modelliert werden. Diese Unterscheidung war für die Bestimmung sprachlicher Kreativität wesentlich, denn nach Chomsky kann die Oberflächenorganisation eines Satzes keine volle Repräsentation der grammatischen Beziehungen geben, die bei der Bestimmung seines semantischen Inhalts eine Rolle spielen (vgl. z.B. Chomsky 1981b). Dem simplen Fakt, dass in divers realisierten Sätzen die gleichen semantischen Relationen encodiert sind, konnte so durch internale Repräsentationen Rechnung getragen werden.

Während sich diese Argumentation für die Kreativität der Sprache hauptsächlich auf die syntaktische Komponente mit ihren rekursiven Erzeugungsregeln bezog und sich hier durchsetzte, galt Ähnliches nicht für die Phonologie. Da die phonologische Komponente das sprachliche Subsystem ist, welches direkt an der Schnittstelle zu auditiven und artikulatorischen Systemen verortet ist, muss für eine „Galilean-style Phonology“ (Hale 2007) entsprechend eine Abgrenzung zwischen physikalischer Implementierung und Realisierung auf der einen Seite und kognitiver Verarbeitung auf der anderen mit hinreichender Trennschärfe herausgearbeitet werden.

Die direkte Schnittstelle phonologischer Repräsentationen zu physikalisch messbaren Ereignissen in Artikulation und Audition führte in Vergangenheit und Gegenwart immer wieder zu Versuchen, Phonologie in die Phonetik zu integrieren, sei es nun im Sinne einer direkten Ableitung phonologischer Repräsentationen von artikulatorischen oder auditiven

² Chomsky (1986) gibt eine detaillierte Beschreibung des Zusammenhangs von mentalen Zuständen und Gehirnzuständen und definiert gleichzeitig die Geltungsbereiche von Linguistik und Neurowissenschaften: „Knowing the language L is a property of a person H; one task of the brain sciences is to determine what it is about H's brain by virtue of which the property holds. We suggested that for H to know the language L is for H's mind/brain to be in a certain state; more narrowly, for the language faculty, one module of this system, to be in a certain state S_L. One task of the brain sciences, then, is to discover the mechanisms that are physical realization of the state S_L.“

³ Katz (1996) wirft Chomsky vor, dass er die Endlichkeit externaler Zustände/Daten nur gegen die Endlichkeit von Gehirnzuständen eingetauscht hätte und somit seiner eigenen Argumentation nach die Kreativität von Sprache selbst nicht erklären könne. Gerade an dieser Stelle ist m. E. die multiple Realisierbarkeit und damit die Virtualität mentaler Zustände im Sinne der Computertheorie des Geistes von entscheidender Bedeutung. Auch Gehirnzustände sind unter dieser Perspektive externe Zustände, so dass der Vorwurf nicht stichhaltig ist.

Phänomenen oder durch die gänzliche Leugnung einer eigenständigen phonologischen Komponente. Aufgrund dieser Schnittstellenproblematik muss eindeutig definiert werden, wie die Relation von echtzeitlich produzierten und somit wahrnehmbaren Äußerungen (Tokens) und abstrakten Entitäten (Typen) aufzufassen ist. Die Exploration dieser Relation ist für die Fundierung einer Ontologie der Phonologie (vgl. Bromberger & Halle 2000) wesentlich und zur uneindeutigen Abgrenzung von rein phonetisch basierten Ansätzen des Phonologieerwerbs notwendig. Bromberger & Halle (1986) unterscheiden in Anlehnung an Chomsky strikt zwischen einer physikalistischen und einer mentalistischen Sichtweise von Phonologie. Erstere konzentriert sich unter folgenden Annahmen auf akustische und artikulatorische Typen (ebd.: 512):

- „a) There exists a finite (relatively small) number of articulatory types such that each (normal) utterance will turn out to be a token of one of them or a sequence of tokens of some of them.
- b) And similarly for acoustical types.
- c) Utterances which belong to the same articulatory type belong to the same acoustical type.
- d) This follows from a very simple one-to-one relationship between the articulatory and acoustical characteristics or features on the basis of which types are defined, and hence there exists a set of *phonetic* types definable indifferently in either articulatory terms or acoustical terms.
- e) The rules of phonology can be expressed in terms of these characteristics, and therefore phonology can be done using phonetic types and does not require an ontology or ideology (in the sense of Quine) that goes beyond that required to define *phonetic* types.”

Im Wesentlichen geht der physikalistische Ansatz demnach davon aus, dass phonologische und phonetische Typen identisch sind. Im Gegensatz dazu, steht im mentalistischen Ansatz die Frage im Vordergrund, welche Intentionen bzw. mentalen Zustände durch eine Äußerung realisiert werden. Das physische Ausagieren von Intentionen ist, um auf obige Fragestellungen Bezug zu nehmen, kein Spezifikum der Phonologie-Phonetik-Schnittstelle. Die Intention z.B., einen Ball zu werfen, ist natürlich multipel realisierbar. Und natürlich kann nicht jeder Typ einer motorischen Ausführung als Ausagieren eben dieser Intention gewertet werden (z.B. Naserümpfen etc.). Unmöglich aber ist es, zu versuchen, anhand von anatomischen und physikalischen Begriffen ein Set von notwendigen und hinreichenden Bedingungen zu definieren, was eine Intention des Ballwerfens sei. Es existiert einfach keine spezifizierbare Form der Armbewegung, deren Realisierungen ausschließlich als die Erfüllung der Intention des Ballwerfens gelten könnte. Die Intention des Ballwerfens muss demnach als eine mentale Einheit verstanden werden, auf die durch die variable physikalische Realisierung rückwirkend geschlossen werden kann, die sich aber einer Definition auf Basis einer

Sammlung ihrer spezifischen Erscheinungsvarianten entzieht (vgl. Bromberger & Halle 2000).

In Anlehnung an dieses Beispiel beschäftigt sich die artikulatorische Phonetik aus mentalistischer Sichtweise ausschließlich damit, welche Aktivitäten im Vokaltrakt für das Ausführen einer phonologischen Intention notwendig sind und welche Schallwellen bei dieser Ausführung entstehen. Zwischen phonologischen und phonetischen Typen herrscht also die Beziehung der Nicht-Identität.

Phonologische Typen sind so verstanden mitnichten metaphysischen Konstrukte, sondern im Langzeitgedächtnis gespeicherte Repräsentationen. Bromberger & Halle betonen hierbei, „that phonology is about concrete mental events and states that occur in real time, in real space, have causes, have effects, are finite in number [...]“ (Bromberger & Halle 2000: 21).

Pylyshyn (1984) hat in Bezug auf die Perzeption und eingehende Stimuli, unter anderem Experimente mit dem Necker Würfel (Julesz 1975), stark gemacht, dass man ohne die Annahme von Repräsentationen oder Symbolen überhaupt keine explanativen Generalisierungen für spezifische Wahrnehmungen und Verhalten vornehmen könnte:

„This is not do deny that some causal chain connects stimulation and perception. The point is simply that there exist regularities stateable over perceptual (or cognitive) categories that are not stateable over properties of the stimulation itself. This, in turn, is true because it appears that virtually no physical properties [...] are necessary and sufficient for the occurrence of certain perceptions; yet it is these perceptions that determine regularities in behavior. In fact, there could not be such properties in general since, as perceptionists are well aware, the way something is perceived can vary radically with physically identical stimuli and can be the same with physically very different stimuli.“ (Pylyshyn 1984: 15)

Ohne die Annahme einer symbolisch-lexikalischen und mit ihr eben einer phonologischen Repräsentation wäre so z.B. auch der *phonemic restoration effect* nicht erklärbar. In den Untersuchungen von Warren (1970) wurde das durch Hustgeräusch ersetzte [s] im Wort „legislature“ dennoch wahrgenommen, das Hustgeräusch allerdings von der Hälfte der Probanden sogar außerhalb des Wortes verortet. Diese heute schon klassische Untersuchung zeigt also, dass Wahrnehmung wesentlich durch die interne Verarbeitung kognitiver Systeme und deren symbolische Repräsentationen bestimmt ist und physikalische Stimuli vom wahrnehmenden Organismus nicht einfach deckungsgleich repliziert werden.

Die Annahme symbolischer Verarbeitung ist ein Kernelement der Computational Theory of Mind (vgl. Fodor 1975; Fodor & Pylyshyn 1988), deren wesentliche Annahmen hier nach Berent (2013: 38-39) zusammengefasst werden sollen:

„Computational Theory of Mind (CTM)⁴

a. Structured representations

- i. Mental representations are symbols, either simple (atomic) or complex (molecular).
- ii. Complex mental representations have syntactic form and semantic content.
- iii. The semantic content of complex mental representations depends on the syntactic structure of their parts and their atomic meanings.

b. Structure-sensitive processes

- i. Mental processes manipulate the syntactic form of representation in a manner that is blind to their semantic content.
- ii. Mental processes operate on variables.”

In Berent (ebd.) werden die Annahmen der CTM in eindeutigen Bezug zu phonologischen Repräsentationen und deren Verarbeitung gesetzt, so dass die relevanten Argumentationspunkte an dieser Stelle angebracht werden können.

Wenn Symbole als arbiträre Paarungen von abstrakter Form und semantischem Gehalt definiert werden, so ist eine solche Paarung für phonologische Repräsentation auf den ersten Blick nicht ersichtlich. Berent weist insofern darauf hin, dass der semantische Gehalt phonologischer Repräsentationen kein spezifisch konzeptueller ist, sondern in der differentiellen Interpretation des phonologischen Systems derjenigen Formen besteht, die letztlich ein Phonem repräsentational enkodieren (vgl. ebd.: 39). Da Berents Erläuterungen m.E. an dieser Stelle etwas unklar bleiben, soll hier eine Spezifikation erfolgen. Dass eine jede mentale Repräsentation einen propositionalen Gehalt hat, in diesem Sinne also semantisch als wahr oder falsch bewertbar ist, ist der Ausgangspunkt von Bromberger & Halle (1997). Sie haben entsprechend dieser Prämisse im Rahmen der Prädikatenlogik eine solche semantische Beschreibung phonologischer Symbole vorgenommen. Prädikate denotieren hier keine individuellen Objekte oder Ereignisse, sondern ob für letztere etwas der Fall ist oder nicht. Das Prädikat 'heiß' denotiert insofern kein Objekt, sondern ist wahr für jedes heiße Objekt und nicht wahr für jedes kalte Objekt. In Anwendung des Lambda-Kalküls kann so auch für einzelne phonologische Merkmale eine semantische Interpretation gegeben werden. Für das Merkmal [+rund] kann so das allgemeine Prädikat “ $\lambda x[+rund\ x]$ ” stehen. Wenn von einem α ausgesagt wird, dass es [+rund] ist, so wird dies in der Art “ $\lambda x[+rund\ x](\alpha)$ ” formalisiert, und wenn etwas das Merkmal [+rund] trägt, so kann dies als “ $(\exists y)(\lambda x[+rund\ x](y))$ ” semantisch transparent gemacht werden (vgl. ebd.: 96). Entsprechend können solche semantischen

⁴ Die CTM kann als spezifische Variante der Representational Theory of Mind (RTM) verstanden werden. In dieser werden zwei Hauptannahmen gemacht: Annahme 1: „For any organism O, and any attitude A toward the proposition P, there is a ('computational'/functional) relation R and a mental representation MP such that MP means that P, and O has A if O bears R to MP“ (Fodor 1987: 17). Annahme 2: „Mental processes are causal sequences of tokenings of mental representations.“ (ebd.).

Qualifizierung für jede signifikante phonologische Repräsentation (distinktive Merkmale, Pausen, Silbenkonstituenten, Phonemklassen usw.) gegeben werden.

Wenn phonologische Repräsentationen einen definierbaren propositionalen Gehalt haben, stellt sich die weitere Frage, was in der Phonologie unter atomaren und molekularen Repräsentationen zu verstehen sei. Zur Illustration führt Berent ein Beispiel aus dem Ägyptisch-Arabischen an. In diesem arabischen Dialekt ist der labiale stimmlose Plosiv /p/ nicht lizenziert, während der labiale stimmhafte Plosiv /b/ regelhaft produziert wird. Velare Plosive treten hingegen sowohl stimmhaft als auch stimmlos auf. Eine solche Differenz lässt sich phonetisch begründen, da das Merkmal [stimmlos] bei Velaren leichter erhalten bleibt als bei Labialen (vgl. Hayes & Steriade 2004). Schaut man sich in diesem Dialekt allerdings nun die Produktion von Geminaten an, so wird hier /bb/ realisiert, während die Geminat /pp/ nicht auftritt. Der entscheidende Punkt ist nun, dass nach rein phonetischen Kriterien die Geminat /pp/ leichter zu produzieren wäre als /bb/ oder auch /k/ (vgl. ebd.). Wir haben es hier also eindeutig mit einer Regularität zu tun, die nur aus dem Berechnungs- und Bewertungsverhalten der phonologischen Komponente stammen kann. Im Rahmen der CTM argumentiert Berent (2013: 40) nun, dass hier analog zu der logischen Folgerung „Sokrates ist sterblich“ aus den Prämissen „Sokrates ist ein Mensch“ und „Menschen sind sterblich“, eine sprachsystematische Folgerung stattfindet. Noch darüber hinaus tritt hier eine systematische Folgerung mit einer komplexen mentalen Repräsentation auf. In dem Satz „Plato & Socrates are barred from heaven“ z.B. werden [Plato und Sokrates] mediiert durch die syntaktische Form notwendig als semantisch komplexe Proposition verstanden. Auf das Nichtauftreten der Geminat /pp/ bezogen bedeutet dies, dass sie als eine molekulare Repräsentation die atomare Repräsentation /p/ enthält. D.h., da /p/ nicht lizenziert ist, wird die komplexe Struktur ebenfalls nicht lizenziert und dies zeigt nichts anderes, als dass auch die phonologische Berechnung sensitiv für formale bzw. molekulare Strukturen ist.

Im Zusammenhang mit der Produktion von Geminaten schlägt Berent eine algebraische Berechnung von Variablen vor. Da nicht-identische Kombinationen von Konsonanten vermieden werden sollen, kann in einer mathematischen Folgerung ($X \rightarrow XX$) mittels Variablen die Identität der Konsonanten sichergestellt werden. Wenn in X jeder labiale Konsonant inseriert werden kann, garantiert die Folgerung ($X \rightarrow XX$) den mehrmaligen Abruf des gleichen Konsonanten und somit die Identität. Formal lässt sich dies zusammenfassend einfach fassen: $*/p/ \rightarrow */pp/$ (vgl. ebd.: 38).

Nach Maßgabe obiger Ausführungen kann für die Untersuchung phonologischer Repräsentationen konstatiert werden, dass auch hier das Grundparadigma der Kognitionswissenschaften gelten kann, in dem zwischen einem abstrakten Programm, abstrakten Mechanismen der Symbolverarbeitung und ihrer physikalischen Implementierung und Ausführung unterschieden wird.⁵

Wir müssen am Ende dieser Ausführungen weiter festhalten, dass aus der hier dargestellten wissenschaftstheoretischen Perspektive die Frage danach, zu welchem Zeitpunkt oder über welchen Zeitraum phonologische Repräsentationen aus phonetischen Mustern emergieren, letztlich nicht korrekt ist. Vielmehr ist die entscheidende Frage, ab welchem Zeitpunkt die, als kognitives System verstandene, phonologische Komponente der Grammatik über einem universellen phonetischen Vokabular, den distinktiven Merkmalen, operiert. Von besonderer Relevanz zur Klärung dieser Fragestellung ist, in welcher Form sprachliche Komponenten kognitiv organisiert sind und wie sich Angeborenheit und Lernbarkeit innerhalb einer solchen Organisation zueinander verhalten.

1.2 Die allgemeine kognitive Organisation des Sprachsystems

1.2.1 Modulare Sprachverarbeitung

Modularität ist ein Terminus, der in den Kognitionswissenschaften eine frequente Verwendung findet, aber abhängig von den jeweiligen wissenschaftlichen und philosophischen Axiomen eine unüberschaubare Diversität an Interpretationsmöglichkeiten bietet. Man kann wissenschaftshistorisch die breite explizite Diskussion dieses Begriffes mit dem Erscheinen von Fodors *Modularity of Mind* (1983) bestimmen, auch wenn durch andere Arbeiten von z.B. Chomsky (1975b; 1981a) oder Marr (1982) bereits fundierte Konzepte von Modularität in den Kognitionswissenschaften Einzug gehalten hatten. Aus diesem Grunde soll hier Fodors Konzeption als Ausgangspunkt einer eingehenden Begriffsbestimmung und der Ausarbeitung eines für die vorliegende Arbeit zweckdienlichen Ansatzes fungieren.

⁵ Eben in einem solchen Sinne definiert an einer Stelle auch Chomsky (1981a: 189 u. 190) das methodische Vorgehen der Linguistik: „So gesehen ist die Linguistik eine abstrakte Untersuchung gewisser Mechanismen, ihres Wachstums und ihrer Reifung. Wir können den für die Anfangs-, Zwischen- und Endzustände angenommenen Strukturen in genau demselben Sinn Existenz zuschreiben, wie einem Programm, von dem wir glauben, es sei irgendwie in einem Computer realisiert; oder das wir annehmen, um die mentale Repräsentation eines dreidimensionalen Gegenstandes im Gesichtsfeld zu erklären. [...] Letztlich hoffen wir, Evidenz für physikalische Mechanismen zu finden, die das Programm realisieren; wir können vernünftigerweise erwarten, dass die in der abstrakten Analyse des Programms und seiner Funktionsweise erzielten Ergebnisse entscheidend dazu beitragen, dieses Ziel zu erreichen (und im Prinzip auch umgekehrt, nämlich dass die Informationen über die Mechanismen zu einem Verständnis des Programms beitragen.)“

Fodor entwickelt eine dreiteilige, funktional motivierte Taxonomie kognitiver Prozesse, in der zwischen Transduktoren, Input-Systemen und zentralen Systemen unterschieden wird.

Die Aufgabe von Transduktoren ist es, Informationen einer physikalischen Form in eine andere Form zu konvertieren. Im Falle des kognitiven Systems werden distale Reize in analog strukturierte neuronale Ereignisse konvertiert. Operationen der Transduktoren sind in diesem Sinne typenäquivalent zu physiologischen Zuständen (Pylyshyn 1984: 142). Gerade deshalb bedarf die Identifikation von Transduktoren genauerer Spezifikationen. Fodor & Pylyshyn (1981) weisen darauf hin, dass zum einen bei einer solch reinen Konvertierungsdefinition das gesamte Nervensystem als ein Transduktor begriffen, zum anderen Mechanismen ohne signifikanten funktionalen Bezug zu perzeptiven Leistungen ebenfalls zur Klasse der Transduktoren gerechnet werden könnten. Sie schließen daraus

„[...] that the identification of transducers will have to advert not, in the first instance, to their neurological structure but to their role in the cognitive processes that they subserve.“ (ebd.: 158)

Mit Uttal (1967) bezeichnen Fodor & Pylyshyn somit nur solche neuronale Ereignisse als Transduktoren, die eine funktionale Relevanz für das kognitive System besitzen.

Liefert der Output der so definierten Transduktoren dermaßen Repräsentationen proximaler Reizkonfigurationen, so besteht der Output der Input-Systeme aus symbolischen Repräsentationen distaler Objekte. Input-Systeme sind das, was im eigentlichen Sinne Module des kognitiven Systems genannt wird. Fodor bestimmt eine Liste von neun Merkmalen, die Module von holistisch operierenden zentralen Systemen unterscheidet und welche alle Input-Systeme teilen.

1. Domänenspezifität: Es wird nur ein solcher Input verarbeitet, für den das Input-System spezifiziert ist. So werden im Sprachverarbeitungssystem nur akustische Daten berechnet, die aus sprachlichen Äußerungen stammen. Isoliert man z.B. aus einer sprachlichen Äußerung eine Sequenz, die in einer solchen von Testpersonen als Einsatz eines Konsonanten wahrgenommen wird, so wird ein und derselbe physikalische Reiz als Pfeifton wahrgenommen (vgl. Liberman et al. 1967).
2. Informationelle Abgeschlossenheit: Andere Informationen wie z.B. aus Überzeugungsprozessen oder aus Hintergrundwissen haben keinen Einfluss auf die Verarbeitung der Input-Systeme. Die Muller-Lyre-Illusion illustriert dies eindrucksvoll. Entgegen eingehender Prüfung und besseren Wissens kommt man immer wieder zu dem „falschen“ visuellen Eindruck, dass die untere Gerade länger als die obere sei.
3. Zwangsläufiges Operieren: Der Input der Transduktoren wird automatisch und quasi „reflexhaft“ berechnet. Die Ausführung der Berechnung kann nicht willentlich beeinflusst werden, so dass eine Äußerung als Satz und ein visuelles Feld aus Objekten in einem dreidimensionalen Raum bestehend wahrgenommen werden muss.

4. Begrenzte Zugänglichkeit für zentrale Prozesse: Die Annahme ist, dass bei einer von den Transduktoren ausgehenden Bottom-up-Verarbeitung der Top-down-Zugriff der zentralen Prozesse immer mehr abnimmt. So kann bei einem kürzlich gehörten Satz meist nicht mehr die automatisch produzierte syntaktische Struktur erinnert werden, wohl aber die globale Bedeutung, so dass nur eine sinngemäße Wiedergabe möglich ist.
5. Schnelligkeit: Die Schnelligkeit der Verarbeitung unterscheidet Module von zentralen Prozessen, die für bestimmte Problemlösungen oder Interpretationen teilweise sehr lange Zeiträume in Anspruch nehmen. Musterbeispiel für die schnelle Verarbeitung von Input-Systemen sind Shadowing-Experimente (Marslen-Wilson 1973), in denen die Wiederholungszeit von gehörten Äußerungen zum Teil eine Viertelsekunde beträgt und sogar verstanden werden.
6. Seichter Output: Informationen werden über ein syntaktisches Kalkül berechnet, semantische Interpretationen, wie z.B. die Zuordnung von Wahrheitswerten, erfolgen erst durch die zentralen Systeme.
7. Implementierung in fester neuronaler Architektur: Vor allem das Merkmal der informationellen Abgeschlossenheit legt nahe, dass auf neuronaler Ebene fest verdrahtete Verbindungen zur Vereinfachung des Informationsflusses bestehen.
8. Charakteristische Störungsmuster: Selektive Ausfälle bei Schädigungen des Gehirns, wie z.B. Aphasien und Agnosien, zeigen, dass eine funktionale Distinktheit von Vermögen vorhanden ist. Die Nichtbeeinträchtigung von Gedächtnis-, Aufmerksamkeits- oder Problemlösungsleistungen unterminieren die mangelnde Erklärungskraft eines holistischen Ansatzes für die visuelle und sprachliche Verarbeitung.
9. Uniforme Ontogenese: Umweltbedingungen und Persönlichkeitsfaktoren scheinen auf die Funktionsfähigkeit der Input-Systeme keinen Einfluss zu haben. Bei Beeinträchtigungen liegen wohl vielmehr endogene Faktoren zugrunde. Dies lässt den Umkehrschluss zu, dass auch der relativ gleichförmige ungestörte Spracherwerb ebenfalls endogen, d.h. durch eine genetisch enkodierte Universalgrammatik, bedingt ist.

Was man mit Colthart (1999) betonen muss, ist, dass Fodor keine eindeutige Definition von Modularität anbietet und die Liste der Charakteristika nicht notwendig in Gänze erfüllt sein muss.⁶ Sind die meisten Charakteristika als optional anzusehen, so ist allerdings informationelle Abgeschlossenheit für Fodor die einzige Eigenschaft, die ein kognitives System notwendig besitzen muss, um als Modul zu gelten (Fodor 1983: 71; Fodor 2000: 63). Es ist gerade die Betonung dieser Systemeigenschaft, die nach Coltheart (1999) für einige Verwirrung in Bezug auf die Modularitätshypothese gesorgt hat. Denn Fodor verneint zum einen ausdrücklich, dass Module aus verschiedenen Subprozessen zusammengesetzt sein könnten, verweist allerdings zum anderen explizit auf den „phonemic restoration effect“, der einen klaren Top-down-Prozess innerhalb der Sprachverarbeitung darstellt (Fodor 1983: 65). In der Verneinung, dass Module zusammengesetzt sein können, hätten wir somit in diesem Falle zwei domänenspezifische Verarbeitungskomponenten (spezifischer phonetische vs.

⁶ „One interpretation might be this: given that a system has any of the properties in question, then the likelihood is considerable that it has all the rest... However, I doubt that a claim that strong could be empirically sustained, since it is reasonably easy to think of psychological processes that are fast but not encapsulated, or involuntary but not innate, and so forth. The present contention, in any event, is relatively modest: it's that if a psychological system has most of the modularity properties, then it is very likely to have all of them.“ (Fodor 1983: 137)

lexikalische Verarbeitung), die außerdem mehrere der o.g. Charakteristika erfüllen, welche aber nicht als Module bezeichnet werden dürften. Coltheart (1999) bestimmt deshalb Domänenspezifität als Hauptmerkmal für modulare Verarbeitung. Was überhaupt als eine modulare Verarbeitungskomponente gelten kann, sollte empirisch, z.B. durch Beobachtungen doppelter Dissoziationen, in der klinischen Forschung festgestellt werden. Ob und wie viele andere Merkmale das jeweilige kognitive Verarbeitungssystem zusätzlich besitzt, ist dann ebenfalls eine empirisch zu beantwortende Frage (vgl. ebd.: 119).

Den Schritt weg von einer Moduldefinition, die wie im Ansatz Fodors auf die Ebene ganzer Vermögen (sprachliche, visuelle, auditive Verarbeitung etc.) fokussiert, hin zu der Annahme einer komplexen Architektur aus Submodulen vollzieht auch Jackendoff (2000; 2002) in seiner Konzeption der Repräsentationalen bzw. Struktur-beschränkten Modularität. Nimmt man eine solche Dekomposition der Sprachverarbeitung vor, so gerät aber die Annahme informationeller Abgeschlossenheit in Widerspruch zu Erklärungen folgender Art (Jackendoff 2000: 12):

„The output of the syntactic processor is sent to the semantics.”
(1) syntactic processor → semantic processor”

Auf den Punkt gebracht, besteht das Problem in einer Divergenz zwischen domänenspezifischer und informationell abgeschlossener Verarbeitung. Domänenspezifität bedeutet ja nichts anderes, als dass nur ein spezifisches strukturelles Format von einem Modul verarbeitet werden kann. Dieses Format ist dann allerdings von einem anderen Modul nicht lesbar, da es ebenfalls nur mit einem spezifischen Format arbeitet. In diesem Sinne sind Module blind für die Informationen, die in anderen Komponenten verarbeitet werden.

Dass also Informationen solchermaßen nicht nur abgeschlossen wären, sondern auch im eigenen System eingeschlossen, d.h. für andere Systeme nicht interpretierbar, ist für Jackendoff eine der größten Lücken verschiedener Modularitätskonzeptionen. Damit bei Informationeller Abgeschlossenheit überhaupt Information von einem System in ein anderes übertragen werden kann, muss man somit sogenannte Schnittstellenmodule annehmen, die über Schnittstellenregeln operieren (Jackendoff 1997: 23-24):

„Components of an interface between system A and system B (an A-B interface)
a. A set of representations in system A to which the interface has access (the system B interface level(s) of system A or BILA)
b. A set of representations in system B to which the interface has access (the system A interface level(s) of system B or AILB)
c. A set of A-to-B correspondence rules, of the form illustrated in (1), which create the correlation between BILA and AILB.

- (1) General form of correspondence rules
 Configuration X in BILA
 Must/may/preferably does correspond to
 Configuration Y in AILB.“

Als Beispiel diene hier eine Formalisierung einer Schnittstellenregel für die Korrespondenz einer phonetischen Repräsentation mit einem entsprechenden motorisches Programm (ebd.):

„Configuration X in phonetic representation
 Must/may/preferably does correspond to
 Configuration Y in motor instructions.“

Die Schnittstellenmodule sind also selbst informationell abgeschlossen, qua definitionem aber bi-domänenspezifisch. Solche bi-domänenspezifischen Schnittstellenmodule stellen die Korrespondenzbeziehungen zwischen Informationen in genetisch angelegten Modulen innerhalb des Sprachsystems her und verknüpfen diese auch mit angeborenen Modulen außerhalb des Sprachsystems (z.B. mit visueller Wahrnehmung und konzeptueller Struktur). In neueren Veröffentlichungen ist es Jackendoff (2002: 131) so möglich, das Lexikon über die Schnittstellenmodule zu definieren:

„[...] the function of lexical items to serve as interface rules, and the lexicon as a whole is to be regarded as part of the interface components.“

Insofern etablieren lexikalische Einträge die Korrespondenz von syntaktischen Konstituenten mit phonologischen und konzeptuellen Strukturen. Jackendoffs Parallele Architektur des Sprachsystems mit genannten Schnittstellenkomponenten stellt sich in einem groben Umriss folgendermaßen dar, genauer soll sie allerdings in Kapitel 1.6 dargestellt werden:

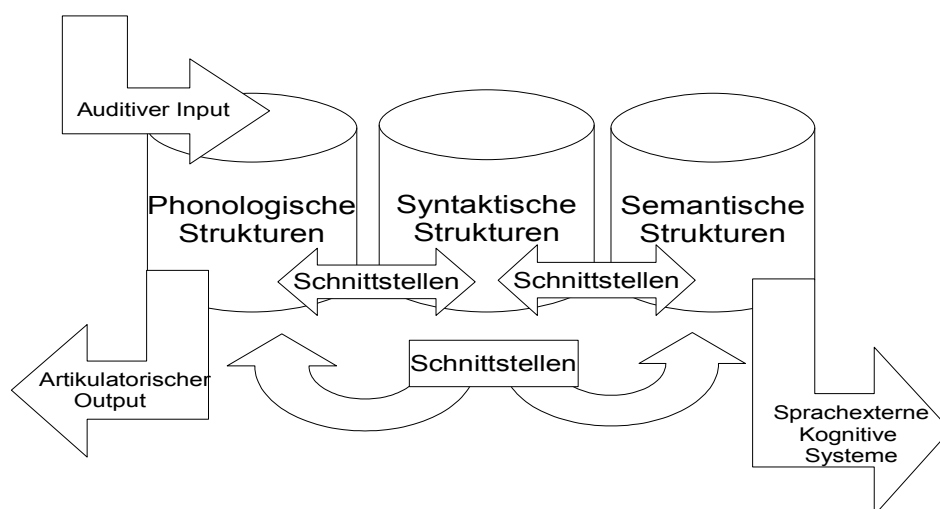


Abbildung 1: Grundriss der Parallelen Architektur (Jackendoff 2002).

Da die empirischen Untersuchungen zur Teilaktivierung und -schädigung der phonologischen Verarbeitung Legion sind (vgl. für einen Forschungsüberblick über die Messung früher und schneller Aktivierung phonologischer und prosodischer Diskriminationsprozesse anhand von Event-related Potentials (ERPs) z.B. Friederici (2005); für Störungen im Erwachsenenalter exemplarisch Caramazza & Berndt (1983) und für eine Übersicht Ellis & Young (1991); für Störungen im Erwerb exemplarisch der Fall P. in Dümig & Leuniger (2013) und für eine Übersicht Yavaş (1998)) und die theoretischen Argumente in Kapitel 1.1 für eine mentale Berechnung von phonologischen Repräsentationen mit neuronaler Implementierung dargelegt wurden, soll im weiteren Verlauf die phonologische Komponente als kognitives Modul mit einem weiterhin submodularen Aufbau verstanden werden.

Für die Modularitätshypothese als Ganze gilt, dass sie vielen Forschern vor allem durch zunehmende Erkenntnisse über Strukturmerkmale und Organisationsweisen des Gehirns als obsolet bis widerlegt gilt. Wenn kognitive Module angenommen werden, dann werden sie meist als Endprodukt bzw. spätere stabile Systemzustände der Gehirnentwicklung verstanden. Da mit Argumentationen auf der Beschreibungsebene neuronaler Entwicklungsvorgänge letztlich auch die Untersuchungsmöglichkeit früher phonologischer Repräsentationen und Berechnungen als Teil eines Phonologie-Moduls negiert werden, soll im nächsten Kapitel der speziell der Frage nach dem Stellenwert der Modularitätshypothese in Hinblick auf den frühen Spracherwerb nachgegangen werden.

1.2.2 Modularität im Spracherwerb

In der Diskussion darüber, ob modulare Verarbeitung und Repräsentationen angeboren sind, findet sich häufig der Verweis auf die Ausbildung neuronaler Verschaltungen innerhalb der ersten Lebensjahre. Unter dem Schlagwort der neuronalen Plastizität wird auf Fodors siebtes Kriterium einer festen neuronalen Implementierung modularer Verarbeitung abgehoben (Fodor 1983: 98-99), ein Kriterium, das er in späteren Publikationen nicht mehr verwendet (vgl. z.B. Fodor 1998; 2000). Vielmehr gesteht er in Bezug auf die Möglichkeit einer prädeterminierten diskreten Lokalisierung mentaler Funktionen im Gehirn explizit das Scheitern von Lokalisationstheorien in der Traditionslinie von Gall ein:

„The empiricists won the battle, of course; but my guess is they will lose the war.“ (Fodor 1999)

Die Frage, die sich allerdings stellt, ist, ob mit dem Verweis auf neuronale Plastizität zugleich auch die Möglichkeit modularer Verarbeitung in der frühen Kindheit ad absurdum geführt wurde oder ob es sich bei dieser „Widerlegung“ nicht vielmehr um einen der vielen Neuromythen (vgl. Schulte 2001) handelt, die gegenwärtig im wissenschaftlichen Diskurs immense Verbreitung gefunden haben.

Ausgangspunkt der Argumentation gegen modulare Verarbeitung im frühen Kindesalter sind verschiedene Untersuchungen zum reifenden Gehirn. Schon Freud (1897) hat darauf aufmerksam gemacht, dass spezifische Hirnschädigungen im Kindesalter nicht die gleichen Symptome evozieren, wie sie bei erwachsenen Aphasikern zu beobachten sind. In einer frühen Untersuchung von Basser (1962) an über einhundert Patienten mit frühen fokalen Hirnläsionen, worunter sich 30 Patienten mit Hemisphärektomien befanden, ließ sich für die Sprachverarbeitung kein Unterschied zwischen links- oder rechtsseitiger Läsionen feststellen. Nach einer kurzen Phase mit Produktionsschwierigkeiten verschwanden die Symptome nach einigen Tagen bis Wochen. Bates et al. (1997) fanden ebenfalls, dass die meisten Kinder mit fokalen unilateralen Hirnläsionen einen Spracherwerb im Normbereich durchliefen.

Dass die frühe Sprachverarbeitung bi-hemisphärisch distribuiert stattfindet und sich linkshemisphärische Dominanz in der Sprachverarbeitung erst sukzessive ausbildet, konnten z.B. Neville & Bavelier (2000) mittels Elektroenzephalographie (EEG), zeigen. In einer Studie mit sprachunauffälligen Kindern im Alter von 20 Monaten, 28-30 Monaten und 36-42 Monaten wurde die elektrische Aktivität der Verarbeitung von Wörtern der offenen (Nomen, Adjektive, Verben) und geschlossenen Klasse (Artikel, Präpositionen, Pronomen, Konjunktionen) untersucht. Mit Hilfe der aus dem EEG gewonnenen ereignis-korrelierten Potentiale (EKP) konnte dargestellt werden, dass im Alter von 20 Monaten noch eine bi-hemisphärische Verarbeitung beider Wortklassen stattfindet, mit 28-30 Monaten Funktionswörter tendenziell mehr linkshemisphärisch verarbeitet werden und mit 36-42 die Dominanz der frontotemporalen Verarbeitung in der linken Hemisphäre für Funktionswörtern wie bei Erwachsenen ausgebildet ist.

Die Plastizität in der funktionalen Ausrichtung der Hirnhemisphären hat sein Pendant in Selektionsprozessen auf neuronaler Ebene. Edelman (1992) unterscheidet hier drei Phasen:

1. Entwicklungsselektion: Diese Phase ist stark durch genetische Prozesse beschränkt, da sich die grundlegende Neuroanatomie einer Art ausbildet. Von den Neuronen verzweigen etliche Axone in alle Richtungen und bilden ein sog. Primäres Repertoire an neuronalen Schaltkreisen. Dieses Repertoire zeigt allerdings eine enorme interindividuelle Variation, da verschiedene molekulare Prozesse, stochastische

Schwankungen der Zellbewegung, Zelltod und von Aktivität abhängige Koppelungen von neuronalen Verbindungen wirksam sind.

2. Erfahrungsselektion: In dieser Phase verändert sich die allgemeine Anatomie nicht mehr. Durch Verhalten werden hier aber präformierte Verschaltungen selektiv verstärkt oder geschwächt. Das so gebildete neuroanatomische Muster bildet das Sekundäre Repertoire.
3. Reentry: Primäre und sekundäre Repertoires bilden Karten, d.h. Funktionsverbände (z.B. visuelle Karte V3 für Farbsehen), die mannigfaltig reziprok verschaltet sind.

Für die Sprachverarbeitung kann folgendes Entwicklungsmuster festgehalten werden (nach Huttenlocher 1994; Huttenlocher & Dabholkar 1997). Die synaptische Dichte erreicht in den sprachrelevanten kortikalen Gebieten (linker Heschl Gyrus, Broca- und Wernicke-Areal) im Alter von 12 Monaten ein Maximum, wobei die Elimination der Synapsen in der späten Kindheit (ca. 3-5 Jahre) beginnt und in der mittleren Adoleszenz beendet ist. Zieht man die obigen Ausführungen zur Lateralisierung hinzu, so scheint also eine Koinzidenz von synaptischer Dichte bzw. Verschaltungsoptionen und funktionaler Plastizität zu bestehen.

Es sind Ergebnisse wie die oben dargestellten, die eine Reihe von Autoren zum Schluss kommen lässt, dass eine computerfunktionalistische Perspektive und mit dieser Modularität entweder gänzlich widerlegt sei (z.B. Edelman 1992; Szagun 2006) oder modulare Verarbeitung (anders als bei Fodor definiert) sich zumindest in der Kindheit durch einen Prozess der Modularisierung erst herausbilden müsse (z.B. Karmiloff-Smith 1992).

Wie Fodor betont, steht hinter dieser Sichtweise eine relativ naive Grundannahme:

„[...] an unwary reader might suppose it's pretty clear, in general, how knowledge is encoded in the nervous system; so if we don't find the right neural connectivity when we look at the infant's brain, that shows that there's no innate content there.“ (Fodor 1998: 145)

Das Problem hierbei ist also, dass niemand weiß, in welchen spezifischen zerebralen Konnektivitätsmustern spezifische mentale Zustände enkodiert sind. Poeppel & Embick (2005) besprechen genau dieses Problem anhand der Neurolinguistik. Wie soll man sich eine gegenseitige Abbildbarkeit von Beschreibungsvokabularen oder gar eine Identität derselben vorstellen? Entsprechen Repräsentationen distinktiver Merkmale Dendriten, Silben Neuronen, Morpheme Zellensembles, Nominalphrasen Zellpopulationen und Sätze ganzen kortikalen Kolumnen? Sie kommen zu einem ernüchternden Schluss:

„[...] to our knowledge, there is not a single case of a successful reduction in these terms in the domain of language [...]“ (ebd.: 106).

Sie fassen dieses Problem, dass nämlich das Beschreibungsvokabular der Linguistik eben nicht auf dasjenige der Neurowissenschaften abbildbar ist (ein Problem, das allerdings grundsätzlich auf das Verhältnis von Kognitionswissenschaft und Neurobiologie zutrifft), als Ontological Incommensurability Problem (OIP):

„Ontological Incommensurability Problem (OIP):

The units of linguistic computation and the units of neurological computation are incommensurable.“ (ebd.: 106)

Das OIP erweist sich nun allerdings ausschließlich als Stolperstein für die Autoren, die ihre Argumente im Wesentlichen auf Grundlage der Kommensurabilität von Gehirnvorgängen und Sprachverarbeitungsprozessen konstruieren. Sie halten die Proposition P für notwendig wahr, dass für jedes mentale Prädikat einer erklärenden psychologischen Beschreibung mindestens eine logische Beziehung zu einem beobachtbaren Verhalten (Token) bestehen muss (vgl. Fodor 1968: 51). Das OIP trifft mentalistische Ansätze nicht, da hier explizit bestritten wird, dass P notwendig wahr ist. Gegen modulare Verarbeitung kann also nicht durch Rekurs auf die Hirnreifung argumentiert werden. Aber selbst wenn man die Reifungsprozesse im Gehirn genauer beleuchtet, scheinen viele der Argumente, wie Carruthers (2006) zeigt, nicht stichhaltig zu sein. Wie oben dargestellt, finden im Alter von drei bis vier Jahren und in der Adoleszenz Selektionsprozesse aufgrund von funktionalen Nutzungshäufigkeiten statt, bei denen 40 Prozent der Neurone und synaptischen Verbindungen absterben, wobei das gesamte Neuronenrepertoire schon in den sieben ersten pränatalen Entwicklungsmonaten gebildet wurde (vgl. Webb et al. 2001). Carruthers (2006: 161) hebt nun hervor, dass diese initiale Platzierung neuronaler Verbindungen keineswegs zufällig ist. Sie wandern vielmehr zu späteren Positionen im Gehirn und senden Axone nach Maßgabe eines komplexen physikalisch-chemischen Signalsystems zu spezifischen Regionen (vgl. Thoenen 1995). Es ist also keineswegs der Fall, dass zu Beginn der Entwicklung ein gleichförmiges Netz neuronaler Verbindungen gegeben wäre, welches anschließend durch Erfahrung modifiziert wird. Vielmehr scheint die Rolle von Erfahrung zu sein, inhärent präformierte funktionale Systemgrenzen zu präzisieren, indem Neuronen mit gleicher Funktion stärker verknüpft werden (vgl. Carruthers 2006: 161). Selbst Selektionsvorgänge sind stärker von endogenen Mechanismen determiniert, als man früher angenommen hat. So konnte z.B. gezeigt werden, dass der Zusammenschluss von Neuronen zu basalen okularen Dominanzkolumnen in einer ersten Entwicklungsphase vorgeburtlich durch spontane neurale Aktivität der beiden Retinas gelenkt wird. Somit ist die basale Struktur kortikaler Regelkreise schon vor jeglicher visueller

Erfahrung festgelegt (vgl. Katz et al. 2000). Selbst wenn man also modulare Verarbeitung von Repräsentationen durch neuronale Plastizität widerlegen wollte, sprechen einige Daten dagegen, dass letztere so geartet ist, dass keine basalen neuronalen Strukturen für erstere vorhanden wären.

Eine relativ moderate Schlussfolgerung, die man aus den obigen Ausführungen ziehen kann, ist, dass man einen reziproken Abgleich zwischen kognitiven Modellvorstellungen und neuronalen Befunden vornehmen sollte. In diesem Sinne unterscheidet Hagoort (2008) zwischen einer *upward adequacy* (neuronale Prinzipien müssen adäquat sein in Hinblick auf Verhaltensdaten und der von diesen Daten abgeleiteten kognitiven Architektur) und einer *downward adequacy* (die kognitive Architektur muss in Hinblick auf neurobiologische Erkenntnisse adäquat sein). In einem solch balancierten Abgleich können die jeweiligen Modellvorstellungen unter Zugrundelegung des OIP selbstverständlich nur approximativ sein. Einseitige Rückschlüsse und Verengungen wie z.B. diejenigen von Edelman (1992), der mentale Leistungen auf spezifische Selektionsprozesse auf neuronaler Ebene reduziert, werden durch diese Adäquatheitskriterien allerdings vermieden. Allerdings sind durch die Proliferation der Neurowissenschaften reduktionistische Ansätze in der gegenwärtigen Forschungslandschaft durchaus Usus, weshalb es noch dringlicher gilt, sich konzeptionell von allen Spielarten des Reduktionismus abzugrenzen. Ein Meilenstein der Argumentation gegen diese bildet Fodor & Pylyshyn (1988), deren grundlegende Argumentation im nächsten Kapitel angeführt werden sollen.

1.2.3 Die Irrelevanz konnektionistischer Modelle

Bisher war die Darstellung vor allem auf die konkrete neuronale Implementierung von Sprache bezogen. Da in der Psycholinguistik konnektionistische Sprachmodelle eine weite Verbreitung gefunden haben (vgl. z.B. Rumelhart & McClelland 1986, Schade 1999, Smolensky 1999), in der Erforschung der Sprachproduktion bis heute verwendet werden (z.B. Chang 2002, Dell & Sullivan 2004, Levelt et al. 1999) und ihre Passung zu aktuellen theoretischen Modellen der Sprachwissenschaft, wie z.B. der Optimalitätstheorie (vgl. z.B. Prince & Smolensky 1997), betont wird, gilt es an dieser Stelle den Fokus noch etwas klarer auf die postulierten Verarbeitungsmechanismen zu lenken. Auch wenn es keine direkte Abbildung von konnektionistischen Modellen auf das reale Gehirn geben kann, da letzteres sich in seiner biologischen Selbstorganisation grundsätzlich von konnektionistischen Computermodellen unterscheidet (vgl. die Kritik von Edelman (1992)), so sind doch bezogen

auf die Erklärbarkeit mentaler Operationen die Argumentationsfiguren von Konnektionisten und Neurowissenschaftlern relativ deckungsgleich. In diesem Sinne impliziert alle Kritik am Konnektionismus auch solche Erklärungsmodelle, die nach Maßgabe biologisch fundierter neuronaler Konnektivität und Aktivität mentale Prozesse darstellbar bis erklärbar machen wollen.

Die wesentliche Grundvoraussetzung, die konnektionistische Modelle machen, ist die, dass die Operationen innerhalb eines konnektionistischen Netzwerks hinreichend sind für die Beschreibung mentaler Prozesse. Innerhalb dieses Netzwerkes sind Knoten über Assoziationen mit anderen Knoten verbunden und alle Bestandteile können, je nach Art des Modells, mit stochastisch beschreibbaren Gewichtungen, d.h. frequenzgetriebenen Bereitschaften zur Aktivierung bei einem jeweils gegebenen Input ausgestattet sein. Wie unterschiedlich die Modellierung auch immer sei, die Grundfigur einer kausalen Aktivierungsweitergabe von einem Knoten zum anderen ist allen gemein. In dieser raumzeitlichen, kausalen Aktivierung kann ein aktivierter Knoten dem anderen folgen bzw. können mehrere Knoten gleichzeitig und/oder parallel durch einen Input aktiv werden.

Zwei wesentliche Unterschiede werden nun von Fodor & Pylyshyn (1988) zwischen klassischen Theorien und konnektionistischen Theorien ausgemacht. Klassische Theorien nehmen, aufbauend auf dem Modell der Turing-Maschine, eine kombinatorische Syntax und Semantik wie auch eine Struktursensitivität von Prozessen an. Unterschieden wird in klassischen Theorien weiter zwischen strukturell atomaren und strukturell molekularen Repräsentationen. Letztere bestehen aus syntaktischen Konstituenten, die selbst wiederum atomar oder molekular sein können. Der semantische Gehalt einer komplexen Repräsentation wird entsprechend gemäß dem Kompositionalitätsprinzip (vgl. Frege 1923) berechnet. Struktursensitiv sind klassische Modelle in dem Sinne, als dass die zuvor beschriebene kombinatorische Struktur von Repräsentationen es möglich macht, Operationen auf der Basis von strukturellen Beschreibungen durchzuführen. Der entscheidende Punkt ist nun, dass in konnektionistischen Modellen keine Kombinatorik vorhanden ist und es dementsprechend keine Struktursensitivität geben kann. Eine Liste von Repräsentationen ist völlig unbrauchbar zur Erklärung sprachsystematischer Prozesse. Ein zuvor aktivierter Knoten ist nur kausal qua Aktivierungsweitergabe mit dem nächsten Knoten assoziiert, strukturelle Relationen wie Implikation, Interferenz, Konstituenz usw. können in einem solchen mono-operationalen Modell nicht hergestellt werden. So kann z.B. das Verstehen einfacher Sätze wie „John loves Mary“ mit einer sukzessiven Aktivierung von Knoten, die für (John), (loves) und (Mary)

stehen, nicht erklärt werden. Die jeweilige Einzelaktivierung ist blind für die ihr folgende. Auch in neueren Arbeiten wird versucht, logische wie auch rein sprachliche Operationen über Aktivierungen von Knoten, Neuronen oder Neuronenverbänden zu erklären:

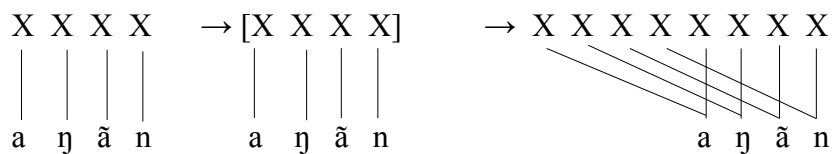
„Assume that there are two neural elements each of which becomes active when the individual sees one specific object, called A and B. A third neuron receives input from both and can be activated by either input. Now, activity of the third neuron signals that A or B is present. If the third neuron has a high threshold for becoming active, it may need both inputs to be active, in which case its activation would index “A and B”, that both objects are present. In this sense, a neuron can be said to mechanistically realise, or “calculate”, logical operations, “or”, “and”, “not” and, with slightly more elaborate wiring, “either-or”, identity and temporal statements such as “always” or “once”. The examples illustrate that our brain is equipped with built-in mechanisms relevant for logical operations.“ (Pulvermüller 2012: 448)

Wie am Beispiel deutlich werden sollte, wird auch hier der entscheidende Punkt vollkommen verfehlt. Die kausale Aktivierung eines Neurons durch zwei andere Neuronen ist nicht gleichzusetzen mit der logischen Operation der Konjunktion. Der Index „A und B“ im obigen Beispiel setzt voraus, dass die Objekte A und B in der Aktivierung des dritten Neurons selbst repräsentiert blieben, was nur auf Basis einer gespeicherten symbolischen Repräsentation und nicht auf derjenigen einer indifferenten Aktivierungsenergie geschehen könnte. Und selbst wenn man eine solch mechanische Realisierung zugestehen würde, müsste es eine etablierte Assoziation von A und B mit dem dritten, neutralen Neuron existieren, die individuell gelernt werden müsste. Damit hätten wir keine allgemeine Regel mehr, sondern müssten für alle Kombinationen der Objekte dieser Welt die Regel der Konjunktion lernen. Und dies würde letztlich auch für alle anderen logischen Operationen in der gleichen Art und Weise geschehen. Die Absurdität eines derartig lebenslangen Lernprozesses dürfte unmittelbar einleuchtend sein.⁷

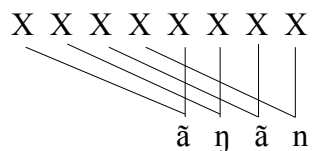
Für das Verstehen des Satzes wird letztlich eine Symbolverarbeitung und mit ihr eine kombinatorische Syntax mit den jeweiligen Konstituenten benötigt. Insofern ist der Schluss von Fodor & Pylyshyn, dass konnektionistische Modelle nur eine Rolle für die

⁷ Wie Fodor (1999: 68-69) betont scheint diese Erkenntnis der Absurdität derartiger neurowissenschaftlicher Modellierungen im Bildungssystem, insbesondere Universitäten und etablierten Medien noch nicht erlangt worden zu sein. Die immense finanzielle Förderung solcher Modellierungen erscheint insofern gleichfalls absurd: „Science is expensive, and it’s largely publicly funded, and there’s never enough money to do all the research that might be worth doing. In particular, brain imaging is expensive compared to other ways of trying to find out about the mind. If you put your money (which is to say: our money) into the elaborate technology required to establish neural localisations of mental functions by imaging techniques, you almost certainly take it out of other kinds of psychological research. Likewise in respect of the time and money that is required to train people to do the science; graduate students, too, are a limited resource. So I’m increasingly concerned when I find that yet another Tuesday has come around, and the *Times* is still interested in the neural localisation of mental functions and I still can’t figure out why. It occurs to me that maybe we’re heavily invested in finding answers to which we don’t know the corresponding questions.“

Implementationsebene eines kognitiven Systems, nicht aber für dessen Operationen auf der kognitiven Ebene spielen können, notwendig gültig. Auch wenn Fodor & Pylyshyn kein phonologisches Submodul erwähnen, so lässt sich doch ableiten, dass eine regel- und (nichtlineare) strukturbasierte phonologische Komponente ebenfalls nicht mit einem konnektionistischen Netzwerk zu erfassen ist. Ein Beispiel von Frampton (2009: 8-9) soll dies in aller Kürze verdeutlichen. Im Malaiischen ist die reduplizierte Form von *aŋã* (Träumerei) die Form *ãŋããŋã* (Ambition). Im Malaiischen existiert eine Nasalisierungsregel, bei deren Anwendung Vokale nach Nasalen nasalisiert werden. Worauf schon Onn (1976) hingewiesen hat, ist, dass es keine Erklärung dafür gibt, warum die Nasalisierungsregel am ersten Vokal appliziert wird. Nach Frampton kann man dieses Form der Reduplikation nur erklären, wenn man repräsentationale Zwischenebenen annimmt. In seinem Modell, der Distributed Reduplication, die in Kapitel 1.7.2 genauer erläutert werden soll, werden Elemente der X-Ebene, welche die zeitliche Abfolge encodieren, durch Einfügung von Klammerung auf einer nächsten Ebene kopiert. Bevor nun die Struktur repariert wird, indem jedem X-Slot nur ein Phonem zugeordnet wird, wird nun die Nasalisierungsregel wirksam.



Bevor nun die Struktur repariert wird, indem jedem X-Slot nur ein Phonem zugeordnet wird, wird nun die Nasalisierungsregel wirksam. Auf dieser abstrakten Zwischenebene befindet sich qua Assoziationslinien der letzte Nasal vor dem initialen Vokal.



Nur mit der Annahme solcher repräsentationaler Zwischenebenen, was kognitiv nur mit einem Modell möglich ist, das die Trennung von Programm und Datenspeicher im Sinne einer Turing-Maschine beinhaltet, sind solche diskreten und regelgeleiteten Verarbeitungsschritte möglich. Konnektionistische Modelle haben keine symbolisch basierte Speichermöglichkeit. Symbolisch encodierte Strukturen können insofern nicht mittels Regeln auf andere Strukturen

abgebildet werden, da hier nur von einem Punkt einer Repräsentationslistung zum nächsten gesprungen werden kann. Die Erklärung solcher regelgeleiteten Formen kann also in einem konnektionistischen Modell nicht geleistet werden und muss durch arbiträre ad hoc Assoziationen ersetzt werden. Dadurch erscheint zum einen der Spracherwerb als eine völlig unplausible, nie endende Adjustierungsprozedur durch einen Input, der enormen Frequenzschwankungen unterliegt, zum anderen ist für die wissenschaftliche Erklärung kognitiver Prozesse im Sinne von allgemeingültigen Explanationen nichts gewonnen.

Auch für die Beschreibung des phonologischen Teilsystems der Sprache ist demnach eine konnektionistische Modellierung abzulehnen und eine klassische Architektur vorzuziehen.

Im Folgenden sollen weitere Argumente für ein mit Symbolen operierendes kognitives System der Phonologie herausgearbeitet werden.

1.3 Die Struktur phonologischer Repräsentationen

1.3.1 Phonologie als sprachliches Submodul mit spezifisch repräsentationaler Verarbeitung

Bromberger & Halle (1989) sehen eine grundsätzliche Differenz zwischen phonologischen und syntaktischen Repräsentationsformen. Phonologie beschäftigt sich mit der Beziehung zwischen Oberflächenrepräsentationen, die als Input für die artikulatorischen und auditiven Systeme dienen, und abstrakten, im Gedächtnis gespeicherten, zugrundeliegenden Repräsentationen. Das Berechnungsvokabular ist auf beiden Repräsentationsebenen gleich, da es sich beide Male um phonetische Information bzw. phonetische Merkmale handelt. Die Beziehung zwischen den Ebenen ist definiert als Derivation auf Basis von extrinsisch geordneten Regeln, da die zugrundeliegenden Repräsentationen den bei einer aktuellen Äußerung aktivierten Oberflächenrepräsentationen klar vorausgehen müssen.

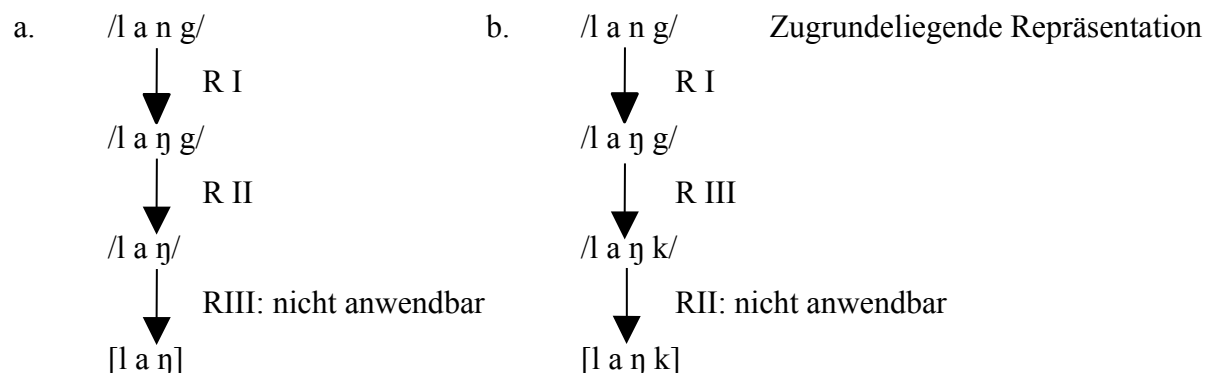
Eine phonologische Regel kann im Rahmen eines Derivationsprozesses einen Anwendungsbereich für eine zweite Regel schaffen oder vergrößern (feeding order) bzw. umgekehrt den Input für eine zweite Regel verkleinern (bleeding order). Ein Beispiel aus dem Deutschen ist die Abfolge von Nasalassimilation (R I), g-Tilgung (R II) und Auslautverhärtung (R III) (nach Ramers 2001)⁸.

⁸ Die hier angegebene Form linearer Regeln schließt eine Regelfolge nicht-linearer Regeln nicht aus (vgl. Wiese 1996, Kap. 7)

$$\begin{array}{l}
 \text{R I: } \begin{pmatrix} +\text{nas} \\ -\text{lab} \end{pmatrix} \rightarrow [+ \text{hint}] / \text{ ______ } \begin{pmatrix} -\text{son} \\ -\text{kont} \\ +\text{hint} \end{pmatrix} \\
 \text{R II: } \begin{pmatrix} -\text{son} \\ -\text{kont} \\ +\text{hint} \\ +\text{sth} \end{pmatrix} \rightarrow \emptyset / \begin{pmatrix} +\text{nas} \\ +\text{hint} \end{pmatrix} \text{ ______ }
 \end{array}$$

$$\text{R III: } [-\text{son}] \rightarrow [+ \text{gesp}] / \text{ ______ }]_{\sigma}$$

In Beispiel a. (s.u.) aus dem Standarddeutschen sieht man, wie R I und R II in einer feeding order stehen, R II und R III aber in einer bleeding order, da die Auslautverhärtung nach Anwendung von R II nicht greifen kann. In nächsten Beispiel b. zeigt sich hingegen eine andere Regelfolge in einer dialektalen Variante des Norddeutschen, in der nun eine bleeding order von R III und R II zu sehen ist, da durch die Anwendung von R III keine g-Tilgung mehr stattfinden kann (ebd.: 72-73):



Das Prinzip, welches eine spezifische Anordnung von Regeln sicherstellt, ist folgendes:

„Phonological rules are ordered with respect to one another. A phonological rule R does not apply necessarily to the underlying representation: rather, R applies to the derived representation that results from the application of each applicable rule preceding R in the order of the rules.“ (Bromberger & Halle 1989: 58-59)

Der syntaktischen Komponente fehlt diese extrinsische Ordnung von Regeln. In der Government and Binding-Theorie (abgekürzt GB-Theorie; Chomsky 1981b) werden die Beschränkungen durch Prinzipien bestimmt, welche von S-Struktur, D-Struktur und LF erfüllt werden müssen, um eine grammatisch wohlgeformte Struktur zu erzeugen. Über Move-a wird Information zwischen den Repräsentationsebenen geteilt, wobei diese Transformationsoperation die vorhandenen Beschränkungen wie das Theta-Kriterium oder das Empty

Category Principle nicht verletzen darf. Die Anwendung von solchen Prinzipien kann aber nicht auf eine strikt sequentielle Anwendung von Regeln zurückgeführt werden, so dass man nicht davon sprechen kann, dass eine Repräsentation von einer anderen durch solche Regelanwendungen deriviert wird, also eine Repräsentation der anderen notwendig vorausgehen muss.

Dass das Phonologie-Modul zudem mit einem anderen Berechnungsvokabular arbeitet als das Syntax-Modul, die jeweiligen Repräsentation insofern nicht direkt aufeinander abbildbar sind, soll auch an folgendem Beispiel verdeutlicht werden (vgl. Jackendoff 2007a: 5):

Syntax:

[Sesame Street] [is [a production [of [the Children's Television Workshop]]]]

Phonology:

[Sesame Street is a production of] [the Children's Television Workshop]

or

[Sesame Street] [is a production] [of the Children's Television Workshop]

Die syntaktische Struktur besteht aus einer Nominalphrase (NP) *Sesame Street*, einer Verbalphrase (VP) mit *is* als Kopf und einer NP, in die die weitere NP *the Children's Television Workshop* eingebettet ist. Die Intonationskontur kann nun auf verschiedene Weise realisiert werden und es muss keine Passung mit den vorhandenen syntaktischen Konstituenten gegeben sein. Die erste Variante übergeht klar die syntaktischen Grenzen der VP und NP in *Sesame Street is a production of*, so dass keine Zuordnung von Intonationsphrasen und syntaktischen Konstituenten stattfindet. Auch wenn keine Passung vorhanden ist, so gibt es doch Beschränkungen in Form von Korrelationen, die in den Schnittstellenmodulen definiert sein müssen. Die Produktion folgender Intonationsgrenzen ist z.B. ausgeschlossen:

*[Sesame] [Street is a] [production of the Children's][Television Workshop]

Die definierte Korrelation könnte nach Jackendoff (ebd.) in Form einer Schnittstellenregel wie folgt aussehen (vgl. auch Gee & Grosjean 1983; Hirst 1993; Truckenbrodt 1999):

a. Prosodic well-formedness: [Utterance IP IP ... IP]

b. Syntax–intonation correspondence

Syntax: [XP W1 ... Wi ... Wn]

Prosody: [IP W1 ... Wi]

where W1, ... Wn are the words in syntactic constituent

XP (and i may equal n)

Regel a. besagt hier, dass eine Äußerung aus einer Reihung von nicht eingebetteten Intonationsphrasen (IP) besteht. Nach Regel b. beginnt eine IP am Anfang einer syntaktischen Konstituente, kann aber noch innerhalb derselben enden. Die IP überschreitet aber auch nicht das Ende der längsten Konstituente mit der sie startet. Im obigen ungrammatischen Beispiel findet eine Verletzung der Korrelation statt, weil die zweite IP mit dem Nomen *Street* beginnt, das nicht am Anfang einer syntaktischen Konstituente steht. Jackendoff vertritt also ebenfalls klar die Haltung, dass das phonologische System eine eigenständige Komponente darstellt:

„[...] phonological structure requires its own set of basic units and combinatorial principles [...]. In other words, phonology is generative in the same sense that syntax is. Phonological structure may lack the high-powered recursive embedding of syntax, but it is hierarchical and generative nonetheless.“ (ebd.: 5).⁹

In diesem Kapitel wurde anhand von Regelanwendungen auf umschriebene Merkmalsmengen und prosodischen Einheiten wie Intonationsphrasen, die grundlegende Differenz von phonologischer und syntaktischer Verarbeitung herausgearbeitet. Unterscheiden sich nicht aber die hier angeführten Beispiele der Derivationen mittels phonologischer Merkmale und prosodische Regelanwendungen sowie auch die berechnungsrelevanten repräsentationalen Einheiten ebenfalls in markanter Art und Weise? Bei einer relevanten Differenz ergäbe sich die Schlussfolgerung, dass das Phonologie-Modul weiter dekomponierbar wäre.

Die Identifikation solch möglicher Subkomponenten innerhalb des Phonologie-Moduls durch die Bestimmung wissenschaftlich signifikanter Unterschiede soll im nächsten Kapitel geleistet werden.

⁹ Die Anschauung, dass die phonologische Komponente ihre Einheiten anders berechnet als die syntaktische, ist vor allem von van der Hulst (z.B. 2000; 2003; 2005) kritisiert worden. Er übernimmt zwar die parallele Organisation der Komponenten von Jackendoff, sieht aber den Unterschied nur in der Verwendung unterschiedlicher sprachlicher Einheiten und nicht in der Anwendung verschiedener kombinatorischer Prinzipien. Er spricht in diesem Sinne auch von einer phonologischen Syntax.

1.3.2 Prosodische Phonologie vs. Segmentale Phonologie

Wie McMahon (2007) ausführlich darlegt, besteht die Frage, ob Phonologie ein in sich homogenes Untersuchungsgebiet ist. Sie postuliert zwei völlig differente Forschungsfelder, deren Unterschiede in der Subsumtion unter den Terminus Phonologie selten klar gesehen werden. Auf der einen Seite ist nach McMahon die prosodische, auf der anderen Seite die segmentale Phonologie. Eine solch starke These der Unabhängigkeit beider Felder wird von McMahon unter Rekurs auf Forschungsergebnisse aus der Spracherwerbsforschung, der Sprachpathologie, der Primatenforschung, und der Neurologie vertreten.

Dass prosodische Merkmale einer Zielsprache von Kindern (rezeptiv) früher als segmentale Merkmale erworben werden, ist ihrer Meinung nach ein Zeichen dafür, dass Prosodie eine substantiellere angeborene Komponente darstellt (ebd.: 165)¹⁰. Im phonetischen Signal ist prosodische Information unterdeterminiert, d.h. z.B., dass Silbengrenzen in diesem nicht identifizierbar sind. Dennoch gelingt die Platzierung von Betonung nur durch die adäquate Silbifizierung eben dieses Signals (vgl. Dresher & Kaye 1990). Alles was das Kind hingegen für den Erwerb des Vokal- und Konsonantensystems einer jeweiligen Zielsprache benötigt, ist ihrer Meinung nach im sprachspezifischen phonetischen Signal enthalten.

Zudem ist in vielen Studien gezeigt worden, dass es im gestörten Phonologieerwerb zu Dissoziationen der beiden phonologischen Systeme kommen kann, d.h. prosodische Defizite bei relativ intakter segmentaler Verarbeitung (z.B. Harris et al. 1999) und vice versa (z.B. Snow 2001; Wells & Peppé 2003). Detaillierte diagnostische Tests sind derart konzipiert, dass sie solche Dissoziationen erfassen (z.B. Bernhardt & Stemberger 2000).

Mit anderen Primaten teilt der Mensch bestimmte Eigenschaften prosodischer Markierung. Phonetisch messbar gibt es während einer Äußerung eine stetige Abnahme der Grundfrequenz, um schließlich in einen raschen Abfall am Ende einer Äußerung zu münden. Zurückführen lässt sich dieses Muster auf artikulatorische Voraussetzungen, d.h. auf den Abfall des subglottalen Drucks im Zuge der Ausatmung, welcher wiederum die Stimmlippen langsamer schwingen lässt. Hauser & Fowler (1992) zeigen, dass dieses prosodische Muster auch bei anderen Tieren zwischen zwei- und dreifachen Signalarufen auftritt, bei Meerkatzen

¹⁰ McMahon ist in ihrer Argumentation m.E. an manchen Stellen nicht sehr differenziert. Dass auch der Prozess des Phonemerwerbs hochautomatisch abläuft und eine kategoriale Wahrnehmung voraussetzt, wird in ihrer Darstellung nicht deutlich. Sie folgt hier leider einer vereinfachten Gegenüberstellung von angeboren vs. erlernt, die man aber wohl kaum mit Oberflächenstrukturen begründen kann. In Kapitel 2.1.2 wird gezeigt, dass auch segmentale Merkmale angeboren sein müssen, sich die phonologischen Systeme somit nur in ihrer repräsentationalen Form unterscheiden.

gar analog zur menschlichen Kommunikation das Ende einer konversationellen Einheit markiert. Dies lässt darauf schließen, dass sich konversationell genutzte prosodische Muster lange vor der Produktion von Phonemen entwickelten. Letztere scheint allerdings kein Analogon im Tierreich zu haben, ein Aspekt, der in der vielbeachteten Diskussion zwischen Hauser, Chomsky & Fitch (2002) und Pinker & Jackendoff (2005) um die Frage, welche Komponente des Sprachsystems humanspezifisch sei, keinen geringen Stellenwert einnimmt. Erstere unterscheiden zwischen einem Sprachvermögen in weiterem Sinne (FLB: faculty of language - broad sense) und einem Sprachvermögen in engerem Sinne (FLN: faculty of language - narrow sense). FLB umfasst ein senso-motorisches (phonetisch/phonologisch) und ein konzeptuell-intentionales (semantisches/pragmatisches) System wie auch FLN, wobei die beiden ersten Komponenten bei anderen Tierarten zu gewissem Grade auch beobachtet werden können. FLN hingegen ist ausschließlich menschliches Sprachvermögen und, so die Nullhypothese, als Rekursion definiert. Die Kritik von Pinker & Jackendoff diesen Aspekt betreffend ist nun, dass Phonologie als System (bei ihnen allerdings nicht getrennt in prosodische und segmentale Phonologie) in der Klassifikation überhaupt nicht berücksichtigt wird, sondern nur Perzeption und Artikulation. Selbst bei der Betrachtung dieser Verarbeitung rekurren Hauser, Chomsky & Fitch auf Evidenz, die auf intensiven Trainingseffekten mit Tieren basiert. Solche Evidenz ist in dem Sinne kritisch, als dass z.B. Sprachlaute auch von auditiven Systemen diskriminiert werden können, die nicht für solche Distinktionen spezifiziert sind:

„For example, a mammalian circuit that uses onset asynchrony to distinguish two overlapping auditory events from one event with a complex timbre might be sufficient to discriminate voiced from unvoiced consonants.“ (Pinker & Jackendoff 2005: 207)

Ohne alle Argumente der Diskussion an dieser Stelle auszuführen, soll hier festgehalten werden, dass prosodische Aspekte, die mit kommunikativen Funktionen gekoppelt sind, auch bei anderen Tierarten beobachtet werden können. Was allerdings nicht beobachtet werden kann, ist die Verwendung von bedeutungslosen Einheiten zur Bedeutungsdistinktion. So verwenden zwar einige Vogelarten Kombinationen von bedeutungslosen Einheiten zur temporalen Verlängerung von Gesängen, dennoch dient diese Verlängerung der Markierung sozialer Dominanz und nicht zur Bedeutungsunterscheidung zwischen einzelnen Gesängen (vgl. Hauser 1996). Somit bleiben die Einheiten hier weiter Bestandteil der Funktion tierischer Kommunikationssysteme, in denen Flucht-, Paarungs- und soziale Interaktionssignale unterschieden werden können (vgl. Bickerton 2009).

Auch hirnanatomisch sind die beiden phonologischen Systeme unterschiedlich distribuiert. Die gängige Dichotomie von emotionaler und sprachlicher Prosodie scheint sich zwar nicht uneindeutig auf die Hirnhemisphären abbilden zu lassen (die Darstellung von McMahon zeigt die starke Tendenz, prosodische Verarbeitung als Ganzes rechtshemisphärisch zu verorten), dennoch ist im Gegensatz zu segmentaler Verarbeitung ein klarer Unterschied zu erkennen. Das Modell von Sidtis & Van Lancker Sidtis (2003) unterscheidet verschiedene prosodische Dimensionen bzw. Funktionen, die sich graduell von einer rechtshemisphärischen Verarbeitung nicht-gegliederter, d.h. kontinuierlicher Intonationsinformation zu einer linkshemisphärischen, diskret-kombinatorischen Verarbeitung gliedern lassen (s. Abb. 2).

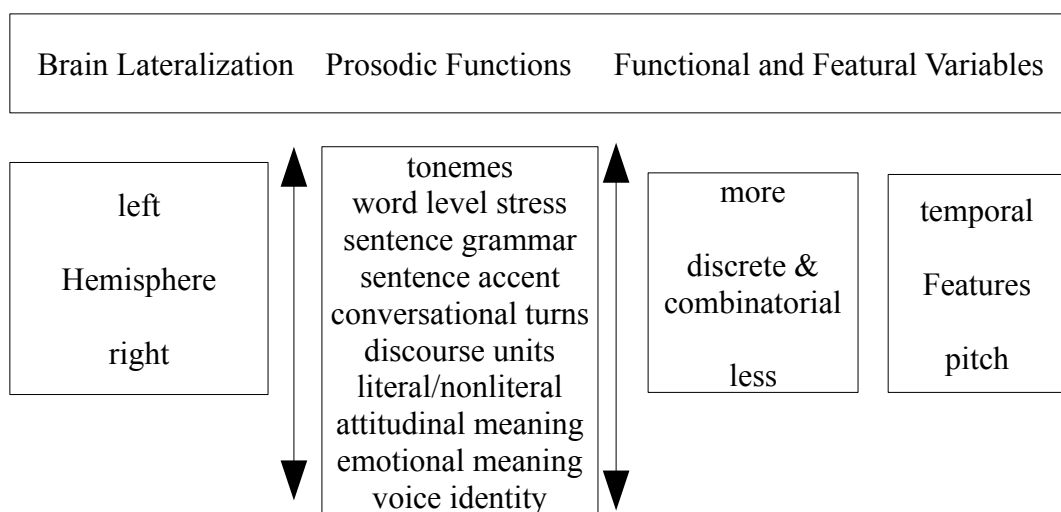


Abbildung 2: Lateralisierung prosodischer Funktionen (nach Sidtis & Van Lancker Sidtis 2003: 98).

Prosodische Verarbeitung ist in diesem Sinne heterogen und bihemisphärisch organisiert. Interessanterweise sind eminente Abhängigkeiten zwischen prosodischer und segmentaler Verarbeitung vorhanden. Zeigen Patienten mit kortikalen Läsionen der linken Hemisphäre und Patienten mit Läsionen der rechten Hemisphäre beide Schwierigkeiten bei der Verarbeitung von Satzprosodie, so sind die Leistungen von Patienten mit rechtshemisphärischer Schädigung signifikant schlechter, wenn segmentale Information herausgefiltert wird (Bryan 1989). Somit kann man mit Friederici (2011: 1376) feststellen: „The less segmental information there is available, the more dominant the RH“ (Anm. des Autors: RH = rechte Hemisphäre).

Es muss hier ausdrücklich festgehalten werden, dass segmentale Information notwendig ist, damit prosodische Information als genuin sprachliche verarbeitet werden kann. Nur in der Koinzidenz von segmentaler und prosodischer Information wird letztere linkshemisphärisch verarbeitet. In neueren Arbeiten ist der linke inferiore frontale Gyrus (LIFG) als das kortikale Gebiet identifiziert worden, welches für die On-line-Verarbeitung sämtlicher sprachlicher Informationen relevant ist (vgl. Hagoort 2005; Hagoort & Levelt 2009). Er umfasst das klassische Broca-Areal (Brodman-Areal 44 und 45) wie auch die Brodman-Areale 47 und 6, wobei die phonologische Silbifizierung in den Arealen 44 und 6 stattfindet (vgl. Hagoort 2005; Indefrey & Levelt 2004). Die phonologischen Wortformen, die in die Silbifizierung eingehen, sind hingegen im linken posterioren Temporallappen gespeichert. Für eine verarbeitungsbasierte Sicht der Linkslateralisierung prosodischer Information sprechen auch Ergebnisse aus Untersuchungen an Sprechern von Tonsprachen. Linkslaterale Aktivierungen lassen sich nur bei sprachbezogenen Aufgaben beobachten (Gandour 2000; Gandour et al. 2004). Sie sind hierbei aber zusätzlich von rechtshemisphärischer Aktivierung begleitet, vor allem im rechten inferioren frontalen Gyrus. Dieser ist bei wechselnder Toninformation bei gleicher Vokalinformation, die im Chinesischen den Träger der Toninformation darstellt, regelhaft aktiv. Ausschließlich linkslaterale Aktivierungsmuster zeigen sich nur bei der Verarbeitung von wechselnder Vokalinformation bei gleicher Toninformation (vgl. Liu et al. 2006).¹¹

Mit McMahon (2007: 181) kann man abschließend eindeutig konstatieren:

„The conclusion to which these various types of evidence cumulatively lead is that prosodic and segmental phonology are separate components neurologically, acquisitionally, and in terms of their relationship with non-human primate vocal communication, where precursors can be found straightforwardly for prosody, but only in a far more rudimentary form for segmentals.“

Angesichts der hier angebrachten Evidenz bleibt die eindeutige Schlussfolgerung, ein phonologisches Modell zu bevorzugen, dass die prosodische und segmentale Phonologie einbezieht, die Relationen zwischen den Systemen in adäquater Art und Weise erfasst, hierbei aber keine Reduktion der Systeme aufeinander vornimmt oder eine Hierarchie aus beiden erstellt. Ein solches Modell stellt die dreidimensionale oder 3D-Phonologie dar, die aus oben genannten Gründen, wie auch aus im nächsten Kapitel darzulegenden Erwägungen, in dieser Arbeit verwendet werden soll.

¹¹ Die divergente neuronale Distribution bei der Verarbeitung phonologischer Information beschränkt sich natürlich nicht nur auf Vokal- und Toninformation. So werden z.B. bei Sprechern des Chinesischen auch Silben und Phoneme in unterschiedlichen Regionen des Gehirn verarbeitet (vgl. z.B. Siok et al. 2003).

1.3.3 Dreidimensionale Phonologie

Die Bezeichnung dreidimensionale Phonologie oder 3D-Phonologie mag den ein oder anderen Leser etwas irritieren. In einem bekannten Aufsatz führten Halle & Vergnaud (1980) den Terminus explizit in den wissenschaftlichen Diskurs ein. Durchgesetzt haben sich aber vor allem die Bezeichnungen autosegmentale oder nichtlineare Phonologie, wobei erstere meist unter letztere subsumiert wird. Erst in neueren konzeptuellen Diskussionen, vor allem durch die Arbeiten von Raimy & Cairns (2009), erlebt die dezidierte Verwendung des Terminus 3D-Phonologie eine Art Renaissance. In diesem Kapitel soll ausgearbeitet werden, warum konzeptuelle Abgrenzungen zu anderen Ansätzen theoretisch sinnvoll und in Hinblick auf die Ergebnisse des vorherigen Kapitels zwingend notwendig sind. Zu diesem Zwecke ist es aber unumgänglich, die historische Genese der phonologischen Ansätze schlaglichtartig zu beleuchten.

Im einflussreichen Werk *The Sound of Pattern of English (SPE)* von Chomsky & Halle (1968) operieren phonologische Regeln über linear geordnete und durch Grenzsymbole (vor allem Morphemgrenzen) unterbrochene Segmentfolgen, welche durch spezifische Merkmalsmengen definiert sind. Die Merkmalsmengen werden als ungeordnete Bündel und die Silbe noch nicht als eigenständige phonologische Repräsentation mit systematischer Bedeutung verstanden. Das SPE-Modell enthält somit nur eine Repräsentationsebene der phonologischen Struktur, weshalb es auch als lineares Modell bezeichnet wird. In den 70er Jahren des 20. Jahrhunderts entstand ein neues Beschreibungsverfahren phonologischer Repräsentationen, das unter dem Namen *Autosegmentale Phonologie* die Kodierung phonologischer Information auf mehreren Ebenen postulierte (vgl. vor allem Goldsmith 1979, 1990). Hier stand vor allem das Phänomen im Vordergrund, dass manche Merkmale diskontinuierlich erscheinen, d.h. sie können sich autonom über Merkmalsmengen und Positionen hinweg verteilen. So z.B. in der Bantusprache Bakwiri, in der bei einer Spielsprache die Silben von zweisilbigen Wörtern vertauscht werden (Durand 1990: 244; die kontrastiven Töne sind mit kleinen Strichen über den Vokalen gekennzeichnet):

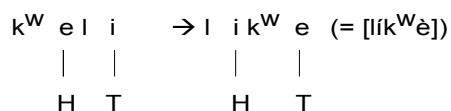
[k^wélí] → [lík^wé]

Bedeutung: „tot“

[k^wéli] → [lík^wè]

Bedeutung: „fallend“

Es könnte nun vermutet werden, dass der Vertauschungsprozess nicht nur die Silben, sondern automatisch auch die mit ihnen assoziierten Töne erfasst. Man kann allerdings beobachten, dass *ì* zu *í* und *é* zu *è* wird. D.h. die Tonmerkmale werden neu zugeordnet, weil sie in ihrer Position verbleiben und die Merkmale von /i/ und /e/ werden mit ihnen assoziiert (Hall 2000: 157; H = hoher Ton, T= tiefer Ton):

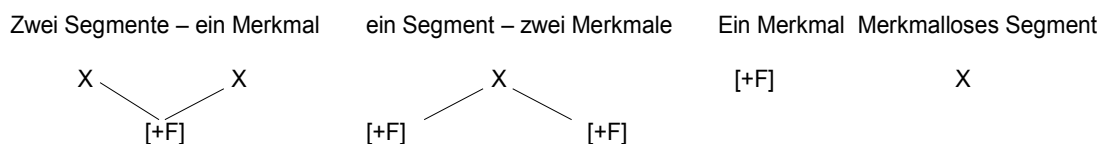


Die Assoziation von einer spezifischen segmentalen Merkmalsmenge und dem jeweiligen Tonmerkmal findet demnach zwischen zwei voneinander unabhängigen Ebenen statt. Tonmerkmale sind also in einem solchen Sinne unabhängig von der Segmentschicht, als dass sie auf einer eigenen Ebene repräsentiert sind. Den Zusammenhang der beiden Ebenen in obigen Beispiel kann man in einem Assoziationsschema formalisieren (F=Feature/Merkmal):

Ein Segment – ein Merkmal



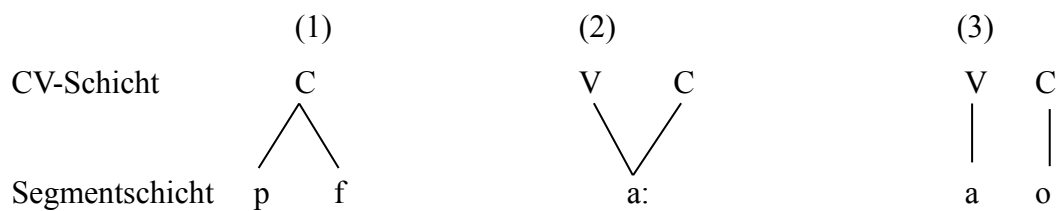
Selbstverständlich existieren noch weitere mögliche Assoziationsrelationen zwischen den Ebenen (vgl. Dinnsen 1997):



Diese Idee nichtlinearer Assoziationen zwischen mehreren phonologischen Merkmalsebenen hatte einen enormen Einfluss auf spätere Theorieentwicklungen. Usus ist in der Generativen Phonologie seit Clements & Keyser (1983) die Mindestannahme von zwei miteinander assoziierten Ebenen: Eine, auf welcher segmentale Information repräsentiert ist, und eine weitere, die Quantitätsinformationen encodiert. In dem sogenannten CV-Modell der Silbe wird die

CV-Schicht zwischen Segment- und Silbenebene verortet (vgl. ebd.). C und V stehen in diesem Modell nicht einfach nur für *consonant* und *vowel*, also für Schließ- und Öffnungsphase des Artikulationsablaufs, sondern stellen Einheiten dar, welche die zeitliche Dauer von Segmenten repräsentieren.

Wird die Merkmalsmenge einfacher Konsonanten und kurzer Vokale 1:1 mit einem C, bzw. einem V assoziiert, so kann man mittels der CV-Schicht auch die differente Dauer von komplexen Segmenten (wie Affrikata) (1), Langsegmenten (2) und Diphthongen (3) abbilden.



Mithilfe dieser Notation ließen sich z.B. in einer Sprache wie dem Türkischen disparate Phänomene wie, Degemination, Vokalepenthese und Vokalkürzung als Beiprodukte der strukturellen Organisation von Segmenten in CVC-Silben beschreiben. Genauer haben alle genannten Prozesse einen einzigen Effekt, nämlich extrasilbische Elemente aus der jeweiligen Repräsentation zu entfernen (vgl. ebd.: 61; Abb. 3).

Akkusativ	Nominativ	Ablativ		
hiss -i	his	his -ten	'Gefühl'	(Degemination)
devr -i	devir	devir -den	'Transfer'	(Vokalepenthese)
zamaan -i	zaman	zaman -dan	'Zeit'	(Vokalkürzung)

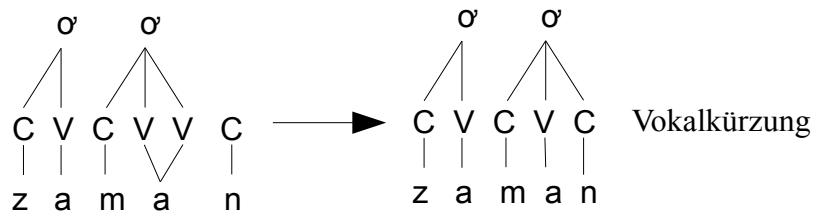
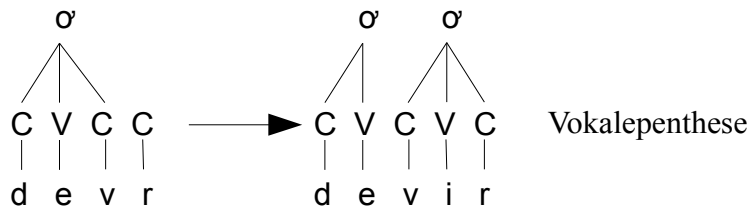
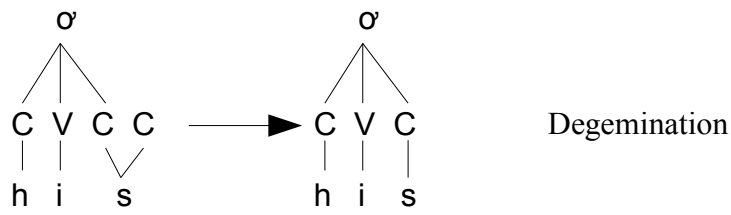


Abbildung 3: Degemination, Vokalepenthese und Vokalkürzung im Türkischen.

Die kanonische minimale und maximale Silbenform einer Zielsprache konstituiert sich nach diesem Modell aus den möglichen Assoziationen zwischen der CV- und Segment-Schicht. Für das Deutsche lässt sich nach Wiese (1996: 38) folgende kanonische Form bzw. Silbenstruktur annehmen:

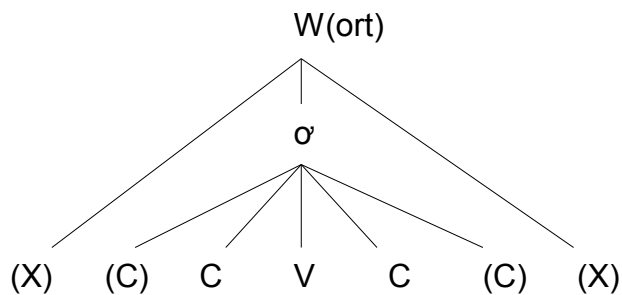
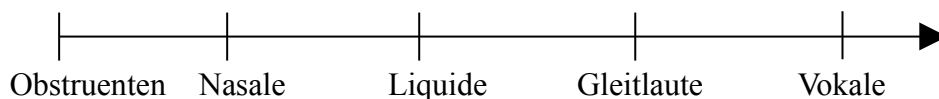


Abbildung 4: Maximale Silbenstruktur nach Wiese (1996).

Die interne Gliederung der Silbe wird nach Maßgabe von Sonorität organisiert, die im nächsten Kapitel detailliert besprochen werden soll. Sie soll hier vorerst als Grad informationeller Saliens einer gegebenen Merkmalsmenge in Abhängigkeit von einem phonologischen Kontext definiert werden. Diese informationelle Saliens korreliert mit Graden der Wahrnehmbarkeit.¹² Phoneme werden innerhalb einer Silbe regelhaft nach zunehmender bzw. abnehmender Sonorität angeordnet. Diese Regelmäßigkeit wird vom Sonoritätsprinzip (SSP) (engl. Sonority Sequencing Principle oder auch Sonority Sequencing Generalization) erfasst (Selkirk 1984: 116):

„Sonority Sequencing Generalization (SSG)
 In any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values.“

Man kann mit Hall (2000: 224) deshalb durchaus festhalten, dass Sonorität in der Silbenphonologie als wichtigstes universelles phonotaktisches Prinzip gelten kann. Allerdings gilt es, hier eindeutig zu differenzieren. Das Sonoritätsprinzip soll nicht alle möglichen Silben von Sprachen spezifisch definieren, sondern als notwendige Beschränkung den Kombinationsmöglichkeiten von Phonemen in Silben unterliegen (vgl. Goldsmith 1990: 111). Die Anordnung von Lauten in Skalen nach der Eigenschaft Sonorität wird als *Sonoritätshierarchie* bezeichnet, wobei an dieser Stelle beispielhaft ein Vorschlag von Clements (1990) stehen soll:



Bei gegebenem Vokal bildet dieser notwendig den Silbengipfel. Es stellt sich hieraus die Frage, inwiefern auf der CV-Schicht vokalische „Identität“ repräsentiert sein muss. In vielen Arbeiten (z.B. Clements & Keyser 1983, Clements 1990, Wiese 1996) wird das Merkmal [silbisch] auf der CV-Ebene verortet, solchermaßen also der Unterschied zwischen Konsonanten und Vokalen und mit diesem der Silbengipfel markiert. So folgt die Silbifizierung nach Wiese 1996 (51-53) folgenden Regeln:

Skeletal structure

Assign X to all root nodes.

¹² Sonorität wird häufig mit Wahrnehmbarkeit in Bezug auf Schallfülle definiert. Ein solch auditives Kriterium wie auch Prominenz nach Maßgabe der Enge des Ansatzrohres wird aufgrund der Wirksamkeit von Sonorität in Gebärdensprache (vgl. Kap. 1.4.2) in dieser Arbeit nicht als sinnvoll erachtet.

Syllable head assignment

- a. Assign [+syllabic] to an X, unless another X immediately precedes.

[-consonantal] [-consonantal]

- b. Join all X dominating [-consonantal] into a nucleus N.

Default for skeletal positions

Assign [-syllabic] to every X.

Besonders aufschlussreich für die Frage nach der Verortung des Merkmals [silbisch] ist hier ein Beispiel von Guerssel (1986), der eine Asymmetrie zwischen Vokalen und Gleitlauten im Berberischen festgestellt hat (die Darstellung folgt Goldsmith 1990: 153). Betrachtet man die Minimalpaare /wci/ (geben) und /ucy/ (aufwachen), so sind sie in Kontexten wie

ur uci-x 'Ich habe nicht gegeben'
ur uci-x 'Ich bin nicht aufgewacht'

homophon, zeigen aber in anderen Kontexten ein breites Spektrum an Varianten:

a) ur y-uci wma 'mein Bruder hat nicht gegeben' /wci/
oder ur i-wci wma
aber *ur y-ucy uma
 *ur i-wcy uma
b) ur y-uci wma 'mein Bruder ist nicht aufgewacht' /ucy/
oder ur y-ucy uma
aber *ur i-wci wma
 *ur i-wcy uma

Guerssel nimmt nun an, dass die möglichen Varianten auf einen Silbifizierungsprozess zur Bildung von CV(C)-Silben mit obligatorischem Onset zurückzuführen sind, der frei von links oder rechts starten kann. Wort-mediale leere Onsets werden per Default mit y gefüllt. Hieraus ergeben sich bei der zugrundeliegenden Form (Abb.5), die Silbifizierungsvarianten 1 (Abb. 6) und 2 (Abb. 7) (s. ebd. 154).

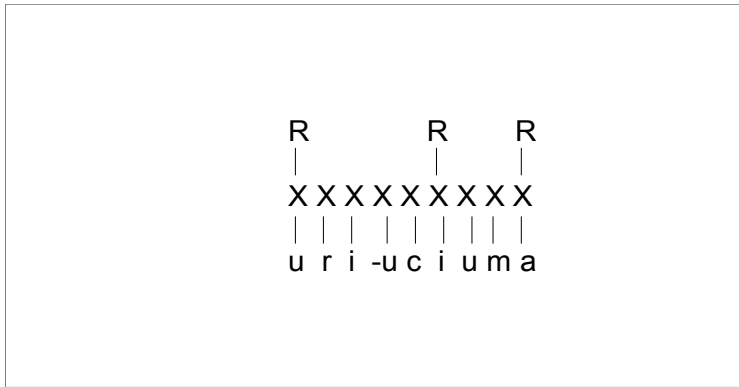


Abbildung 5: Zugrundeliegende Form von 'Mein Bruder hat nicht gegeben' /wci/.

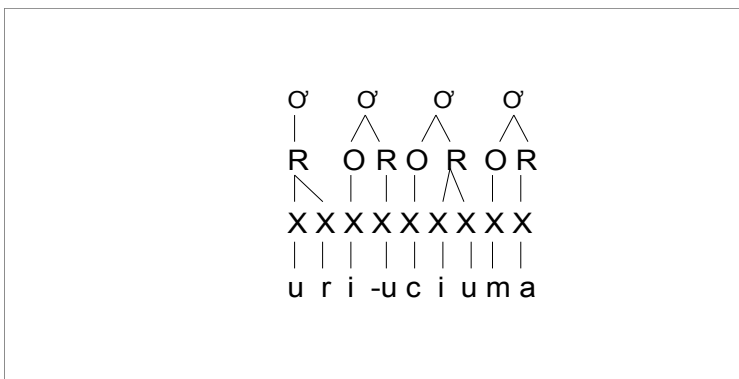


Abbildung 6: Silbifizierungsvariante 1 /ur y-uci wma/.

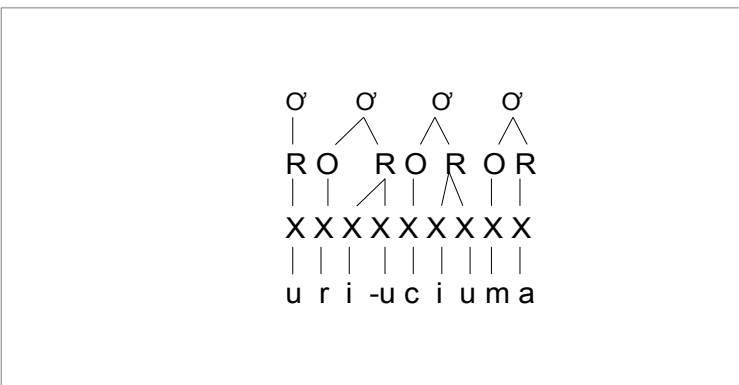


Abbildung 7: Silbifizierungsvariante 2 /ur i-wci wma/.

Die Asymmetrie von Vokalen und Gleitlauten lässt sich so sehr gut beschreiben. In Abb. 6 wird das zugrundeliegende links stehende /i/ mit einer Onset-Position assoziiert, weshalb es per Default als [y] realisiert wird. Der letzte Vokal /u/ nimmt nicht die Nukleus-Position ein und wird deshalb als Gleitlaut [w] produziert. In Abb. 7 betrifft dies alle zugrundeliegenden Phoneme /u/, während alle /i/s mit Nukleus-Positionen assoziiert sind und deshalb auch als [i] umgesetzt werden.

Folgendes wird an diesem Beispiel deutlich: Das Merkmal [silbisch] kann als segmentales Merkmal verstanden werden und nur die wirklichen hohen Vokale sind entsprechend in obigem Beispiel für das Merkmal [+silbisch] spezifiziert. Dementsprechend bilden auch nur sie konstant den Silbennukleus. Man benötigt deshalb auch keine zusätzlichen Regeln zur Bestimmung desselben. Insofern erscheint es plausibler von einer X-Ebenen-Theorie auszugehen (engl.: X-tier-theory; vgl. z.B. Levin 1985), deren X-Elemente, wie wir später sehen werden, im Grunde nur eine Schnittstellenfunktion für dreidimensional organisierte Repräsentationen einnehmen. Nach Goldsmith (1990: 65 und 151) kann man so aus einem lexikalischen Eintrag linear geordneter Merkmalsmengen¹³ sowohl Quantität als auch Qualität, d.h. Silbenkonstituenten, ableiten. Für den Prozess der Ableitung nimmt er einen Splitting-Prozess an, der in Abbildung 8 dargestellt ist. Hier wird das Merkmal [silbisch] von der jeweiligen Merkmalsmenge abgetrennt, so dass es gesondert in die genannten Berechnungen eingehen kann.

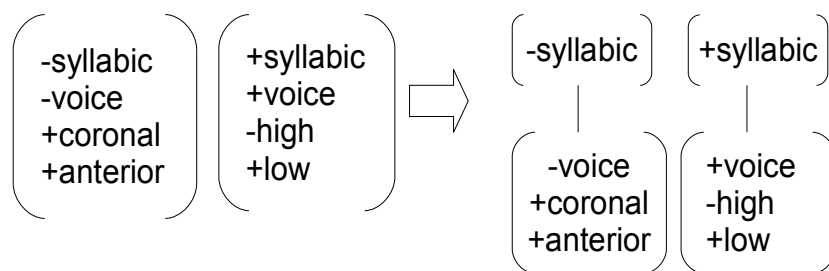


Abbildung 8: Splitting nach Goldsmith (1990).

Für die Spracherwerbsforschung ist die Frage danach, ob die CV-Ebene nicht nur eine einfache Schnittstellenfunktion hat, sondern auch einen eigenen repräsentationalen Merkmalsgehalt besitzt, von keiner geringen Bedeutung. Man führt durch die Annahme eines eigenen repräsentationalen Merkmalsgehaltes dieser Ebene natürlich in einem gewissen Sinne eine Top-down-Perspektive ein, da hierdurch segmentale Merkmale von höher liegender, prosodischer Information lizenziert wird. Diese Perspektive ist in Hinblick auf die vorherige theoretische Besprechung wie auch auf die externe Evidenz aus dem vorherigen Kapitel fraglich, kann zudem aber im Zusammenhang mit prosodischer Lizenzierung theoretisch noch umfassender kritisch beleuchtet werden. Prosodische Lizenzierung beschreibt nichts anderes, als dass in einem Top-down-Verhältnis niedrigere Konstituenten zu höheren Konstituenten gehören und niedrigere Einheiten in der Hierarchie durch höhere sanktioniert werden müssen,

¹³ Dieser Ansatz wird von Raimy (2000) aufgegriffen und für die Beschreibung von Reduplikationen fruchtbar gemacht (vgl. Kap. 1.7.1).

damit sie letztlich phonetisch interpretierbar sind (vgl. Itô 1986). D.h. nach der prosodischen Hierarchie in Abbildung 9, müssen z.B. Silben in Füße und Füße in phonologische Wörter eingebettet sein (vgl. Selkirk 1984; Nespor & Vogel 1986).

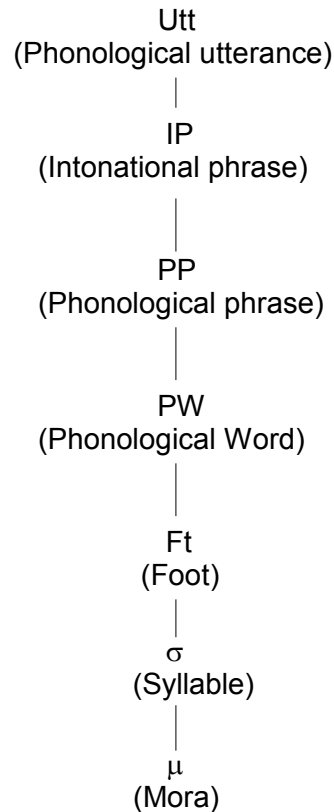


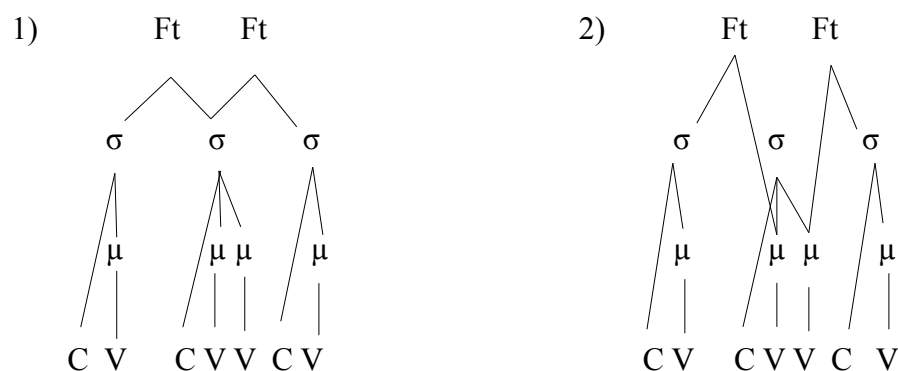
Abbildung 9: Prosodische Hierarchie.

Cairns (2009) bezeichnet ein solches Modell als 2D-Modell. Betonung wird in einem solchen Modell direkt aus der Geometrie der prosodischen Hierarchie abgeleitet. Terminale Segmente werden direkt durch Silbenknoten dominiert. Extrasilbische Segmente oder Appendizes, welche durch Verletzung des SSP oder der maximalen Silbenstruktur nicht silbifiziert werden können, werden in diesem Modell durch höhere Knoten dominiert.

Nach Cairns (2002; 2009) ist das 2D-Modell aber empirisch inadäquat, weil es Verletzungen des Prinzips der Silbenintegrität (SI) nicht erklären kann (Cairns 2002: 1):

„Syllable Integrity prohibits: ... $(v)v[v]$..., where syllable boundaries are shown by square brackets and foot boundaries by parentheses. Equivalently: No language may make a contrast between tautosyllabic $v\acute{v}$ and $v\acute{v}$.“

Blevins (1995) vertritt z.B. die Position, dass in keiner Sprache ein Betonungscontrast zwischen $\acute{v}v$ und $v\acute{v}$ besteht. Im Südlichen Paiute aber, einer Shoshonensprache der Uto-Aztekischen Sprachfamilie, werden adjazente bzw. tautosilbische Vokale unterschiedlichen prosodischen Füßen zugeordnet. In Südlichen Paiute wird das Betonungsmuster anhand von Vokalen und nicht Silben berechnet. Ein Beispiel von Cairns (2002) soll den Prozess illustrieren. Ein Wortstamm der Form /papapaa/ weist zwei tautosilbische Vokale auf. Findet eine Affigierung des Suffix /-ppapaa/ statt (/papapaa + ppapaa/), so fällt die Betonung auf den zweiten der tautosilbischen Vokale (/papàpaàppapàa/). Wird zusätzlich ein monomorphisches Element zur Linken an den Stamm affigiert (/nam+papapaa+ppapaa/), so fällt die Betonung nun auf den ersten der tautosilbischen Vokale (/nampàpappàappàpaa/). Zudem wird der zweite Vokal entstimmlicht ([nampàwawàA^hpàwaA]), was ein völlig differentes Verhalten der tautosilbischen Folge $\acute{v}v$ gegenüber $v\acute{v}$ zeigt und somit eine klare Verletzung des Prinzips der Silbenintegrität darstellt. Entscheidend ist nun, dass das Modell der Prosodischen Hierarchie diese Verletzung nicht erfasst. Wenn man z.B. eine ambifüßige Struktur wie in 1) annimmt, ist nicht klar, dass der linke Vokal der gemeinsamen Silbe der Kopf des linken Fußes ist und der rechte Vokal der gleichen Silbe als unbetonter zum rechten Fuß gehört. Eine Struktur wie 2) zeigt eine uneindeutige Zuordnung der Füße zu den Moren, zurück bleibt hier aber eine ungeparste Silbe, welche auch die entsprechenden Moren dominiert. Wie Cairns betont, ist diese „so hopelessly ad hoc and unnatural“ (Cairns 2002: 2), dass man eine solche Struktur nicht den Optionen der UG zuordnen kann.



Ein solches Problem ist mit der Anwendung eines 3D-Modells nicht gegeben. In diesem sind disparate sprachliche Informationen (segmentale, silbische (im Sinne von Silbenrepräsentationen) und metrische) auf verschiedenen Ebenen (engl. plane) orthogonal um eine Zentralachse, die X-Ebene, angeordnet und über diese miteinander assoziiert (s. Abb. 10). Betonungsmuster werden im 3D-Modell durch metrische Gitter aus Asterisksen dargestellt

(Lieberman 1975). Ein gängiges Beispiel für das Auseinanderfallen von metrischer und silbischer Information stellen Schwa-Silben dar. Diese sind aus metrischer Perspektive inert, bilden dennoch aber regelhaft den Kopf bzw. Nukleus einer Silbe. Ein Prozess der prosodischen Lizenzierung, wie auch überhaupt die Annahme einer strikt geordneten prosodischen Hierarchie, wird deshalb auch von Vertretern eines solchen Modells explizit abgelehnt (z.B. Halle & Idsardi 1995: 339).

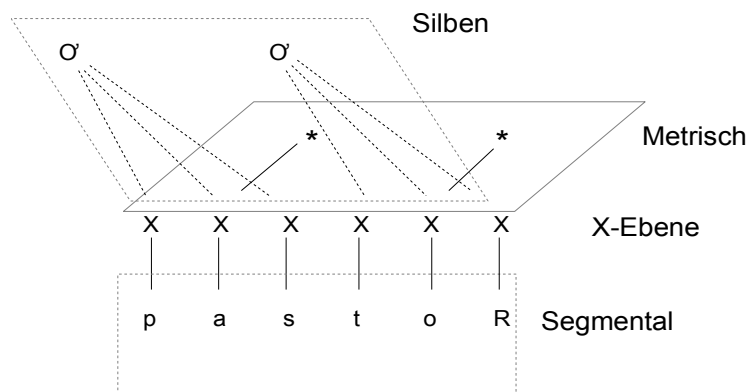


Abbildung 10: Beispiel einer dreidimensionalen Struktur.

Das Beispiel aus dem Südlichen Paiute kann in diesem 3D-Modell problemlos analysiert werden (nach Cairns 2002: 3). Der Kopf des Fußes ist in Abbildung 11 auf Linie 1 repräsentiert, so dass man hier das jambische Muster erkennen kann. Auf Linie 0 werden die Fußgrenzen durch Klammern angezeigt. Man erkennt die eindeutige Verletzung des Prinzips der Silbenintegrität im Südlichen Paiute durch das Nichtübereinstimmen von der ersten Fuß- und Silbengrenze (für andere Beispiele der Verletzung des Prinzips der Silbenintegrität s. z.B. Halle 1990).

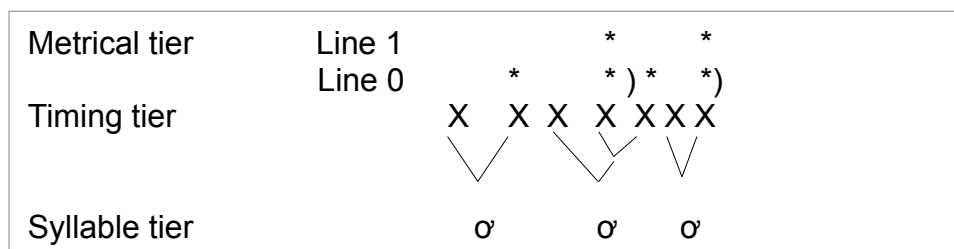


Abbildung 11: Dreidimensionale Repräsentation im Südlichen Paiute (ohne Segmentebene).

An dieser Stelle soll ein Beispiel für einen möglichen Algorithmus nach Halle & Idsardi (1995) angeführt werden. Metrische Konstituenten werden hier als Gruppierungen von

Projektionen derjenigen Einheiten verstanden, die Betonung tragen können. Zwei Mechanismen werden in diesem Modell benötigt. Projektion generiert eine Linie aus Taktschlägen und Konstituentengrenzen aus einer jeweiligen niedrigeren Ebene. Ein anderer Mechanismus konstruiert Konstituenten durch die Platzierung von Grenzen. Beide Anwendungen werden durch Parameteroptionen beschränkt. Demgemäß werden in einem ersten Schritt betonbare Elemente (hier Silben) auf Ebene 0 projiziert:

Line 0 mark projection

Project a line 0 element for each syllable head.

Dass manche Silben prominenter als andere sind, wird mit dem Syllable boundary Projection Parameter erfasst:

Syllable Boundary Projection

Project the *left/right* (L/R) boundary of *certain syllables* onto line 0.

Mit dem Head Location Parameter wird nun das jeweilige Element einer (durch Klammerung gekennzeichneten) Konstituente identifiziert, das als einziges auf Ebene 1 projiziert werden kann.

Head Location Parameter

Project the *L/R*-most element of each constituent onto line 1.

Der folgende Parameter gibt an, zu welcher Seite hin eine metrische Kette geschlossen wird.

Edge-Marking Parameter

Insert a *L/R* bracket to the *L/R* of the *L/R*-most element in the string.

Für das Deutsche könnte die Anwendung des Algorithmus auf einzelne Wörter wie folgt aussehen. Jede Silbe erhält eine Markierung auf Ebene 0 (Line 0 projection).

0	X	X	X	X	X	X	X	X
	ba	na	nə	ʔe	lə	fant	ho	sə

Es werden zudem die Köpfe gekennzeichnet. Im Deutschen ist es umstritten, ob schwere Silben (CVV oder CVC) (Uhmann 1991) oder Vollvokale Prominenz repräsentieren (Féry 1986). Da nach der Theorie von Wiese (1996) jeder Vollvokal gleichzeitig eine CVC Silbe bildet, wird hier mit dieser Annahme diese Differenz nicht weiter diskutiert. So erhalten alle Silben mit Vollvokal, die gleichzeitig CVC-Silben sind, eine Silbengrenzenmarkierung (Syllable Boundary Projection).

0	(X	(X	X	(X	X	(X	(X	X
	ba	na	nə	ʔe	lə	fant	ho	sə

Die so gekennzeichneten Köpfe werden nach Maßgabe des Head Location Parameters (hier gilt das meist links liegende Element einer durch Grenzen markierten Konstituente) auf die nächste Ebene projiziert.

1	X	X		X		X		X
0	(X	(X	X	(X	X	(X	(X	X
	ba	na	nə	ʔe	lə	fant	ho	sə

Nun wird die Belegung des Edge-Marking-Parameters relevant, der im Deutschen eine Klammer um das meist rechts liegende Element der Ebene 1 verlangt.

1	X	(X		X		(X	(X	
0	(X	(X	X	(X	X	(X	(X	X
	ba	na	nə	ʔe	lə	fant	ho	sə

Daraus folgt nach letzter Projektion folgendes Betonungsmuster:

2		X				X	X	
1	X	(X		X		(X	(X	
0	(X	(X	X	(X	X	(X	(X	X
	ba	na	nə	ʔe	lə	fant	ho	sə

Es soll an dieser Stelle nicht um Feinheiten des hier nur exemplarisch angewandten Modells gehen, sondern um das Prinzipielle des Verfahrens (für eine etablierte Darstellung der Anwendung metrischer Gitter im Deutschen vgl. Féry (1986) und Uhman (1991)).

Auf Einheiten wie Mora und Fuß kann in diesem Modell verzichtet werden, d.h. es verlangt weniger repräsentationale Einheiten zur Beschreibung der Betonungszuweisung. Im Zusammenhang mit diesem Verzicht fällt auch die Annahme weg, dass der Fußaufbau im Deutschen strikt von rechts beginnt, Silben also von rechts nach links geparst werden (vgl. z.B. Penner et al. 2006). Das metrische Modell kann aber durch den Edge-Marking-Parameter trotzdem die potentielle Rechtsorientierung der Hauptbetonung im Deutschen erfassen. Dies

ist für die Interpretation von Spracherwerbsdaten von Relevanz, da vor allem die Einheit Fuß in vielen Ansätzen eine zentrale Rolle einnimmt (vgl. Demuth & Fee 1995; Fee 1996; Fikkert 1994).

Letztlich stellt sich noch einmal die Frage nach der Terminologie. Wie wir gesehen haben, sind auch Ansätze, die einem 2D-Ansatz folgen, nichtlineare phonologische Theorien. Nichtlineare Phonologie als Überbegriff verdeckt also wichtige Unterscheidungen. Auch die Bezeichnung Autosegmentale Phonologie verdeckt viele feinkörnige Unterschiede. Zum einen kann eine autosegmentale Theorie agnostisch gegenüber der Frage sein, welche Notation für die Betonungszuweisung verwendet wird, also Baum- oder Gitterrepräsentationen.¹⁴ Zum anderen kann ein autosegmentaler Ansatz durchaus jegliche Hierarchisierung von Merkmalen negieren, so dass jedes Merkmal unabhängig voneinander mit einer Silbenposition assoziiert ist (z.B. Goldsmith 1976). In der hier zu vertretenden 3-D-Theorie sind Hierarchien durchaus möglich, allerdings nur im jeweils eigenen Berechnungsvokabular der drei angenommenen (und durch die X-Ebene verbundenen) Ebenen (vgl. Abb. 10). Auf Silbenebene die Hierarchie von Onset, Nukleus und Reim, auf der metrischen Ebene eine Prominenzhierarchie und auf segmentaler Ebene die Merkmalsgeometrie. In Abbildung 12 wird die theoretische Zu- bzw. Einordnung der 3D-Phonologie als Teiltheorie der nicht-linearen Phonologie noch einmal stark simplifiziert illustriert.

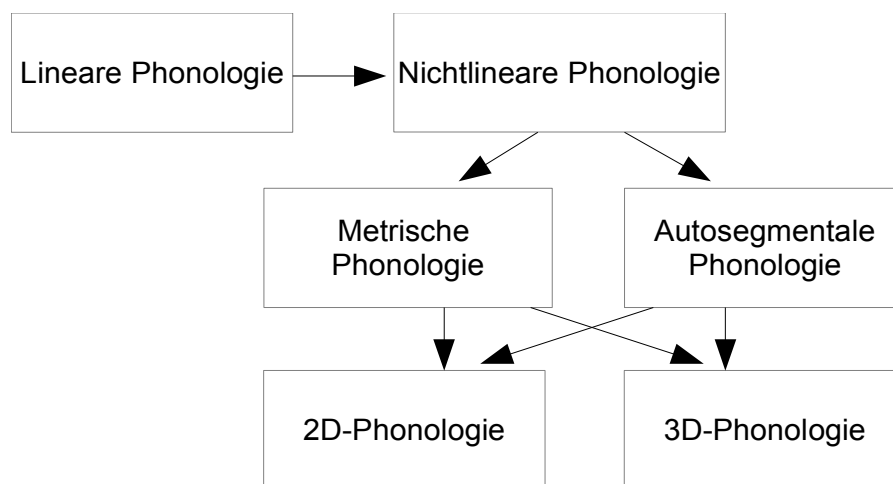


Abbildung 12: Teiltheorien der Phonologie.

¹⁴ Auch wenn z.B. Goldsmith (1990: 193) die Annahme die Annahme eines metrischen Gitters als autonome Struktur favorisiert.

Mit dem Aufbau des 3D-Modells genügt man somit der Unterscheidung von Prosodischer und Segmentaler Phonologie, die in Kapitel 1.3.2 durch verschiedene Evidenz begründet wurde. Es wurde bereits darauf hingewiesen, dass es für den Silbifizierungsprozess wie auch für die Zuweisung von Betonung eine überaus wichtige Frage ist, nach welchen Kriterien segmentaler Gehalt in die X-Ebene integriert wird. Das viel diskutierte und theoretisch oft unscharf definierte Konzept Sonorität nimmt für diesen Integrationsprozess eine zentrale Rolle ein, weshalb es einer genaueren Betrachtung unterzogen werden soll.

1.4 Sonorität

1.4.1 Definitiorische Annäherungen

Zur Beschreibung der internen Organisation der Silbe, d.h. für die Abfolgeregularitäten von Phonemen innerhalb derselben, wird meist das Sonoritätsprinzip (SSP) angenommen. Die prominente Definition von Selkirk (1984: 116) soll an dieser Stelle noch einmal wiederholt werden:

“Sonority Sequencing Generalization (SSG (Anm. d. Verf.: auch SSP))
 In any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values.”

Die Anordnung von Segmenten folgt einer Skalierung nach Maßgabe von Sonoritätswerten, die als Sonoritätshierarchie bezeichnet wird. Es gibt verschiedene Darstellungen der Hierarchie, wahrscheinlich auch, weil unklar ist, wie genau Sonorität definiert bzw. motiviert werden soll. Hier zwei Beispiele unterschiedlicher Anordnung:

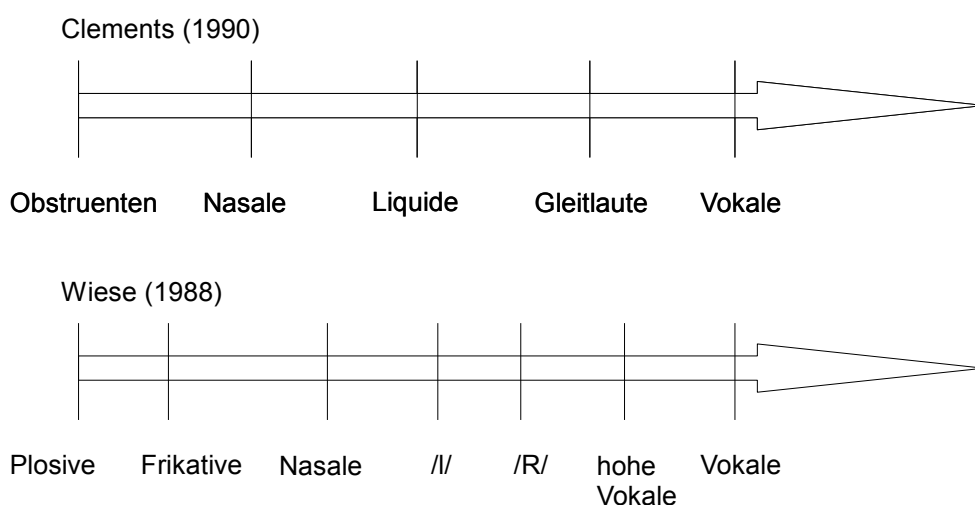


Abbildung 13: Beispiele für Sonoritätshierarchien.

Das Sonoritätsprinzip selbst wie auch die Sonoritätshierarchie sind allerdings neutral gegenüber der Frage, was Sonorität selbst sei, d.h. nach welchen phonetischen oder phonologischen Merkmalen oder Eigenschaften sie operationalisierbar ist. Meist wird sie, wie Eisenberg (2000: 104) feststellt, nach den Vätern des Sonoritätsgedankens, Eduard Sievers (1901) und Otto Jespersen (1913), artikulatorisch oder auditiv interpretiert:

„Unter Sonorität kann entweder ein Maß für die >Schallfülle< oder auch >Schallstärke< von Lauten verstanden werden, zum anderen so etwas wie der artikulatorische >Öffnungsgrad<. Im ersten Fall fasst man die Silbenfolge einer sprachlichen Äußerung als einen Wechsel von laut und leise oder schallintensiven und weniger intensiven Abschnitten des Lautkontinuums auf, im zweiten Fall als eine Folge von Öffnungs- und Schließbewegungen. Beide Betrachtungen haben ihren guten Sinn, nur bleibt unklar, wie man aus einer von ihnen oder aus beiden gemeinsam das Sonoritätskontinuum ableiten kann.“

Evidenz gegen eine solche modalitätsabhängige Definition von Sonorität ist allerdings eindeutig durch Gebärdensprachen gegeben. Analog zu den Untersuchungen der Lautsprache werden auch in der Gebärdensprachforschung unterschiedliche Sonoritätshierarchien angenommen (vgl. auch Corina & Sandler 1993; Sandler 1993). In Blevins (1993) wird z.B. folgende Anordnung von geringer Sonorität zu hoher Sonorität angenommen:

1: dominant hand as a static articulator < 2: non-dominant hand as a static articulator < 3: non-static articulator (= handshape or orientation change) < 4: path movement

Wenn auch in der Gebärdensprachforschung ebenfalls unterschiedliche Sonoritätshierarchien vorgeschlagen werden, so verbindet doch alle Ansätze die Grundannahme, dass visuelle Salienz ausschlaggebend für die Skalenanordnung ist. Spezifischer bedeutet dies, dass die Handbewegung (Movement) bei gleichbleibender Handform und Handstellung in Gebärden visuell salienter ist als eine Handkonfiguration, d.h. Handform und Handstellung, an einem bestimmten Ausführungsort (Hold).

Das repräsentationale Format von Silben in der deutschen Gebärdensprache (DGS) lässt sich nach Happ & Vorköper (2006: 57) demnach pointiert erfassen:

„Die Silbe ist immer eine Kombination aus den Silbenpositionen H (Hold) und M (Movement).“

In Gebärdensprachen folgt die interne Organisation der Silbe also den gleichen abstrakten Prinzipien wie in Lautsprachen, d.h., es ist ein Silbenkern, nämlich M, mit einem Maximum an Sonorität gegeben und vorangehenden und folgenden, weniger sonoranten, Elementen, den jeweils besetzten Hold-Positionen (vgl. Abb. 14). Diese regelhafte Abfolge findet sich

entsprechend innerhalb der maximalen Silben HMH der Deutschen Gebärdensprache (DGS) (nach Happ & Vorköper 2006: 57):

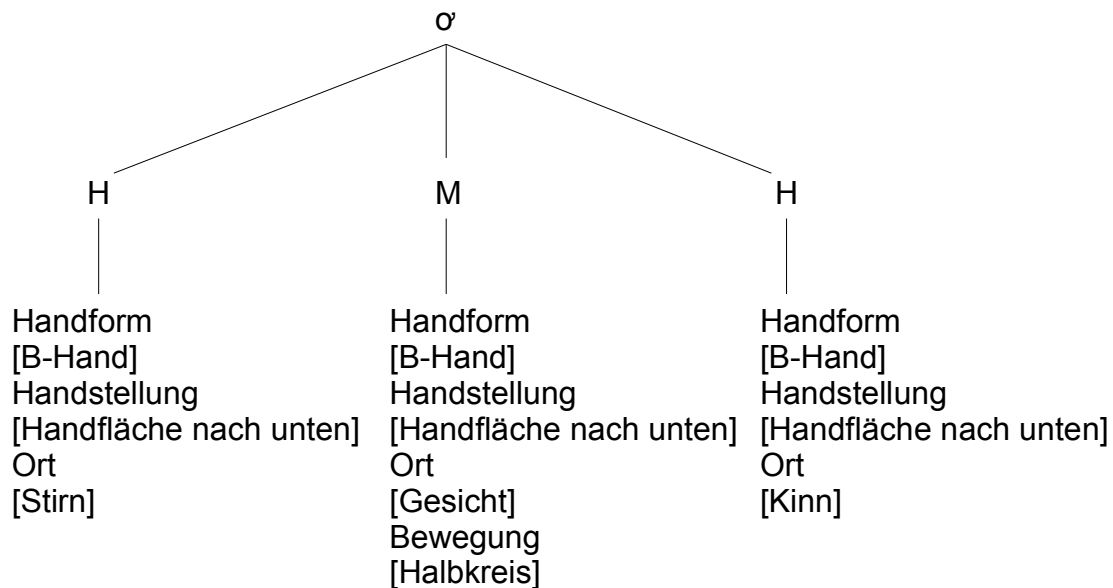


Abbildung 14: Silbenstruktur von VATER.

In Anbetracht der Indifferenz der sprachlichen Verarbeitung gegenüber der jeweiligen Modalität, d.h. ihrer grundsätzlichen Modalitätsneutralität, mit der die Klassifikation und Berechnung von Silbenpositionen erfolgt, ist m.E. die Rede von auditiver oder visueller Salienz irreführend. Da für die zentrale Verarbeitung die Modalität nicht ausschlaggebend ist, sondern offensichtlich nur eine gestufte, kategorial berechnende Salienz an sich, wird hier die im vorigen Kapitel schon eingeführte Bezeichnung der informationellen Salienz vorgezogen. In diesem Sinne kann man weiterführend konstatieren, dass Sonorität ein kognitives Prinzip sein muss und es daher gilt, eine Definition von Sonorität vorzuziehen, die der dargestellten Modalitätsneutralität Rechnung trägt und einer rein grammatischen Klassifikation folgt. Unter Zugrundelegung von Jackendoffs Schnittstellenmodulen (vgl. Kap.1.2.1) korrespondieren somit phonologische Repräsentationen respektive Merkmale mit externalen artikulatorisch-auditiven und motorisch-visuellen Input-Output-Konfigurationen. Wenn man Sonorität als ein kognitives Prinzip definiert, stellt sich weiter die Frage, auf welcher repräsentationalen Ebene Sonoritätsbeschränkungen wirksam sind.

Wiese (1996) und Neef (2004) argumentieren dafür, dass Sonoriätsbeschränkungen auf der CV-Ebene zu verorten sind. Das SSP wird nach Wiese (1996: 59) im Deutschen häufig durch Sonoritätsplateaus wie auch durch Sonoritätsumkehrungen verletzt. So findet man z.B.

Affrikata wie /pf/ sowohl im Onset der Silbe als auch im Silbenreim (s. Abb. 15). Nach dem SSP bilden solche Affrikata ein Sonoritätsplateau, welches nicht vorkommen dürfte. Auch eine Analyse, welche Plosive und Frikative in der Sonoritätshierarchie differenzieren würde, hätte zur Folge, dass nun im Silbenreim eine Sonoritätsumkehrung auftreten würde, welche wiederum das SSP verletzt. Hieraus schließt Neef (2004: 258), „that affricates are units with respect to sonority [...]. Since affricates constitute units on the skeletal tier but two elements on the segmental tier, sonority applies on the skeletal tier only.“ Auch Langvokale stellen nach Neef ein Problem dar (ebd.: 259-260), da sie in ihrer üblichen Assoziation mit zwei Position auf der CV-Ebene eine Sonoritätsplateau bilden würden. Eine Lösung besteht für ihn darin, dass durch die Assoziation mit einer V- wie auch mit einer C-Position, mit letzterer Position selbst das Abfallen der Sonoritätskontur enkodiert ist.

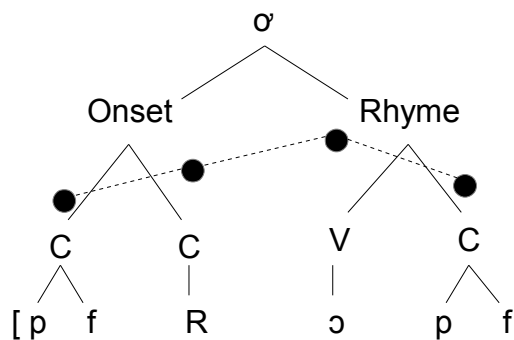


Abbildung 15: Silbenstruktur mit Sonoritätskontur nach Neef (2004).

Es lassen sich m.E. mehrere kritische Einwände gegen die Analyse von Neef erheben. Vor allem scheint hier eine gravierende, notationsbedingte Fehlinterpretation vorzuliegen. Die Argumentation, dass Affrikata eine Position auf der CV-Ebene einnehmen, auf Segmentebene aber hingegen zwei Elemente, ist in Anbetracht von neueren Theorien zur Merkmalsgeometrie nicht haltbar. In diesen werden Affrikata strukturell durch ein Mehr an Merkmalen, aber nur einen Wurzelknoten repräsentiert (vgl. z.B. Van Oostendorp 2000; Van Oostendorp & van de Weijer 2005). D.h., wir haben auf segmentaler Ebene in Passung mit der X- oder CV-Ebene durch den Wurzelknoten eine Position gegeben, so dass kein Widerspruch zu dem von Goldsmith angenommenen Splitting-Prozess besteht (vgl. Kap. 1.3.3). Insofern greift natürlich auch nicht Neefs Schlussfolgerung, dass Sonorität auf der CV-Ebene bestimmbar sei. Eine in diesem Zusammenhang interessante Beobachtung wurde von Lléo & Prinz (1996, 1997) für den Spracherwerb gemacht. Sie fanden, dass der Erwerb des Silbenonsets der von

ihnen beobachten Kinder eine regelhafte Folge zeigte, nämlich von einzelnen Phonemen über Affrikata zu Konsonantenclustern (vgl. Abb. 16). D.h., Affrikata bzw. die sie konstituierenden Merkmalsmengen nehmen eine wichtige Funktion für den Ausbau von Silbenpositionen ein. Wie genau dies geschieht, ist fraglich; spekuliert man, könnte es einfach sein, dass das „überschüssige“ Merkmal der Affrikata die Bindung durch einen weiteren Wurzelknoten initiiert. Solche Phänomene der Bauplanüberfrachtung sind aus dem Syntaxerwerb wohl bekannt (vgl. Tracy 2000: 34). So zeigen folgende Selbstkorrektur-Beispiele wie die *dazə*-Vorläuferkonstruktion als ganzheitlicher deiktischer Platzhalter neue strukturelle Positionen erschließt (ebd.: 33):

J. (2;4) [dazə] BAUernhof \ ... da IS das bauernhof\

J. (2;4) [daza] schava (=Scheuche) ... da IS-i vogelscheuch\

In diesem Sinne können auch Affrikata ein (konventionelles) segmentales Übergangsphänomen darstellen, was u.a. sprachtherapeutisch nutzbar gemacht werden kann (vgl. Gierut & O'Connor 2002).

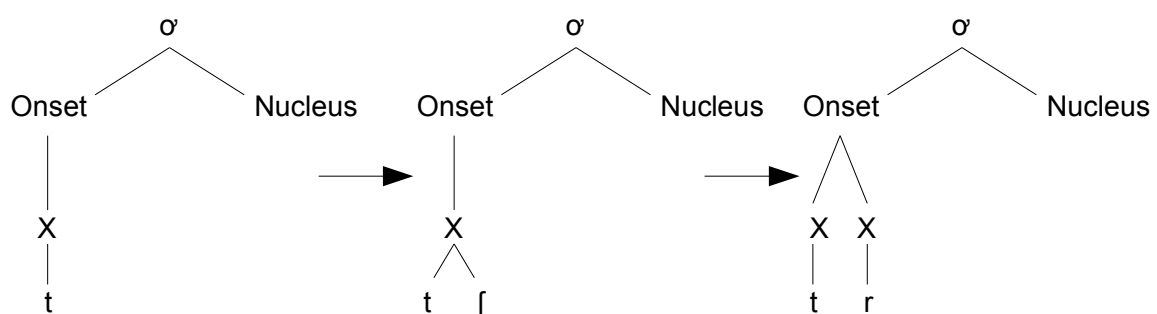


Abbildung 16: Der Erwerb von komplexen Silbenonsets über Affrikata.

In Bezug auf Langvokale sind analoge Einwände gegen die Argumentation von Neef zu erheben, hier gestaltet sich allerdings aufgrund der Unbestimmtheit des Gegenstands die Diskussion schwieriger. Wie Wiese (1996: 15) feststellt, ist nur für kurze Vokale der selbständige phonematische Status unstrittig, Langvokale hingegen verhalten sich hinsichtlich ihrer Distribution wie Kombinationen aus Kurzvokal und Konsonant. Die Entdeckungsprozedur der Minimalpaarbildung verweist allerdings durchaus auf bedeutungsunterscheidende Kontraste durch Langvokale. Aufgrund dieser kontrastiven Funktion wurde in früheren Arbeiten das Merkmal $[\pm \text{lang}]$ verwendet (vgl. Chomsky & Halle 1968), allerdings bleibt die Frage der positionalen Distribution wie auch des Sonoritätsabfalls

mit diesem Merkmal unbeantwortet. Wiese (1996: 154) wie auch Hall (2000: 249-254) verweisen darauf, dass Längung über die X-Positionen spezifiziert werden kann. Zum einen stellt sich aber bei einer zugrundeliegende Markierung von Langvokalen für zwei X-Positionen, wie Wiese sie vorschlägt, die Frage, wieso Formen segmentaler Information nicht die gleiche Funktion haben können bzw. der Rekurs auf eine autonome X-Ebene und damit eine Verdopplung von Entitäten im Sinne von Ockhams Rasiermesser nicht vermieden werden kann. Zum anderen erscheint der Verweis von Hall auf kompensatorische Längung (s. Abb. 17; Hall 2000: 53) auch nicht zwingend eine autonome X-Ebene zu verlangen.

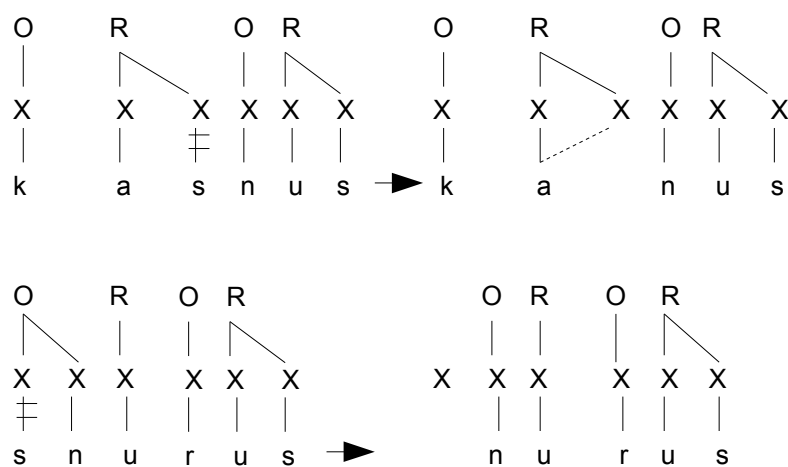


Abbildung 17: Kompensatorische Längung über die X-Ebene.

Da hier eine Tilgung dem Spreading des Vokals vorausgeht, steht dies nicht im Widerspruch dazu, dass durch Splitting zuvorderst die Position spezifiziert wurde. Vielmehr scheint das entscheidende Charakteristikum von Vokalen zu sein, zwei Wurzelknoten bilden zu können. Eine Top-down-Erklärung, wie Neef sie anbietet, scheint nicht zwingend notwendig und folgt letztlich der theoretischen Annahme, die Selkirk (1990: 126) als One-Root Theory of Length bezeichnet. Länge wird hier, wie oben dargestellt, durch eine vom segmentalen Gehalt gänzlich autonome CV- oder X-Ebene quantifiziert. Die von ihr vorgeschlagene „Two-Root Theory of Length“ verortet hingegen aufgrund der Möglichkeit zweier Wurzelknoten, die mit einer Merkmalsmenge assoziiert sind, die positionale Spezifikation von Langvokalen bzw. ihre Längung auf segmentaler Ebene.¹⁵ Mit dieser Theorie lässt sich z.B. sehr gut die Diphthongisierung von Langvokalen beschreiben. Während in der Theorie von Hayes (1990)

¹⁵ Selkirk (1990) verzichtet gänzlich auf eine X- oder CV-Ebene. Zudem verwendet sie in ihrem Modell Morarepräsentationen. Die Details ihres Modells sollen hier aber nicht genauer untersucht werden.

Diphongisierung eine Derivation mittels der Manipulation von Linearität signalisierenden Indizes und doppelten X-Ebenen-Positionen darstellt, ist sie in der Theorie von Selkirk mit den Prozessen Delinking und Spreading ausschließlich auf segmentaler Ebene zu erfassen (ebd.: 135-138).

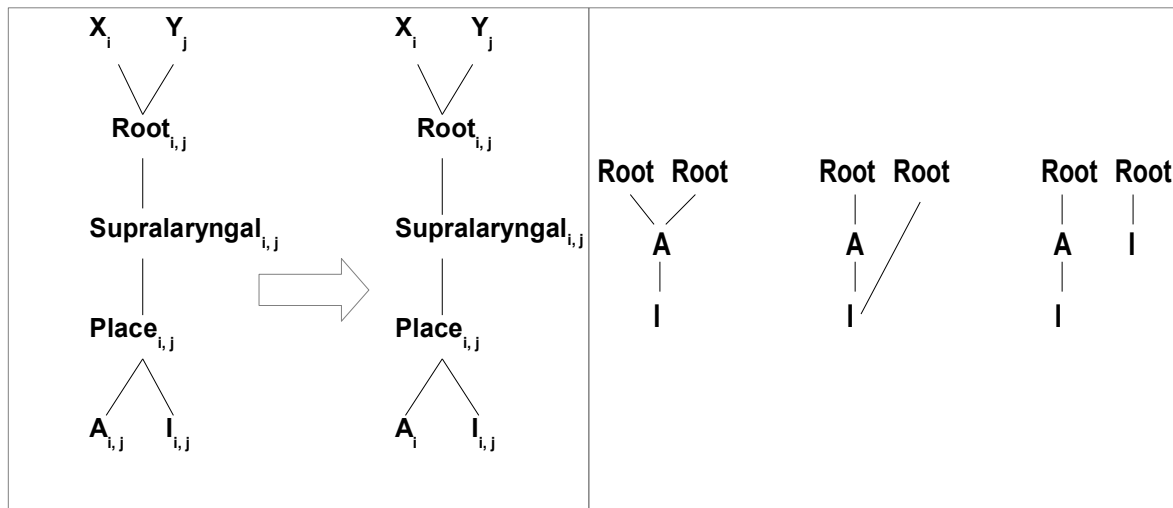


Abbildung 18: Diphongisierung nach Hayes und Selkirk.

Wie man in Abbildung 18 erkennen kann, wird nach der Theorie von Hayes (links) ein Index für die Skelettposition X gelöscht, während nach Selkirk (rechts) das primäre Vokalmerkmal A entkoppelt wird und das sekundäre Merkmal I zum zweiten Wurzelknoten streut.¹⁶

Insofern erlaubt die Two-roots-Theorie von Selkirk einerseits eine homogenere Beschreibung von Länge, indem sie mit den Standard-Prozessen (Delinking, Spreading) nichtlinearer Phonologie beschreibbar ist, andererseits genügt sie Ökonomiekriterien dadurch, dass keine weitere Ebene mit eigenem Gehalt postuliert werden muss. Selkirk selbst nutzte die Theorie vor allem zur Beschreibung von Geminaten. Wie van Oostendorp (2000) gezeigt hat, kann man sie aber auch auf Vokallänge übertragen. Das Sonoritätsprofil von Langvokalen muss so nicht mit Neef top-down durch die Assoziation mit einer C-Position erklärt werden, sondern kann berechnet werden, indem segmentale Ebene und X-Positionen zusammengedacht werden.

Ein weiterer m.E. kaum diskutierter Aspekt ist, dass Sibilifizierungsregeln nur adäquat operieren können, wenn die zu sibilifizierende segmentale Information bereits eine geordnete Sequenz darstellt. Ein bloßes Konglomerat segmentaler Information brächte im Zuge der

¹⁶ Folgende Merkmale werden von Hayes angenommen: /i/ = I; /u/ = U; /e/ = IA; /o/ = UA; /a/ = A. Diese Notation soll hier aber nur zum Verständnis der Beispiele verwendet werden.

Silbifizierung zwar wohlgeformte, aber inadäquate Formen in Bezug zur intendierten Zielform hervor. So würde ohne Verletzung eines auf der CV-Schicht verorteten Sonoritätsprinzips z.B. *Schiff* als *Fisch* realisiert. Eine explizite Spezifikation der Anordnung segmentaler Information wird z.B. von Raimy (2000) in seiner Precedence-Based Phonology zur Erklärung reduplikativer Prozesse angenommen (vgl. Kap. 1.7.1).

Letztlich stellt sich die Frage, in welcher Art Sonorität auf der CV-Ebene repräsentiert werden soll. Irgendeine Form der Merkmalspezifikation müsste auch Neef auf dieser Ebene postulieren, da sie per definitionem als rein kategoriale Repräsentation temporaler Information indifferent gegenüber spezifisch phonematischen Merkmalsgehalt ist, somit nur Länge und Konsonant-Vokal-Differenz repräsentiert. Mit der Hinzunahme weiterer Merkmale wäre nur Redundanz in die Beschreibung eingeführt, da mit der segmentalen Ebene nun zwei sonoritätsklassifizierende Spezifikationen gegeben wären. So bliebe m.E. nur der Weg einer post-phonematischen, phonetischen Klassifikation von Sonorität, mit welcher man aber wichtige sprachsystematische Phänomene nicht erfassen könnte wie z.B. sonoritätsbasierte Clusterreduktionen von Kindern (Ohala 1999).

Die obigen Ausführungen zu Sonorität sollten zwei wesentliche Punkte deutlich gemacht haben:

1. Das Sonoritätsprinzip kann wegen seiner Wirksamkeit in Gebärdensprachen nur als ein kognitives Organisationsprinzip verstanden werden. Als Bestandteil der Universalgrammatik muss es so im Rahmen einer grammatischen Klassifikation betrachtet werden.
2. Die CV-Ebene muss nicht angenommen werden, um Sonorität und das Sonoritätsprinzip zu erfassen. Vielmehr lassen sich bestimmte Phänomene in diesem Zusammenhang ausschließlich mit Rekurs auf die segmentale Ebene beschreiben.

Wie im Folgenden gezeigt wird, ist mit Clements (1990) eine grammatische Klassifikation von Sonorität auf der Basis von phonologischen Merkmalen gegeben. Sie erfüllt insofern die Kriterien, die eine Sonoritätskonzeption nach den Ausführungen dieses Kapitels erfüllen sollte. Zudem macht sie detaillierte theoretische Aussagen über Art und Vorkommen von Sonoritätsbeschränkungen, die durch empirische Evidenz aus der psycho- und neurolinguistischen Forschung gestützt werden.

1.4.2 Sonorität und phonologische Komplexität nach Clements

Das SSP operiert nach der Theorie von Clements nicht auf Oberflächenstrukturen, sondern steuert im Wesentlichen die Silbifizierung von zugrundeliegenden phonologischen Repräsentationen. Spezifischer handelt es sich um den ersten Silbifizierungsprozess, von Clements als Kernsilbifizierung (engl: core syllabification; vgl. ebd.: 289) bezeichnet. Konsonanten, die im Zuge der Kernsilbifizierung selbst nicht assoziiert werden, können in diesem Sinne also extrasilbische Elemente darstellen und werden im grammatischen System anders bewertet bzw. erst in einer Resilbifizierung in die Silbe integriert (vgl. Hall 2000: 248). So gehen solche Konsonanten z.B. im Sanskrit auch nicht in Reduplikationsprozesse ein. Andere Konsonanten werden erst auf einer nachfolgenden Ableitungsebene als Affixe inseriert und folglich ebenso derivational nach dem Stamm silbifiziert (vgl. Fujimura & Lovins 1978).

Da Sonoritätsbeschränkungen im Sinne der skalaren Anordnung der Sonoritätshierarchie auf der Ebene der Kernsilbifizierung wirksam sind, kann Sonorität selbst keine phonetisch messbare Eigenschaft von Lauten sein. Im Gegenteil: Sonorität lässt sich durch die Berechnung von binären Merkmalsmengen, also kategorial, erfassen.

Maßgeblich sind für Clements in der Berechnung die Oberklassenmerkmale, die auch im SPE-Modell verwendet wurden (vgl. Clements 1990: 284). Vier Oberklassen (Obstruent, Nasal, Liquid und Gleitlaut/Vokal) werden von Clements durch drei Oberklassenmerkmale [\pm vocoid], [\pm approximant] und [\pm sonorant] definiert. Das hinzugenommene Merkmal [\pm silbisch] ist für ihn durch die Assoziationen der Segmente mit C-Knoten oder V-Knoten der CV-Ebene motiviert. Segmente, die von einem V-Knoten dominiert werden, sind daher Vokale oder silbische Konsonanten mit dem Merkmal [+silbisch], Segmente, die von C-Knoten dominiert werden sind Konsonanten mit dem Merkmal [-silbisch]. Gleitlaute sind unter der Maßgabe aller Merkmale nichtsilbische Vokale. Wenn Clements über das Merkmal [\pm silbisch] zu dem Schluss kommt, dass es das einzige Oberklassenmerkmal sei, das alle anderen kreuzklassifiziert¹⁷, dann hängt dies damit zusammen, dass dieses Merkmal die Wahrscheinlichkeit einer Klasse als Silbengipfel zu fungieren, in Relation zur Oberklassenzugehörigkeit definiert. Bei höherer Sonorität ist die Assoziation mit einem V-Knoten und damit die Merkmalszuweisung [+silbisch] wahrscheinlicher.

¹⁷ Clements (1990: 295): „[...] it is the only major class feature that cross-classifies all others.“

O	< N	< L	< G	< V	
-	-	-	+/-	+	syllabic
-	-	-	+	+	vocoid
-	-	+	+	+	approximant
-	+	+	+	+	sonorant
1	2	3	4		rank

Tabelle 1: Oberklassenmerkmale und Sonoritätsgrad nach Clements (1990).

Der charakteristische Anstieg und Abfall von Sonorität in der Lautrealisierung innerhalb von Silben, dem Sonoritätszyklus (engl: sonority cycle), lässt sich unter Annahme dieser merkmalsbasierten Hierarchie (1=niedrigste Sonorität bis 4=höchste Sonorität) modellieren:

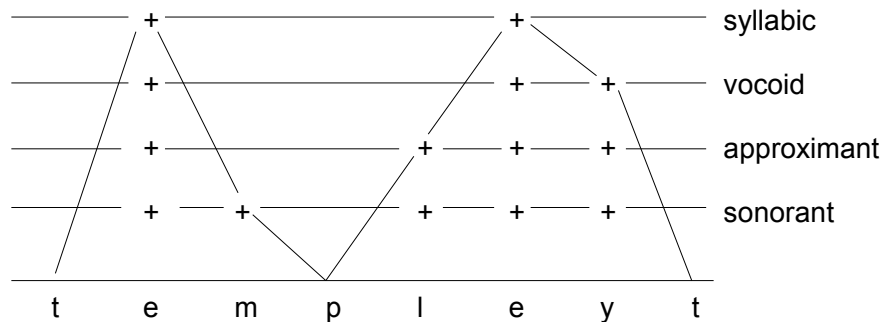


Abbildung 19: Sonoritätszyklus des Wortes template (nach Clements 1990).

Crosslinguistisch kann man nun nach Clements feststellen, dass Silben präferiert werden, dessen Sonoritätsprofil durch einen maximalen Anstieg hin zum Silbennukleus wie nach diesem durch einen minimalen Abfall zum Silbenende charakterisiert sind (ebd.: 301).

Ausgehend von einer solchen optimalen Silbe kann man die Abweichung gegebener Konsonant-Vokal-Verbindungen von dieser berechnen, was äquivalent mit einer Berechnung der phonologischen Komplexität ist. Die Dispersion von der optimalen Silbe wird mit dem Wert D angegeben und auf Basis von sogenannten Demisilben errechnet. Die zugrundeliegende Überlegung ist, dass man eine Silbe in zwei sich überschneidende Demisilben aufteilen kann, zu welchen je der Silbengipfel zugeordnet wird.

So bestünde nach Clements [kran] aus den Demisilben [kra,an], [pa] aus [pa,a] und [ap] aus [a,ap]. Das Sonoritätsprofil der ersten Silbe ist von demjenigen der zweiten unabhängig, somit auch die jeweilige phonologische Komplexität (vgl. ebd.: 303).

Dieser Unabhängigkeit der Demisilben wird mit dem Sonority Dispersion Principle (SDP) Rechnung getragen. Die Distanz oder Dispersion von der jeweiligen Demisilbe zur optimalen Silbe wird mit dem Wert D angegeben. D ist kleiner, wenn die Sonoritätsdistanz maximal ist (so wäre er z.B. bei Obstruent zu Vokal kleiner, da eine Distanz von 4 vorhanden ist, während die Distanz von Liquid zu Vokal nur 2 ist, er somit also höher wäre (vgl. Tab. 2)), und höher, wenn die Sonoritätsdistanz minimal ist (hier gilt der umgekehrte Fall).

Die Bestimmung des Wertes D erfolgt nach Maßgabe folgender Formel, wobei hier d für den Sonoritätsrang und m für die Anzahl der Segmente innerhalb der jeweiligen Demisilbe steht.

$$D = \sum_{i=1}^m 1/d_i^2$$

In Leuninger et al. (2013) ist zur Illustration die Berechnung von D exemplarisch durchgeführt worden:

„Nehmen wir die initiale Demisilbe /pla/ aus /plat/. Für jede Lautkombination dieser Demisilbe wird zunächst der Sonoritätsrang bestimmt, der dem Abstand der Lautklassen innerhalb der Sonoritätsskala [...] entspricht. Wie [...] hervorgeht, haben der Obstruent /p/ und der Liquid /l/ einen Sonoritätsabstand von 2 (Obstruent = 1; Liquid = 3) und somit einen Sonoritätsrang von 2; von /p/ zu /a/ beträgt der Abstand und der Sonoritätsrang 4; und von /l/ zu /a/ wiederum 2. Durch Einsetzen der so ermittelten Sonoritätsränge [...], ergibt sich der folgende Wert für D: $D = (1/2^2) + (1/4^2) + (1/2^2) = 0.56$. Dieser Wert zeigt die Distanz/Dispersion (D) der Demisilbe /pla/ an. Je nach Position der Demisilbe (initial vs. final) wird der Wert einem Komplexitätsrang zugeordnet. Letzterer ergibt sich aus dem optimalen Sonoritätsprofil, in dem ein steiler Anstieg von Sonorität initial (niedriger Wert von D), final hingegen ein flacher Abfall und damit ein höherer Wert von D präferiert ist.“

Unter Zugrundelegung der optimalen Silbe kann durch das Dispersionprinzips bzw. SDP also eine universell gültige Komplexitätsmetrik formuliert werden:

Dispersion Principle

- a. The preferred initial demisyllables minimizes D.
- b. The preferred final demisyllables maximizes D.

Die Komplexitätsgrade für zweigliedrige (CV- und -VC) wie auch für dreigliedrige (CCV- und -VCC) Demisilben sind in Tabelle 2 zusammengefasst.

	D	C
a. Two-member demisyllables		
i. initial:		
OV	0.06	1
NV	0.11	2
LV	0.25	3
GV	1.00	4
ii. final		
VO	0.06	4
VN	0.11	3
VL	0.25	2
VG	1.00	1
b. Three-member demisyllables		
i. initial:		
OLV	0.56	1
ONV, OGV	1.17	2
NLV, NGV	1.36	3
LGV	2.25	4
ii. final		
VLO	0.56	4
VGO, VNO	1.17	3
VLN, VGN	1.36	2
VGL	2.25	1

Tabelle 2: Dispersionswerte (D) und Komplexitätsgrade (C) von Demisilben nach Clements (1990).

Mit Meinschaefer (1998: 48) kann man abschließend hervorheben, dass Sonorität in Clements Konzeption einerseits auf segmentale Merkmale, d.h. auf die Oberklassenmerkmale, zurückgeführt, andererseits in Bezug auf die Silbenstruktur und Silbenposition eines Segments definiert wird, da das Merkmal [silbisch] hinzugenommen werden muss, um alle Sonoritätsbeziehungen zwischen Segmenten einer Silbe zu beschreiben.

Sonorität ist insofern ein Prinzip, „welches die Verbindung zwischen segmentaler Struktur und suprasegmentaler Silbenstruktur herstellt und welches die Integration einer Folge von Segmenten in eine übergeordnete Struktur, d.h. in Silben, ermöglicht.“ (ebd.: 137)

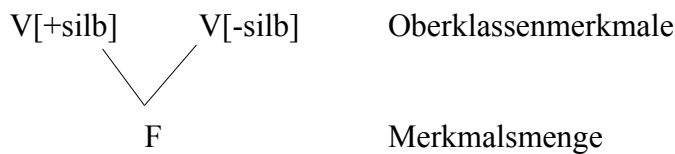
Es müssen zum Abschluss noch einige Erläuterungen in Bezug auf die Diskussion in Kapitel 1.4.1 hinzugefügt werden, d.h., die oben dargestellte Metrik muss zu einer strukturellen Beschreibung in Relation gesetzt werden. Vor allem die Einbettung von Langvokalen und Diphthongen bedarf einer genaueren Spezifikation.

Neef (2004) leitet die Sonoritätsprofile von diesen beiden Elementen aus der CV-Schicht ab. In Kapitel 1.4.1 wurde der Vorschlag gemacht, dass man Sonorität durchaus auf der segmentalen Ebene verorten kann, wenn man für Langvokale die Two-Root Theory von Selkirk (1990) adaptiert. Wie genau kann dies nun aussehen, und inwiefern hängt dies mit Clements Annahmen zusammen?

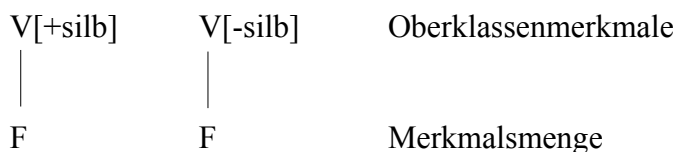
Da in Langvokalen inhärent die Sonorität abfällt, kann dies auf segmentaler Ebene unter der Two-Root Theory nur modelliert werden, indem der zweite Wurzelknoten für das Merkmal [-silbisch] spezifiziert ist. Dieser Wurzelknoten impliziert insofern die Oberklassenmerkmale eines Gleitlautes:

„[...] a syllabic glide is identical to a vowel, or to put it another way, a glide is simply a nonsyllabic vowel.“ (Clements 1990: 294)

Wir erhalten somit folgende Struktur:



Mit diesem Two-Root-Ansatz kann man somit sowohl die vergleichbaren Distributionsmuster von Langvokalen und Vokal-Konsonant-Verbindungen beschreiben als auch den Abfall der Sonoritätskontur innerhalb von nur einem Segment. Weiterhin erhält man eine natürliche strukturelle Affinität zu Diphtongen, wenn man mit Wurzel (1981: 921) und Vater (1992)¹⁸ Diphtonge als Kombinationen aus Vokal und Gleitlaut definiert. Dermaßen kann man mit Selkirk (1990) den Prozess der Diphtongisierung adäquat erfassen (vgl. Kap. 1.4.1: 54). Der wesentliche Unterschied ist hier allerdings, dass zwei disparate Merkmalsmengen mit entsprechend zwei Wurzelknoten in die Silbe integriert werden. Somit stellt sich die Struktur von Diphtongen wie folgt dar:



Diese Analyse vermeidet auch Widersprüche, wie Neef sie zu Recht für bestimmte Ansätze benennt. Ein Beispiel soll dies illustrieren. Hall (2000: 29) sieht Diphtonge als eine Abfolge von zwei Vokalen an, wobei fallende darunter ebenso verstanden werden. Die zweite Hälfte der Vokalabfolge ist nach ihm für [-silbisch] spezifiziert, ohne dass hier ein Gleitlaut angenommen wird. In einem X-Ebenen-Modell, das die Länge mit zwei disparaten X-

¹⁸ „Ein komplexer Nukleus kann aus Vokal + Glide, d.h. aus einem Diphtong, bestehen. Dazu gehören die durch /r/-Vokalisierung entstehenden Diphtonge [...].“ Vater (1992: 126)

Positionen klassifizieren kann, würde diese Analyse allerdings ein Sonoritätsplateau implizieren. In Abb. 20 ist dies am Beispiel der Entwicklung des protogermanischen Diphthongs [ai] (bain) zu [a:] (ba:n) im Altenglischen (ebd.: 258) illustriert:

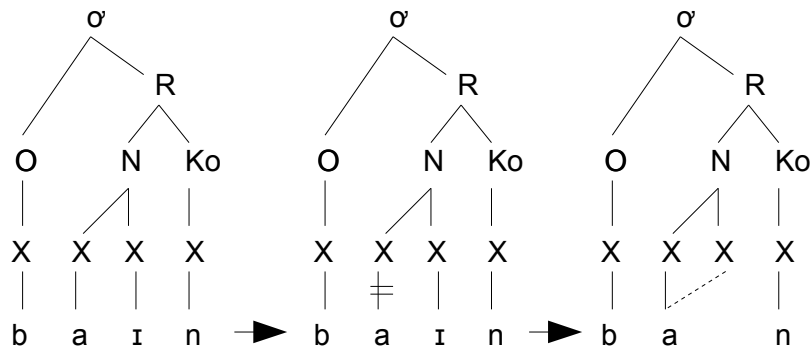


Abbildung 20: Entwicklung des protogermanischen Diphthongs im Altenglischen.

Stellt das zweite, nichtsilbische Element des Diphthongs einen Gleitlaut dar, so haben wir ebenso zwei X-Positionen, allerdings ist kein Sonoritätsplateau vorhanden. Zudem erhalten wir eine oberflächennahe Interpretation des Sonoritätsabfalls. Auch wenn zugrundeliegend Langvokale und Diphthonge bezogen auf die Sonoritätsbeschränkungen gleich behandelt werden, so wirkt sich bei der phonetischen Realisierung die segmentale Ebene aus. Der Abfall der Sonorität wird bei Langvokalen phonetisch letztlich über ein segmentales Element, bei Diphthongen über zwei segmentale Elemente mit jeweils einem Wurzelknoten realisiert. D.h., der Abfall kann bei einem Langvokal inhärent nicht so groß sein, wie dann, wenn er sich zeitlich über zwei Segmente distribuiert, auch wenn die Dauer letztlich gleich ist.

Diese strukturellen Erweiterungen von Clements Theorie sind in Anbetracht der Kritik von Neef notwendig, da Clements selbst eine CV-Schicht angenommen hat und den Aspekt der Vokallänge nicht ausführlich in der Sonoritätskonzeption ausgeführt hat. Für die Berechnung der Komplexität von Langvokalen wie für Diphthonge im Deutschen nimmt er allerdings eine VC-Abfolge an, deren Aufteilung in Demisilben und Komplexitätsberechnung analog zu anderen VC-Abfolgen (s.o.) erfolgt (vgl. ebd.: 327 u. 328; Fußnote 19). In Abbildung 21 ist abschließend die strukturelle wie auch die Demisilbenaufteilung von Langvokalen und Diphthongen illustriert (für eine OV-VN-Abfolge wie z.B. mit Langvokal *Bahn* und Diphthong *Baum*).

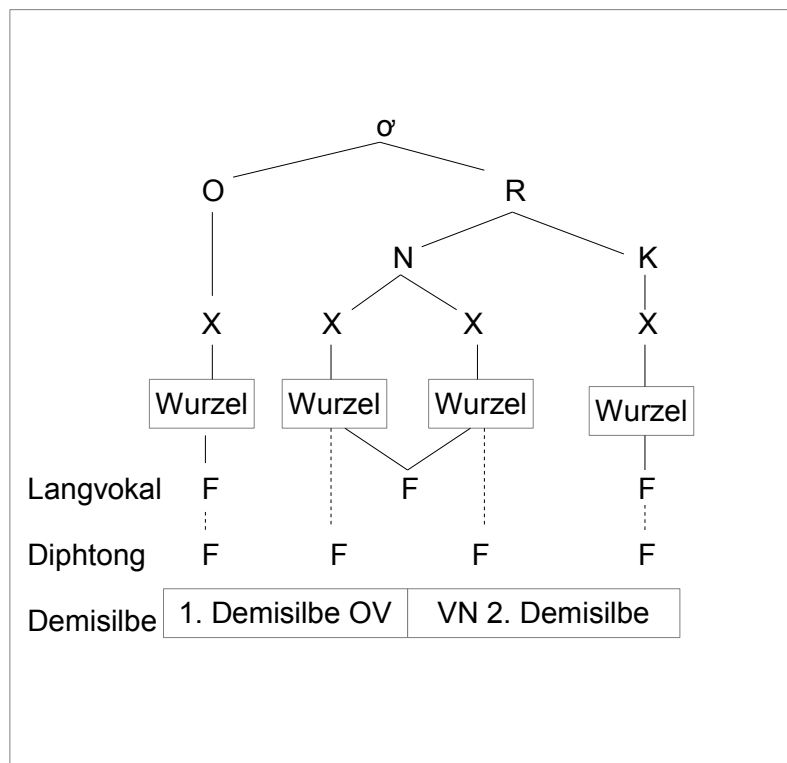


Abbildung 21: Struktur- und Demisilbenangabe für Langvokale und Diphthonge in der Two-Root Theory.

Ein Aspekt, der bisher noch nicht genannt wurde, ist die interne Organisation von Segmenten. Entgegen den Annahmen der linearen Phonologie, die ein Phonem als ungeordnete Bündel binärer Merkmale beschreibt, führte die Beobachtung, dass Merkmale sich zu funktionalen Einheiten gruppieren lassen¹⁹, zu der Entwicklung diverser Merkmalsgeometrien. Benannte funktionale Einheiten lassen sich hierarchisch ordnen (vgl. Abb. 22), so dass höhere Knoten niedrigere implizieren und entsprechend Regelanwendungen gemäß der Hierarchie mit solchen phonologischen Konstituenten operieren. Clements & Hume (1995: 181) fassen dies in folgendem Prinzip zusammen:

- (A) Phonological rules perform single operations only.

Eine Regel, welche die Konstituente C beinhaltet, wird auch d, e, f und g in ihrem Geltungsbereich implizieren, während keine Regel auf c, d und e alleine angewendet werden kann, da diese keine Konstituente bilden (vgl.: ebd.).

¹⁹ So werden z.B. im Englischen die Nasale der Präfixe *syn-* und *con-* vor labialen als [m], vor alveolaren als [n] und vor velaren als [ŋ] realisiert. D.h., die Assimilation betrifft den Artikulationsort als eine funktionale Einheit (vgl. Clements & Hume 1995: 181).

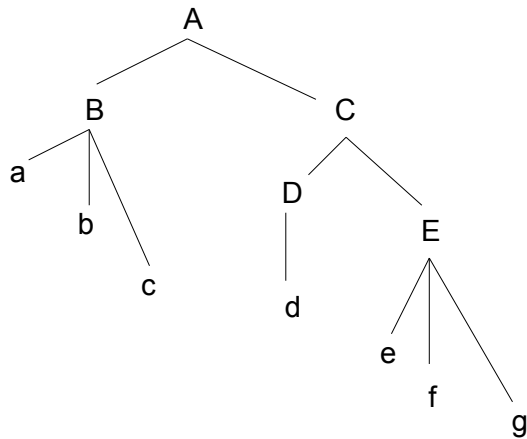


Abbildung 22: Grundaufbau einer phonologischen Merkmalsgeometrie.

Die hierarchische Organisation der Merkmale, ihre Platzierungen auf einzelnen Ebenen und ihre Gruppierungen zu Konstituenten, ist cross-linguistisch uniform:

(B) Feature organization is universally determined.

Prinzip B beinhaltet nach Clements & Hume ebenfalls, dass auf allen Ableitungsstufen einer phonologischen Repräsentation die grundlegende Merkmalsgeometrie projiziert wird und so als Wohlgeformtheitsbeschränkung wirksam ist (ebd.: 182). Aus der Fülle der existierenden Modellvorschläge für Merkmalsgeometrien (z.B. Clements 1985; MacCarthy 1988; Sagey 1986) soll hier das Modell von Clements & Hume (1995) verwendet werden, da es zum einen hinsichtlich der Wurzelknotenmerkmale mit der in dieser Arbeit zu überprüfenden Sonoritätskonzeption von Clements (1990) kompatibel ist, zum anderen im direkten Vergleich mit anderen aktuellen Geometrien eine wohlbemessene Differenziertheit und empirische Fundierung hat.

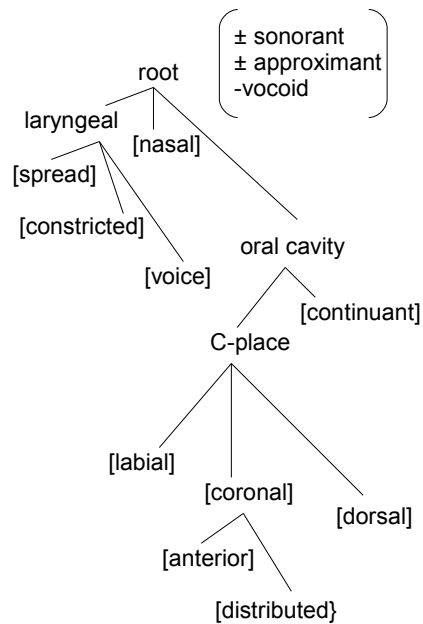


Abbildung 23: Merkmalsgeometrie für Konsonanten nach Clements & Hume (1995).

Auf die einzelnen Argumentationen für die Merkmale und Anordnungen kann und soll im Rahmen dieser Arbeit nicht spezifisch eingegangen werden, da hier vornehmlich die Oberklassenmerkmale bzw. Wurzelknotenmerkmale für die Thematik phonologischer Komplexität eine Rolle spielen, ansonsten aber ausschließlich der prinzipielle hierarchische Aufbau einer segmentalen Repräsentation und die Gliederung in phonologische Konstituenten von argumentativer Wichtigkeit sind. Die Rolle der phonologischen Komplexität in der Sprachproduktion soll nun im nächsten Kapitel an einschlägiger Evidenz aus der Aphasieologie dargestellt werden. Es wird weiter zu zeigen sein, dass der Ansatz einer dreidimensionalen Phonologie wie auch die Komplexitätsmetrik von Clements eine adäquate Passung zu Modellen der Sprachproduktion besitzen.

1.5 Sprachproduktionsmodelle und 3D-Phonologie

1.5.1 Das Sprachproduktionsmodell von Garrett

Das aus Versprecherdaten abgeleitete Sprachproduktionsmodell von Garrett (1975, 1980) kann als Ausgangsmodell für die Diskussion um die Art und Weise der Planung und Berechnung phonologischer Repräsentationen gelten. Es wird als autonomes Modell klassifiziert, wobei in einem solchen eine unabhängige Informationsverarbeitung in einzelnen Einheiten eines (kognitiven) Systems stattfindet. Bevor in einem solchen System weitere Einheiten aktiviert werden, müssen früher aktivierte Einheiten in ihrer Informationsverarbeitung erst zu einem Abschluss gekommen sein. In dem Modell von Garrett werden solche Einheiten als Ebenen dargestellt, wobei diese für diskrete Verarbeitungszeitpunkte von Operationen mit einem spezifischen repräsentationalem Format stehen (vgl. auch Wiegand 1996: 26). Der Informationstransfer erfolgt in diesem Modell streng seriell und top-down, was hier nichts anderes bedeutet, als dass von der obersten Ebene (Botschaftsebene), der intendierten Botschaft, zur untersten Ebene (Phonetische Ebene), den Instruktionen an den artikulatorischen Apparat, fortgeschritten wird.

Das Sprachproduktionsmodell von Garrett beinhaltet eine funktionale und eine positionale Ebene der Satzplanung. Auf der funktionalen Ebene findet der Zugriff auf Lemmata statt (in Abb. 24 unter 'Bedeutung'). Ein Lemma umfasst die syntaktische Information eines Wortes, also Wortklasse, Subkategorisierungsrahmen wie auch flexionsmorphologische Kategorien, wie Person, Kasus, Numerus, Tempus, Modus, Determiniertheit und Komparation. Auf der funktionalen Ebene wird also eine propositionale Struktur des Satzes erstellt und syntaktische Relationen wie auch grammatische Funktionen verarbeitet. Es handelt sich auf dieser Ebene um eine Repräsentation eines abstrakten Satzrahmens, in dem noch keine Positionen für einzelne Elemente fixiert sind.

Erst auf der positionalen Ebene findet eine solche Fixierung durch die Abbildung einer phonologischen Form auf den abstrakten Satzrahmen statt (in Abb. 24 unter 'Form'). Auf dieser Ebene werden die spezifischen Lexem-Informationen abgerufen, d.h. die phonologisch-phonetischen (und graphematischen) Strukturen mit ihrer silbischen Gliederung und Akzentplatzierung im Wort. Erst diese repräsentationale Gesamtstruktur wird abschließend auf die phonetische Ebene für konkrete Instruktionen an den Artikulationsapparat abgebildet.

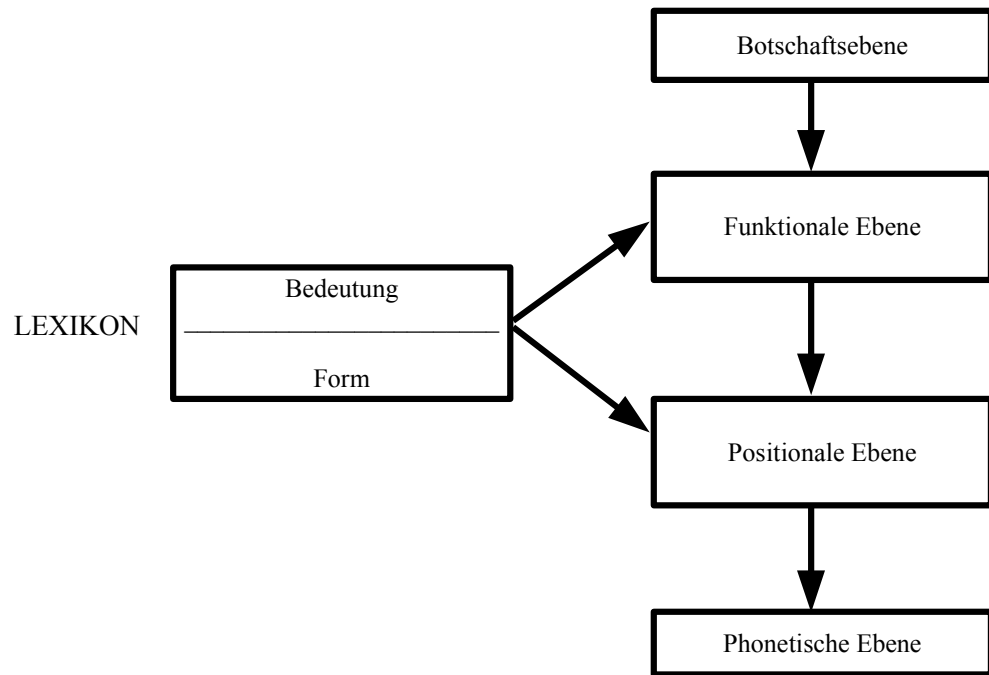


Abbildung 24: Das Sprachproduktionsmodell von Garrett (1975, 1980).

1.5.2 Die Berechnung von Silbenpositionen und segmentaler Information

Poulisse (1999) hat für Versprecher die wichtigsten Forschungsergebnisse zusammengefasst, wobei an dieser Stelle nur diejenigen dargestellt werden soll, die sich auf die positionale, d.h. sublexikalische Ebene beziehen.

Ergebnis	Quellen
Einzelne Segmente sind die wichtigsten Einheiten in der Sprachproduktion	Boomer & Laver 1968; Dell 1986, 1988; Fromkin 1971; Nootboom 1969; Shattuck-Hufnagel 1982
Antizipationen sind häufiger als Perseverationen	Boomer & Laver 1968; Cohen 1966; Fromkin 1971; Nootboom 1969; Van den Broecke & Goldstein 1980
Vertauschungen sind sehr selten	Cohen 1966; Dell 1986; Nootboom 1969
Wenn phonologische Einheiten den Platz wechseln, behalten sie die jeweilige Silbenposition bei (z.B. Onset zu Onset)	Boomer & Laver 1968; Fromkin 1971; García-Albea et al. 1989; Garrett 1975; Laubstein 1987; MacKay 1970; Nootboom 1969; Shattuck-Hufnagel 1982
Wort- oder silbeninitiale Konsonanten sind häufiger in Versprechern zu finden als finale.	Garrett 1975, 1980; Laubstein 1987; MacKay 1970; Shattuck-Hufnagel 1987; Stemberger 1985; Van den Broecke & Goldstein 1980
Fehler finden häufiger in betonten als in unbetonten Silben statt.	Boomer & Laver 1968; Fromkin 1971; Garrett 1975; Nootboom 1969
Wenn zwei Segmente vertauscht werden, besteht häufig eine phonetische Ähnlichkeit	Fromkin 1971; García-Albea et al. 1989; Garrett 1975; Levitt & Healy 1985; MacKay 1970; Nootboom 1969; Shattuck-Hufnagel 1982; Shattuck-Hufnagel & Klatt 1979; Van den Broecke & Goldstein 1980
Zwei Segmente werden eher vertauscht, wenn ihnen identische Phoneme vorangehen oder folgen	Dell 1984; Dell & Reich 1980; Garrett 1975; MacKay 1970; Nootboom 1969
Versprecher resultieren nie in Sequenzen von Phonemen, die in der jeweiligen Sprache nicht möglich sind	Boomer & Laver 1968; Dell 1986, 1988; Fromkin 1971; Garrett 1980; MacKay 1970; Van den Broecke & Goldstein 1980; Wells 1951

Tabelle 3: Wesentliche Ergebnisse der Versprecherforschung zusammengefasst nach Poulisse (1999).

Ausgehend von dieser umfassenden Listung von Forschungsergebnissen kann nach der adäquaten Modellierung des Sprachproduktionsprozesses auf sublexikalischer Ebene gefragt werden. In Shattuck-Hufnagel (1979, 1983) wird diese Ebene und mit ihr mögliche operative Komponenten genauer spezifiziert. Auf Grundlage von Phonemfehlern bei Versprechern entwirft sie das sogenannte *Slots-and-fillers-Modell*. Silbenpositionen (slots) und Segmente (fillers) bilden in diesem Modell zwei unabhängig voneinander berechnete Repräsentationen. Die Repräsentation der Silbe und ihrer Untergliederung in verschiedene Positionen (Onset, Nukleus, Reim (Koda)) stellt den Rahmen bereit, in den die separat abgerufenen Segmente

des lexikalischen Eintrags inseriert werden. Diese Segmente des jeweiligen lexikalischen Eintrags werden kurzzeitig in einem Zwischenspeicher (Buffer) gehalten. Zwei Komponenten werden nun aktiv. Zum einen der Scan Copier, der die Segmente in entsprechende Silbenpositionen kopiert, und zum anderen der Checkoff Monitor, der als Kontrollinstanz bereits kopierte Segmente erkennt und sie nach Identifizierung aus dem Buffer löscht. Mit diesem Modell kann man die „Silbentreue“ von Lautvertauschungen erfassen, da die vertauschten Einheiten die jeweilige Silbenposition der anderen besetzen, wie z.B. in dem Versprecher *C-Moss-Melle* statt *C-Moll-Messe*.

Nicht klar und genauer spezifiziert wird bei Shattuck-Hufnagel allerdings, wie genau der Integrationsprozess vonstattengeht:

„[...] after the generation of a framework of slots, the scan copier examines the set of candidate segments to find the target segment appropriate for the first slot, selects the target, and copies it into the slot.“ (Shattuck-Hufnagel 1983: 132)

Woran aber erkennt der Kopiermechanismus die angemessene Position für ein unabhängig generiertes Phonem? Da der Kopiervorgang explizit als ein Prozess der seriellen Anordnung bezeichnet wird (ebd.: 133), kann man wohl von einem ungeordneten Kandidatenset von Phonemen ausgehen, was die gesamte Zuordnungsaufgabe noch undurchsichtiger macht. Bei aller explanatorischen Stärke dieses Modells bleibt es an dieser Stelle also außerordentlich ungenau. Levelt (1992: 10) bemerkt genau dieses Problem kritisch an: Wenn Silbenposition und Phoneme aus einem phonologischen Code im Lexikon abgeleitet werden, warum erst beide erstellen, um sie dann wieder zusammensetzen? Was er vor allem betont, ist, dass keine Lexeme, sondern phonologische Wörter silbifiziert werden. So wird z.B. *gave it him* als gesamtes phonologisches Wort in der Form /ga-vi-tim/ silbifiziert, so dass alle Lexemgrenzen verletzt werden (vgl. ebd.: 11). Was in den Aufbau des phonologischen Wortes eingeht, so Levelt, sind die abstrakten und auf früherer Stufe gespeicherten Informationen für die Anzahl der Silben, ihre Betonung und ihre Gewichtung (in Moren).²⁰ An diesen metrischen Rahmen werden nun die Phoneme in der Silbifizierung nach Maßgabe des Sonoritätsprinzips von links nach rechts assoziiert (vgl. auch Meyer & Schriefers 1991). U.a. lassen sich so Versprecher wie *peel like flying* ← *feel like playing* erklären, da hier wahrscheinlich zwei phonologische Wörter, nämlich [*feelike*]_o und [*playing*]_o, geplant waren und somit der häufig von Fehlleistungen betroffene Wortonset (vgl. Shattuck-Hufnagel 1987) nicht bei dem auf dieser

²⁰ Dies wird u.a. mit Bezug auf Tip-of-the-tongue-Zustände (TOTs) begründet, da Personen in einem solchen zwar keinen Zugang zur phonematischen Ausbuchstabierung haben, dennoch aber oft Silbenanzahl und Betonung abrufen können (für einen Überblick s. Sauer (2015)).

Ebene liegendem medialem *like* gegeben ist. Levelt betont insofern in Bezug auf Shattuck-Hufnagels Slots-and-fillers-Modell:

„It is therefore not generally the case that phonological encoding consists of filling pre-existing lexical skeletons. Rather, new phonological word frames are constructed, dependent on the context of utterance. It is these newly constructed frames that have to be filled with segmental materials. Hence, there is no paradox.“ (ebd.: 11)

Levelts Analyse auf Basis des phonologischen Wortes ist an dieser Stelle zuzustimmen. Allerdings findet sich auch bei ihm keine eindeutige Klärung zur Frage der Linearität lexikalischer Einträge sowie zur modelltheoretischen Deutung der verschiedenen Fehlleistungstypen. Auch wenn die Wortonsetpräferenz in dem, m.E. grundsätzlich richtigen, Silbifizierungsansatz erfasst wird, so wird nicht deutlich, wie genau die Entstehung der anderen Fehlleistungsmuster in diesen einzuordnen sind. Durch den Silbifizierungsmechanismus, der auf einer linearen Segmentmenge strikt linear von links nach rechts operiert, werden Scan-Copier und Check-Off-Monitor im Grunde obsolet. Schon in Levelt (1989: 351) wird deshalb auf die Komplementarität des Slots-and-filler-Modells mit konnektionistischen Modellen hingewiesen. Spätere Entwicklungen des WEAVER (Word-form Encoding by Activation and VERification)-Modells (Levelt et al. 1999) zeigen aber deutlich, dass Fehlleistungen nicht mehr im Rahmen einer klassischen Architektur mit symbolischer Verarbeitung erklärt werden sollen, sondern auf der Ebene der Implementierung, d.h. durch die Aktivierung informationell spezifizierter Netzwerkknoten (vgl. Kap. 1.2.3). Da der konnektionistische Ansatz in dieser Arbeit als nicht zielführend und fruchtbar für die psycholinguistische Analyse erachtet wird²¹, soll im Zuge dieses Kapitels ein Ansatz mit einer detaillierten Silbifizierungsprozedur vorgeschlagen werden.

Bezogen auf das Slots-and-fillers-Modell von Shattuck-Hufnagel sieht Buckingham (1990) in den Sprachdaten von Jargon-Aphasikern hinreichende Evidenz für eine weitere Produktionskomponente auf sub-lexikalischer Ebene. Nach Schädigung des posterioren Temporallappens produzieren diese Patienten häufig Neologismen an Stelle von existierenden Nomen, Adjektiven oder Verben. Butterworth (1979) schließt aus der Beobachtung, dass vor der Produktion von Neologismen Pausen von bis zu 500 ms im Redefluss auftreten, dass der Zugang zum Lexikon gestört ist. Vor allem sogenannte abstruse Neologismen von Jargon-Aphasikern zeigen keinerlei phonematischen Bezug zu Zielwörtern der jeweiligen Sprache,

²¹ Interessant ist in diesem Kontext, dass Levelt selbst 1991 einen immens kritischen Artikel zum Konnektionismus mit dem Titel „Die konnektionistische Mode“ verfasst hat. Wenn man diesen Titel aufnehmend mit Oscar Wilde sagen möchte „Mode ist jene kurze Zeitspanne, in der das völlig verrückte als normal gilt“, kann man Levelts neuere Arbeiten nun leider nicht mehr ausnehmen.

d.h., ihre Produktion muss durch fehlenden Zugang zum Lexikon initiiert sein. Buckingham (1990) greift in diesem Zusammenhang die Idee eines Zufallsgenerators auf, dessen erste Erwägung ebenfalls auf Butterworth (1979) zurückgeht. Der Zufallsgenerator ist im Slots-and-fillers-Modell auf der Ebene des Buffers verortet, welcher für die Zwischenspeicherung von segmentaler wie auch (Silben-)positionaler Information zuständig ist (*book-keeper* bei Shattuck-Hufnagel, s.o.). Da bei Jargon-Aphasikern eine reine Produktionsstörung vorliegt, kann man im Modell entsprechend ein Input- und Output-Lexikon annehmen. Es ergibt sich somit eine differenzierte Erweiterung des ursprünglichen Garrett-Modells (für eine genauere Beschreibung, s. Dümig & Leuninger 2013):

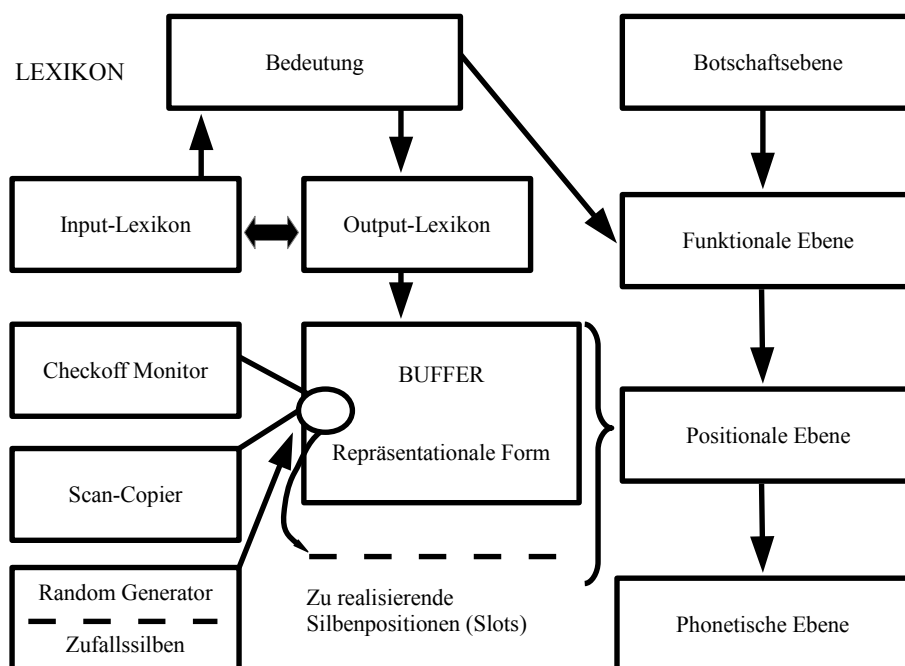


Abbildung 25: Erweitertes Garrett-Modell mit Input- und Output-Lexikon.

Ist der Zugang zum Ziellexikon blockiert, werden vom Zufallsgenerator im Buffer Segmente in die Positionen inseriert und die kompletten Silbenformen in den vorsezifizierten Satzrahmen eingefügt. Von einigen Autoren (z.B. Ellis & Young 1991) wurde der Zufallsgenerator als eine unzulässige Ad-hoc-Annahme kritisiert. Buckingham vertritt hier die Annahme, dass der Zufallsgenerator eine Sprachproduktionskomponente ist, die auch bei

Sprachgesunden vorhanden, aber nicht aktiv ist.²² Wann und wie diese Komponente aktiv ist, wird in späteren Kapiteln eine wesentliche Fragestellung der vorliegenden Arbeit sein.

In Kapitel 1.4.2 wurde herausgearbeitet, dass die Integration von Segmenten in Silbenpositionen über die durch Oberklassenmerkmale definierte Sonorität geschieht. Diese integrative Funktion von Sonorität scheint nach bisherigen Erkenntnissen neuronal fest implementiert und kortikal und/oder subkortikal großflächig repräsentiert, da selbst schwerste Hirnschädigungen und komplette linke Hemisphärektomien die produktive Silbifizierung hinsichtlich Sonorität nicht beeinträchtigen (vgl. Code 1982; Sussman 1984). In welcher Form werden nun aber Sonoritätsbeschränkungen bei den Produktionen des Zufallsgenerators wirksam?

Dies wurde von Christman (1992) an Spontansprachproduktionen und elizitierten Äußerungen von drei Jargonaphasikern untersucht. Produzierte Neologismen (zielbezogene und abstruse) bestehen hier vorwiegend aus initialen CV- und finalen V- und VC-Demisilben. Die C-Position wird initial am häufigsten mit Obstruenten besetzt, final sind übermäßig reine Vokal-Demisilben realisiert. Insgesamt besteht so in den produzierten Neologismen die Tendenz hin zur optimalen CV-Silbe, wie sie Clements definiert. Diese Ergebnisse stützen nach Christman die Anschauung von Sonorität als

- „(1) a hard-wired component of the language system, since its operation was not significantly impaired in phonological systems that have undergone serious impairment;
 - (2) a mediator of phonological construction in all word forms, neologistic or otherwise [...]
- and
- (3) a useful metric in capturing the underlying phonological regularity of words that would otherwise appear to be somewhat randomly constructed [...].” (ebd. 234)

Code & Ball (1994) fanden die gleiche Tendenz zur Bildung optimaler Silben in einem Korpus aus englischen und deutschen nicht-lexikalischen Sprachautomatismen (z.B. [dodidodi]). Das englische Korpus bestand aus 102 Silben, das deutsche aus 119 (vgl. Code 1982; Blanken et al. 1990). In beiden Korpora dominierten initiale OV- und finale V-Demisilben. Vereinfachungen der Komplexität von Silbenstrukturen stellten Romani & Calabrese (1998) bei einem Italienisch sprechenden Broca-Aphasiker, DB, fest. Im Einklang mit den Vorhersagen des Sonority Dispersion Principles wurden bei Substitutionen im Onset Elemente aus weniger sonoranten Lautklassen inseriert. Romani & Calabrese weisen aber

²² „The random generator is merely an alternative way of describing a system of phonological knowledge that all normal speakers possess as part of their cognitive linguistic machinery, and that therefore the mechanism is not created by the lesion at all but rather is simply a normal, albeit underused, capacity that is incorporated at certain points in the speech output of severely anomic Wernicke’s aphasics where content words are sought but unsuccessfully retrieved.” (Buckingham 1990.: 215)

darauf hin, dass im Italienischen keine komplexen Kodas lizenziert sind und nur Liquide und Nasale realisiert werden können. Signifikante Veränderungen können durch diese einzelsprachlichen Bedingungen so nicht gemessen werden. In einer neueren Studie von Stenneken et al. (2005) wurden die abstrusen Neologismen in Spontansprachdaten des deutschsprachigen Patienten KP mit diagnostizierter Jargon-Aphasie untersucht. Die erhobenen Daten wurden mit Daten aus der deutschen CELEX-Datenbank (Baayen et al. 1993) verglichen, wobei produzierte Demisilben in einer Type-Token-Analyse daraufhin untersucht wurden, ob sie eine Tendenz zur optimalen Silbe zeigen. Auch hier waren die Demisilbentypen OV initial und V final die häufigsten Formen. Auch bei komplexeren Demisilbentypen konnte eine Tendenz zur Minimierung phonologischer Komplexität beobachtet werden (ebd.: 289).

Die Produktionen der Jargon-Aphasiker offenbaren also die Integrationsfunktion von Sonorität in einem Zustand, in dem noch keine Adjustierung an ein zielsprachliches Phoneminventar stattgefunden hat bzw. ein solches nicht zur Verfügung steht. In einem solchen Zustand ist eine Verarbeitungspräferenz für nicht-komplexe Demisilben gegeben. Dass die Verarbeitungsleistung des Buffers eng an die Funktion des Dispersionsprinzips gekoppelt ist, lässt sich auch bei anderen Formen der Aphasie beobachten. So fanden Dümig, Konradi & Leuninger (2016) bei einem Nachsprechtest mit zwölf Aphasikern, dass die nach phonologischer Komplexität kontrollierten Items bei zweipositionalen Demisilben nicht signifikant in ihrer phonologischen Komplexität verändert wurden, bei dreipositionalen Demisilben allerdings eine signifikante Reduktion derselben stattfand. Da nur eine geringe Anzahl von Clusterreduktionen vorhanden waren, konnte dies nicht einfach auf Probleme mit der Länge der Items zurückzuführen sein. Da bei zweipositionalen Demisilben keine Reduktion stattfand, muss man ein genuines Problem mit phonologischer Komplexität auch ausschließen. Erklärbar werden die Muster nur, wenn man das gesamte Zusammenspiel von Buffer, Silbifizierung und phonologischer Komplexität betrachtet. Silbifizierung ist nach einer solchen psycholinguistischen Interpretation ein Prozess, der innerhalb des Buffers Silbenpositionen mit einer bereits geordneten Menge von Segmenten assoziiert. Dieser Assoziationsvorgang wird wesentlich von dem Merkmalsgehalt der Segmente, d.h., durch das Dispersionsprinzip beschränkt. Silbifizierung findet nun vom Vokal aus zu den Silbenrändern hin statt, wobei bei mehr X-Positionen auch eine erhöhte Verarbeitungszeit gegeben ist. Bei einer verminderten Buffer-Kapazität kann der Gehalt der X-Positionen an den Silbenrändern nicht mehr adäquat repräsentiert werden und es werden, ähnlich wie bei der Jargon-Aphasie,

Demisilben mit geringerer phonologischer Komplexität als Default produziert. Solch feinkörnige Interaktionen auf sublexikalischer Ebene, die die Wirkmächtigkeit des Dispersionsprinzips offenbaren, bedürfen einer weitaus größeren Aufmerksamkeit in der Aphasiologie wie auch in der Spracherwerbsforschung. Es ergibt sich weiterhin eine andere Interpretation für die Versprecher auf sublexikalischer Ebene.

In Anlehnung an die Darstellung der Silbifizierung in Kapitel 1.4.2 soll hier nun versucht werden, die gegebene Versprecherevidenz für die Sprachplanung auf sublexikalischer Ebene (vgl. Poulisse 1999) in die besagte Konzeption einzuordnen. Von Levelt (1989, 1992) werden folgende grundsätzliche Annahmen übernommen:

- a) die Domäne der Silbifizierung ist das phonologische Wort.
- b) die Silbenanzahl und Betonung ist für Lexeme abstrakt auf einer vorherigen Stufe abgespeichert.
- c) Es findet eine Online-Silbifizierung statt, in der Position und segmentaler Gehalt nicht voneinander getrennt berechnet wird.

Geht man von einem Silbifizierungsprozess aus, der auf einer bereits linearisierten Merkmalsmenge operiert (vgl. Kap. 1.7.1), folgt zuerst das von Goldsmith (1990) beschriebene Splitting, in dem über die Identifikation des Merkmals [silbisch] die Assoziation zur X-Ebene hergestellt wird. Nach der Applikation von Splitting als Subprozess besteht simultan eine positionale wie auch phonematische Spezifikation, d.h., Silbenpositionen wie die Konsonant-Vokal-Distinktion sind gegeben. Diese frühe Planungsebene spiegelt letztlich den Unit-Similarity Constraint (Shattuck-Hufnagel 1979) wider, der besagt, dass die basalen Positionen und Kategorien nicht mehr verändert werden können. Dies bedeutet aber auch, dass Hinzufügungen wie Auslassungen entsprechend auf dieser Berechnungsstufe zu verorten sind. Jede Hinzufügung oder Auslassung schießt eine Berechnung des jeweils neu entstehenden Sonoritätsprofils ein und da keine Verletzungen desselben stattfinden, müssen diese Fehlleistungen zuvor entstanden sein.

Diese weitere wichtige Beschränkung bei Fehlleistungen, nämlich dass niemals Sequenzen von Phonemen erzeugt werden, die in der jeweiligen Zielsprache nicht möglich sind, wird also erst auf der nächsten Ebene gültig. Auf dieser zweiten Ebene wird folglich die Information der Oberklassenmerkmale (und mit ihr der Demisilbenkomplexität) berechnet und assoziiert. Die Demisilbenkomplexität bei Antizipationen, Perseverationen und Substitutionen wird nicht signifikant verändert, wie die Untersuchung von Leuninger et al. (2013) zeigt. Dies spricht dafür, dass die Oberklassenmerkmale als Maßgabe dieser Komplexität erhalten bleiben und auf einer früheren Stufe als die drei benannten Versprechertypen entstehen

müssen. Was allerdings auf dieser Stufe stattfinden muss, ist die Fehlplanung, die zur Entstehung von Vertauschungen führt. Vertauschungen treten selten auf.²³ Dies hat m.E. seine Ursache darin, weil sie relativ gut repariert werden können. Zwei Charakteristika von Vertauschungen sind die Ähnlichkeit der betroffenen Phoneme (meist unterscheiden sie sich in nur einem Merkmal, welches am häufigsten das Merkmal [Artikulationsort] ist) sowie die Ähnlichkeit des phonologischen Kontextes (der folgende Vokal ist meist gleich). Da Vertauschungen nicht zu Verletzungen der Phonotaktik führen dürfen, muss die sie verursachende Fehlplanung also bereits auf Ebene 2 stattfinden. D.h., bei der Silbifizierungsprozedur werden in der linearen Anordnung die Oberklassenmerkmale des sequenziell späteren Phonems zuerst assoziiert. Dies muss nicht aufgrund einer höheren Aktivierung geschehen (Levelt 1989: 346), sondern vielmehr aufgrund symbolischer Identität (Identität der Oberklassenmerkmale) und Struktursensitivität (abstrakt sind die Folgen AB und BA gegeben) die für eine klassische Verarbeitung kennzeichnend sind (vgl. Kap. 1.2.3). Aber da auch hier zuerst nur die Oberklassenmerkmale vertauscht assoziiert werden, so muss auf Ebene 3 der segmentale Rest assoziiert werden. Genau diese zweistufige Entstehung von Vertauschungen, die sie von anderen Versprechertypen unterscheidet, ist nach meiner Hypothese dafür verantwortlich, dass sie am Seltensten auftreten bzw. in Korpora nicht sichtbar werden. Auslassungen und Hinzufügungen auf Ebene 1 sind klar identifizierbar. Fehlplanungen auf Ebene 2 können hingegen auf Ebene 3 noch korrigiert werden. Es besteht auf Ebene 3 die Möglichkeit, dass der segmentale Rest richtig assoziiert wird, d.h., eine Möglichkeit zur Reparatur ist auf dieser Stufe noch gegeben.

Reparaturen sind also bei den letzten beiden Versprechertypen, Antizipation und Perseveration maßgeblich. Nach der Studie von Leuninger et al. (2013) werden Oberklassenmerkmale bzw. phonologische Komplexität nicht signifikant bei der Entstehung von Antizipationen und Perseverationen verändert. D.h., auf Ebene 3 entstehen Antizipationen und Perseverationen durch Doppelassoziation eines segmentalen Rests mit zwei Oberklassenmerkmalsmengen (was man auch mit dem klassischen Terminus Spreading bezeichnen kann). Dass Antizipationen häufiger vorkommen als Perseverationen, lässt sich analog zur Studie von Leuninger et al. (2013) wiederum in Bezug auf den Silbifizierungsvorgang erklären und seiner Verortung im Buffer erklären. Zuerst abgerufene Information wird in Zeitmaßstäben sprachlicher Berechnung der Logik noch länger online gehalten. Findet eine Fehlplanung früher statt, ist

²³ Nach vielen Studien ist dies der Fall. Im Frankfurter Versprecherkorpus mit einer Anzahl von 6970 Versprechern nehmen sie allerdings mit 15 % relativ frequent (wobei hier nicht zwischen Wort- und Phonemvertauschungen unterschieden wird) (vgl. Leuninger 1993, 2005).

der Zeitraum für die Registration derselben größer und eine Reparatur durch Spreading von einer nachfolgenden Merkmalsmenge kann somit in einem größeren Zeitfenster geleistet werden.

Die Vorteile des hier dargestellten Ansatzes als Ergänzung zu Shattuck-Hufnagel (1979, 1983) und Levelt (1989, 1992) sind zusammenfassend diese:

1. Mit Levelt wird der Online-Silbifizierungsprozess als zentral angesehen.
2. Postulierte, sprachexterne Mechanismen wie der Scan-Copier und der Check-off-Monitor fallen durch die zentrale Silbifizierungsprozedur als ursächliche Komponenten für die Erklärung von Versprechertypen weg.
3. Die Konzeption, dass segmentale Information in der Verarbeitung in drei Stufen top down an die X-Ebene assoziiert wird, ist kompatibel mit merkmalsgeometrischen Theorien, ohne dass zusätzliche Mechanismen postuliert werden müssen.

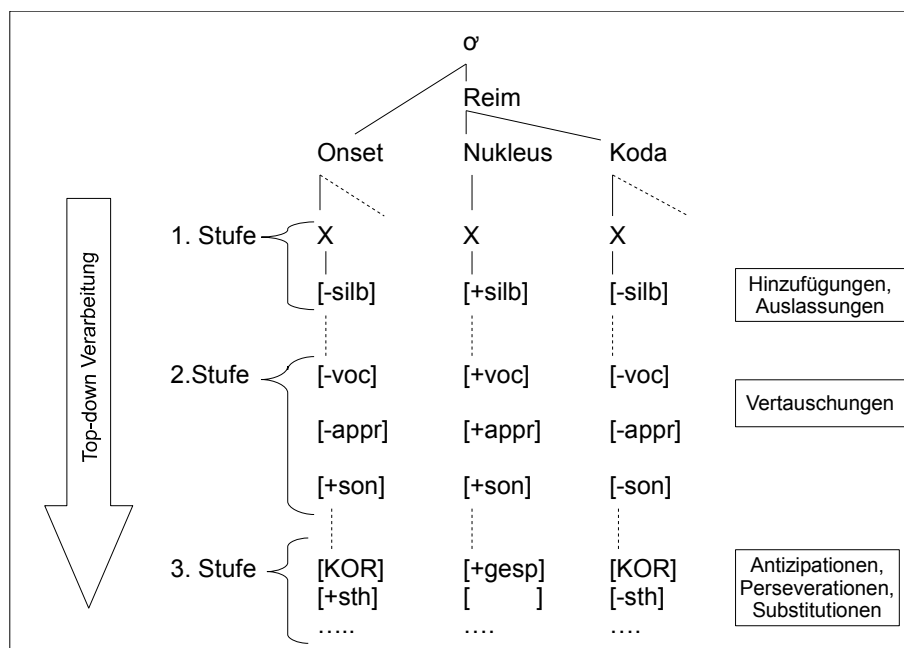


Abbildung 26: Dreistufig konsekutive Top-down-Berechnung von phonologischer Information mit Fehlplanungstypen.

Goldsmiths Annahme eines Splittings, in der nur das Merkmal [silbisch] auf einer ersten Stufe abgespalten assoziiert wird, kann also erweitert werden zu der Hypothese, dass eine dreifache Abstufung von Splitting-Prozessen stattfindet, in der das Merkmal [silbisch], die Oberklassenmerkmale und letztlich spezifische Artikulationsort- und

Artikulationsartmerkmale (die hier als segmentaler Rest bezeichnet werden sollen) konsekutiv assoziiert werden. Dies soll in folgender Hypothese zusammengefasst werden:

Splitting-Hypothese

Das Merkmal [silbisch], die Oberklassenmerkmale und die jeweils spezifischen Artikulationsort- und Artikulationsartmerkmale einer segmentalen Menge werden zeitlich konsekutiv in drei Stufen an definierte X-Positionen assoziiert.

Splitting ist insofern ein Prozess, der einerseits für die Sprachproduktion allgemein gilt, andererseits, wie im Rahmen dieser Arbeit weiter gezeigt werden soll, wesentlichen Stufen der phonologischen Entwicklung im Spracherwerb unterliegt. Analog zu dem Meilenstein-Modell von Tracy (1995) könnte man nach Maßgabe der Splitting-Hypothese so einen Merkmalsausbau im sehr frühen Spracherwerb in Bezug auf die Merkmalsgeometrie erfassen. Merkmale der noch nicht erreichten Integrationsstufe wären für die Hauptberechnungsprozesse der jeweiligen Stufe nicht relevant. Anders formuliert wären auf der 1. Integrationsstufe Silbenpositionen an sich relevant, auf der 2. Integrationsstufe die Anordnung von Phonemen nach Maßgabe von Sonorität und erst auf der 3. Integrationsstufe die Spezifikationen, die für die Zieladäquatheit von Phonemen relevant sind (vgl. Abb. 27).

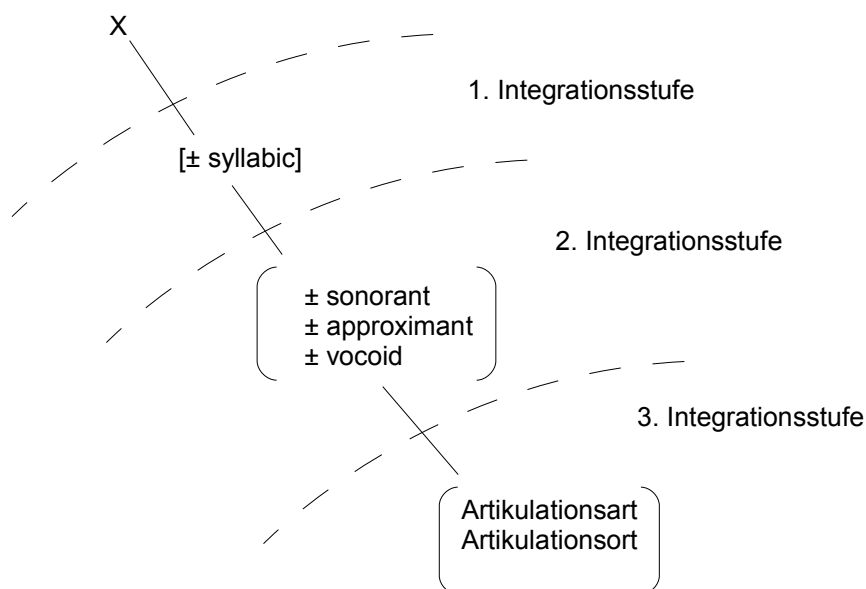


Abbildung 27: Schalenmodell des phonologischen Merkmalsausbaus.

Für Versprecher im Spracherwerb existiert leider eine teilweise widersprüchliche Datenlage. Konsens ist allerdings, dass sich produzierte Fehlleistungen von Versprechern im

Erwachsenenalter nicht derart unterscheiden, als dass man sie nicht mit den gleichen Modellen erfassen könnte (vgl. Poulisse 1999: 45). Dies steht im Einklang mit der Grundannahme dieser Arbeit, dass modulare Sprachproduktionsmodelle auch für die Analyse des frühen Spracherwerbs notwendig und hinreichend sind (vgl. Kap. 1.2.2). Es scheint zudem der Fall zu sein, dass Versprecher im Spracherwerb häufiger vorkommen und Perseverationen häufiger als Antizipationen auftreten (vgl. Smith 1990; Wijnen 1992).²⁴ Eine Deutung der häufigen Perseverationen kann an dieser Stelle nicht gegeben werden, da nicht bekannt ist, unter welchen silbenstrukturellen Bedingungen diese häufiger auftraten.

Letztlich liegt es aber bei einer solchen Datenlage nahe, dass der Abruf und die Integrationsprozesse innerhalb der Sprachproduktion maßgeblich für die Versprecherproduktion sind und nicht die Sprachkompetenz. Und man kann mit einiger Sicherheit annehmen, dass mit geringerem Erwerbsalter die Performanzfaktoren einen höheren Einfluss haben. Vor allem die Fragen danach, wie viel sprachliche Information im Arbeitsgedächtnis respektive Buffer berechnet werden kann und wie lange und in welcher Form sie in diesem zur Verfügung steht, erscheint in Anbetracht obiger Überlegungen zur phonologischen Sprachplanung im Produktionsprozess von zentraler Bedeutung. Diesen Fragestellungen soll im folgenden Kapitel nachgegangen werden.

1.6 Sprachplanung und -verarbeitung, Gedächtnissysteme und die Parallele Architektur

In Kapitel 1.2.1 wurde bereits die Parallele Architektur von Ray Jackendoff in ihren Grundzügen dargestellt. Ein wichtiger Aspekt der Parallelen Architektur ist, dass die diskrete Bildung phonologischer, syntaktischer und konzeptueller Strukturen innerhalb eines entsprechend partitionierten, sprachlichen Arbeitsgedächtnisses stattfindet. Die Parallele Architektur ist somit sowohl als Modell zur linguistischen Analyse als auch als Sprachverarbeitungsmodell konzipiert und inkorporiert in der spezifischen Sequenzierung der sprachlichen Submodule die allgemeine Verarbeitungslogik des Garrett-Modells (vgl. Kap. 1.5.1 und 1.5.2). Dies bedeutet folgende logische Abfolgen in Sprachproduktion und -verstehen:

²⁴ Ähnliche Ergebnisse erhielten Schwartz et al. (1994) bei einem Vergleich der Versprecher eines Jargon-Aphasikers mit Versprechern aus dem London-Lund-Korpus, wobei der aphasische Proband zudem mehr Nicht-Wörter bei Versprechern produzierte.

Logische Direktionalität im Sprachverstehen:

Laut => Phonologie => Syntax => Bedeutung

Logische Direktionalität in der Sprachproduktion:

Laut <= Phonologie <= Syntax <= Bedeutung

Dass sämtliche sprachliche Information in einem zuständigen Arbeitsgedächtnis eingespeist und online gehalten wird, unterscheidet dieses Modell erheblich von anderen Konzeptionen des sprachlichen Arbeitsgedächtnisses, wobei das Modell von Baddeley & Hitch (1974; vgl. auch Baddeley 2003) wohl als das bekannteste gelten kann. In letzterem spielt die sogenannte Phonologische Schleife eine entscheidende Rolle, in der phonologisch analysierte Information in einem phonologischen Kurzzeitspeicher für wenige Sekunden abgelegt und über ein inneres, subvokales Sprechen wieder aufgefrischt werden kann. Diese sprachliche Komponente des Arbeitsgedächtnisses operiert neben dem visuell-räumlichen Notizblock, der entsprechend visuelle Informationen speichert und zur mentalen Manipulation präsent hält. Die dritte Komponente ist die Zentrale Exekutive, die als Restkategorie Prozesse bearbeitet, die von den zwei anderen Komponenten nicht erfasst werden (z.B. Addition von Zahlen). Darunter fällt auch die Steuerung von Aufmerksamkeit und Abrufprozessen aus dem Langzeitgedächtnis. Das grundlegende Modell (Abb. 28) entstand vornehmlich auf Basis von Testungen, in denen es hauptsächlich um das Nachsprechen und Memorieren von unrelatierten Einheiten, d.h. Listen von (Nicht-)Wörtern, Buchstaben oder Zahlen ging.

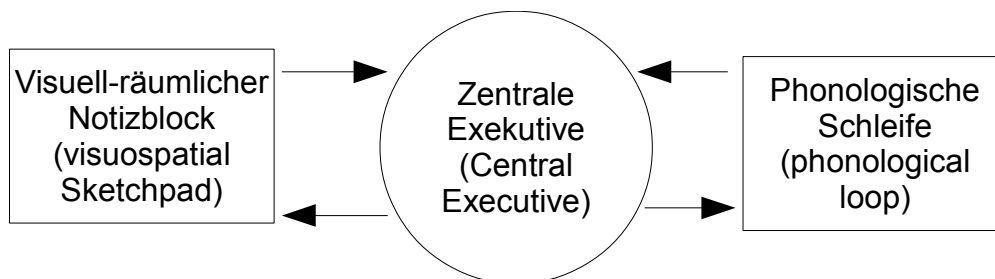


Abbildung 28: Phonologisches Arbeitsgedächtnis nach Baddeley & Hitch (1974).

Wie Jackendoff kritisch anmerkt, ist es bei Weitem nicht hinreichend, das sprachliche Arbeitsgedächtnis auf Basis einer solchen Evidenz zu konzipieren, im Gegenteil verzerrt es geradezu die wesentliche Funktion und Arbeitsweise sprachlicher Informationsverarbeitung:

„What strikes me about this tradition is that it apparently neglects what would appear to be by far the most important function of linguistic working memory: understanding and producing spoken language.“ (Jackendoff 2002: 206)

Schon Miller (1956) hat gezeigt, dass es nicht die pure Anzahl der informationellen Einheiten ist, die maßgeblich Arbeitsgedächtnisprozesse bestimmt, sondern die Art und Weise, wie die jeweiligen Informationen zu einem Datenblock (chunk) zusammengefasst werden. Es ist also gerade die Kombinatorik von Informationen, die wesentlich für die Prozesse innerhalb von Sprachproduktion und Sprachverstehen ist. Wie genau kann man sich aber die Interaktionen von Langzeitgedächtnis und Arbeitsgedächtnis beim Erstellen von neuen Informationseinheiten (die alle alten Einheiten enthalten), dem Chunking, vorstellen? Wichtig für eine Modellierung solcher Prozesse sind die von Miller et al. (1960) entwickelten TOTE-Einheiten, mit deren Hilfe man sogenannte kompensierende Rückkopplungen modellieren kann. Rein formal wird in solchen Einheiten ein Input empfangen, der im Element TEST hinsichtlich eines SOLL-Zustandes evaluiert wird. Ist keine Übereinstimmung gegeben, trifft das Element OPERATE Maßnahmen zur Kompensation dieser Diskrepanz. Diese können aus der Angleichung des IST-Zustandes an den Soll-Zustand bestehen oder aus einer Anpassung des Soll-Zustandes. Ist nach dieser Prozedur, die aus vielen Rückkopplungen bestehen kann, eine Übereinstimmung gefunden, fließt die Information in der Verarbeitungskette weiter (EXIT).

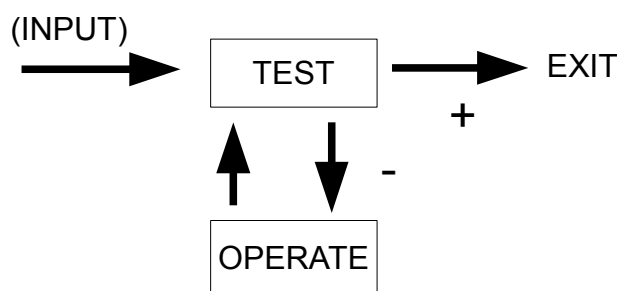


Abbildung 29: TOTE-Einheiten nach Miller et al. (1960).

Solche TOTE-Einheiten kann man nun in umfassende kognitive Modellierungen implementieren. So z.B. in das kognitive Verarbeitungsmodell MEKIV (Rahmenmodell zur elementaren und komplexen menschlichen Informationsverarbeitung) von Hussy (1993) (Abb. 30). Dieses Modell beinhaltet die klassischen Gedächtniskomponenten von Atkinson & Shiffrin (1968), sprich Langzeitgedächtnis, Kurzzeitgedächtnis, Ultrakurzzeitgedächtnis bzw. Sensorisches Register, erweitert durch das Arbeitsgedächtnis und Unterteilungen des

Langzeitgedächtnisses in informationspezifische Typen wie Faktenwissen (epistemische Struktur (deklaratives Gedächtnis)), Veränderungswissen (heuristische Struktur (prozedurales Gedächtnis)) und Bewertungswissen (evaluative Struktur).

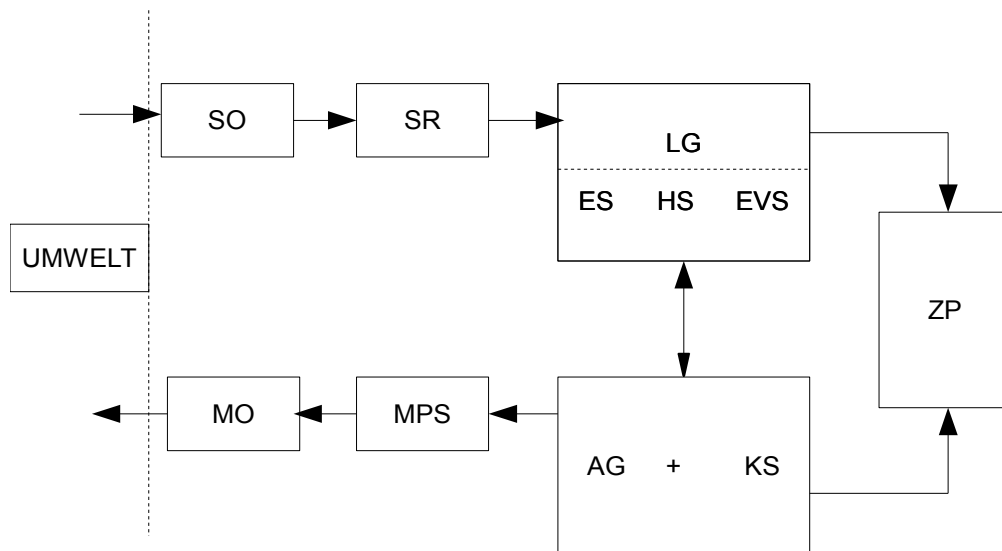


Abbildung 30: MEKIV nach Hussy (1993).

SO = Sinnesorgane SR = Sensorische Register LG = Langzeitgedächtnis
 ES = Epistemische Struktur HS = Heuristische Struktur EVS = Evaluative Struktur
 ZP = Zentraler Prozessor AG = Arbeitsgedächtnis KS = Kurzzeitspeicher
 MPS = Motorisches Programmsystem MO = Motorik

MEKIV ist eine Modellierung, die zum Zwecke der Darstellung von Problemlöseprozessen entworfen wurde. Analog kann man aber m.E. auch sprachliche Verarbeitungsprozesse modellieren. In Bezug auf Chunks wird in MEKIV eine Unterscheidung zwischen Arbeitsgedächtnis und Kurzzeitgedächtnis angenommen. Diese Trennung lässt sich z.B. durch Experimente von Baddeley & Hitch (1974) motivieren, bei denen im Rahmen einer Mehrfachaufgabe eine Problemstellung zum schlussfolgernden Denken sowie gleichzeitig das Behalten von einer Anzahl sinnfreier Silben zu leisten war. Wird die Kapazität des Kurzzeitgedächtnisses durch mehr Silben stärker in Anspruch genommen, sinkt gleichzeitig die Problemlösequalität. Für die Bildung von Chunks wird angenommen, dass im Kurzzeitgedächtnis Abrufhinweise für diese gespeichert werden, so dass das Arbeitsgedächtnis ökonomischer operieren kann. Anstatt sich also eine Reihe von Zahlen monoton einzeln zu merken, können diese unter einer Etikettierung oder einem Index gespeichert werden. In einer Untersuchung von Chase & Ericsson (1982) konnte sich z.B. ein Proband Reihen von Zahlen nach Maßgabe von Rekordzeiten der Leichtathletik merken, die

schon im Kurzzeitgedächtnis etikettiert und im Langzeitgedächtnis gespeichert waren (wie in Abb. 31 die Zahl 987 als Weltrekordzeit für die 100-Meter-Strecke bei Männern).

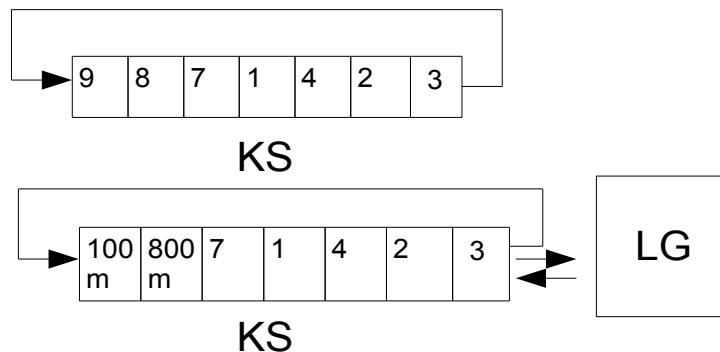


Abbildung 31: Monotone Zahlenliste (oben) vs. Chunking im Kurzzeitgedächtnis (unten)

Die Interaktion von Langzeitgedächtnis mit Kurzzeit- und Arbeitsgedächtnis ist bei allen Problemlöseprozessen relevant. In MEKIV laufen unter der Steuerung einer zentralen Steuerungseinheit, dem zentralen Prozessor, verschiedene Systemzustände für eine Problemlösung ab (Abb. 32). Es erfolgt zuerst eine Identifikation der faktischen Problemstellung über die epistemische Struktur. Ist in dieser für eine Problemstellung keine Lösung zu finden, geht das System in einen anderen Systemzustand über, in dem in der heuristischen Struktur nach Operatoren zur Lösung des Problems gesucht wird.

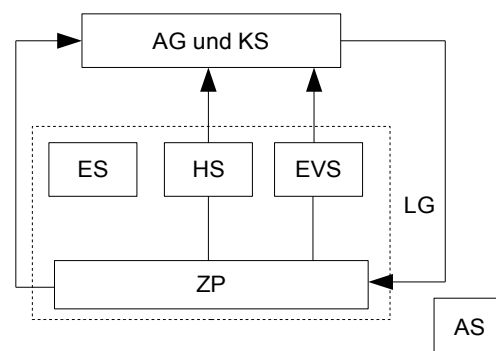


Abbildung 32: Mögliche Rückkopplungsschleife nach Anwendung eines Operators.

Der Inhalt des Arbeitsgedächtnisses definiert hierbei die Auswahl der Operatoren. Ist der Probleminhalt z.B. eine Zahlenfolge, so wird nach mathematischen Operatoren gesucht (addiere, subtrahiere, usw.). Entsprechend wird der Inhalt des Arbeitsgedächtnisses umstrukturiert, wobei hier eben Chunking für eine ökonomische Verarbeitung appliziert

werden kann. Alle Ergebnisse der hier beschriebenen Prozesse werden abschließend durch das Bewertungswissen evaluiert und entschieden, ob sie in das Ausgabesystem gegeben werden können. Falls nicht, können alle Systemzustände durch existierende TOTE-Einheiten, also kompensierender Rückkopplungsschleifen, erneut durchlaufen werden.

Es muss hier betont werden, dass solchen kognitiven Systemen die sukzessive Ökonomisierung der Verarbeitung über längere Zeiträume hinweg immanent ist. D.h., bei kontinuierlicher Informationsverarbeitung gleichen Typs wird das System automatisch selbst größere Einheiten respektive Chunks bilden, um effizientere Lösungen für jeweilige IST-SOLL-Problemstellungen zu finden, wobei vorläufige Lösungen nur den bestmöglichen Kompromiss in Bezug auf angestrebte SOLL-Strukturen darstellen können. Bezogen auf den Spracherwerb bedeutet dies, dass auch hier für die Performanz Chunking ein zentraler Prozess sein dürfte, was auch von Jones (2012: 167) betont wird:

„For the very young child, only a limited amount of spoken information will be stored in short-term memory. Relative to older children, very young children have had little exposure to their native language and therefore they will have little opportunity to chunk phoneme sequences into large chunks. This means children’s early vocabulary acquisition will proceed slowly, because they will find it difficult to represent in short term memory all of the necessary phoneme sequences that constitute a word. Over time, however, children store increasingly larger chunks of phoneme sequences, meaning that over time, children are able to store all of the necessary phonemes for words, phrases, and even whole utterances within their short-term memory, all because they have chunked the constituent phonemes. There is likely to be a “tipping point” during this developmental trajectory whereby the child has now learned enough chunked phoneme sequences to suddenly be able to store and subsequently learn the phoneme sequences for many new words – predicting the so-called vocabulary spurt that we see in children’s vocabulary acquisition.“

Nach Jones ist Chunking der zentrale Prozess in der Entwicklung sprachlicher Performanz, was er u.a. anhand des Erwerbs von Nicht-Wörtern im Computermodell EPAM zeigt (ebd.).²⁵ Hier wurden zwei andere wichtige Faktoren, reine Kapazität und Verarbeitungsgeschwindigkeit, konstant gehalten und nach altersbemessenen Input aus CHILDES nur mit implementiertem Chunkingprozess signifikante Korrelationen mit Leistungen 2-3 jähriger bzw. 4-5 jähriger Kinder erreicht. Es ist also zwischen einem Arbeitsgedächtnis, in dem Informationen kombiniert werden, und einem Kurzzeitgedächtnis, das diese Informationen durch sich etablierende Abrufhinweise ökonomisch organisiert, zu unterscheiden. In der Parallelen Architektur ist diese Unterscheidung nicht explizit gegeben, da der Fokus auf der Kombination von Inhalten im Arbeitsgedächtnis liegt. Die Operation, die für die Kombination

²⁵ „[...] chunking should be considered as an explanation of age-related changes in performance before any other mechanisms of development.“ (Jones 2012: 2)

von sprachlichen Informationseinheiten, aber auch solcher aus anderen kognitiven Domänen, zuständig ist, wird hier als UNIFY bezeichnet:

„[...] the only *procedural* rule will be UNIFY (“clip“ structures together). The counterpart of a traditional “rule of grammar“ in this systems is a lexicalized grammatical pattern with one or more variables in it.“ (Jackendoff 2002: 182)

Strukturbildung findet also qua UNIFY prozedural, d.h. im Arbeitsgedächtnis statt und wird durch die im Langzeitgedächtnis gespeicherten und domänenspezifischen Strukturformate beschränkt und kontrolliert, ähnlich also den heuristischen Strukturen in MEKIV. Diese domänenspezifische und auf dem Langzeitgedächtnis basierte Abgleichungsoperationen werden von Jackendoff (phonologischen, syntaktischen und konzeptuellen) Integrativen Prozessoren zugeordnet (vgl. Jackendoff 2007a). In Anlehnung an das Modell von Baddeley & Hitch (1974) überträgt ein Kopiermechanismus also Langzeitgedächtnisinformtionen ins Arbeitsgedächtnis, so dass die Parallele Architektur den Grundaufbau einer klassischen Architektur mit der Trennung von Programm und Datenspeicher zeigt (vgl. Kap. 1.2.3)²⁶:

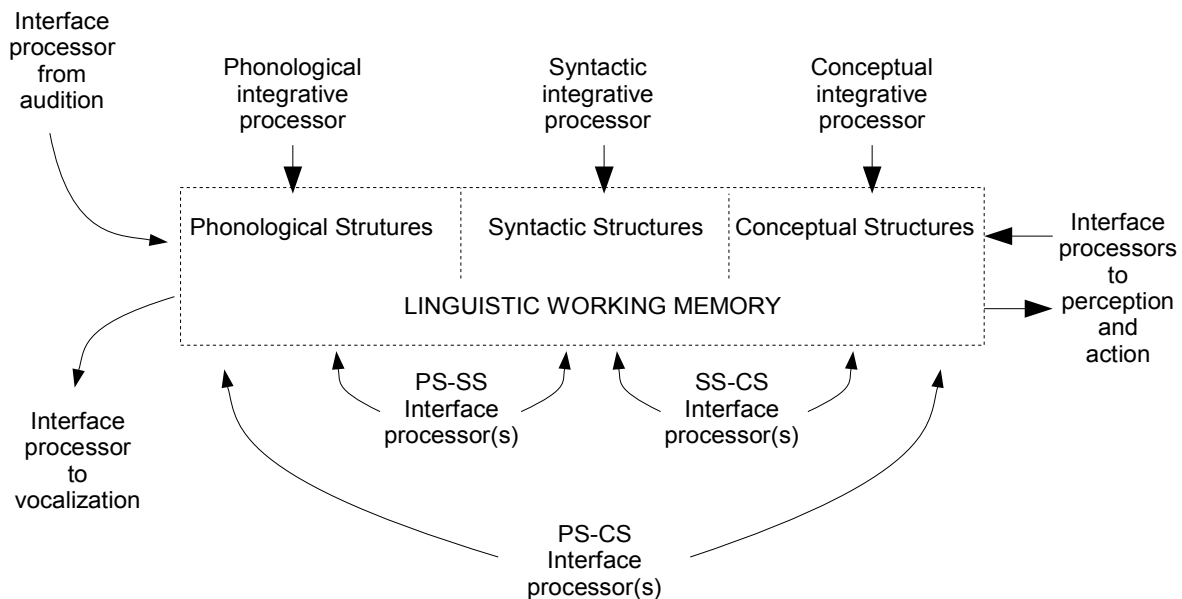


Abbildung 33: Die Parallele Architektur nach Jackendoff (2002).

²⁶ Jackendoff (2002: 205) betont, dass diese Trennung aufgrund des von ihm formulierten „Problem of 2“ logisch notwendig sei. Angenommen ein Lexem wird in einem Satz zweimal realisiert. Wird eine Langzeitgedächtnisrepräsentation im Sinne des Konnektionismus im Netzwerk immanent durch die Aktivierung eines Knotens konzipiert, so können beide Lexeme strukturell nicht unterschieden werden. Der Knoten wird hier während des Satzes nach der erstmaligen Aktivierung des Lexems aktiv bleiben, um die Interpretation desselben zu garantieren und kann somit nicht ein zweites Mal feuern oder aber, er feuert ein zweites Mal und so können beide nicht voneinander unterschieden werden.

Die oben benannte grundsätzliche Direktionalität von Sprachverstehen und Sprachproduktion bedeutet nun nicht, dass in den Submodulen und den spezifizierten Strukturbereichen des Arbeitsgedächtnisses nicht parallel Information verarbeitet werden kann. Vielmehr wird in der Parallelen Architektur eine inkrementelle Verarbeitung angenommen, wobei durch die Schnittstellenmodule ein permanentes Feedback im Sinne einer optimalen Abbildung der gebildeten Strukturen aufeinander gewährleistet ist (vgl. Jackendoff 2007b: 67-68). Wir könnten ambige Strukturen z.B. nicht zügig korrekt parsen, wenn wir nicht kurzfristig konkurrierende Repräsentationen eines Satzes unterhalten würden. So könnte die phonologische Segmentkette *It's not a parent* in a. auch eine zweite Interpretation im Sinne von b. erhalten.

- a. It's not a parent, it's actually a child.
- b. It's not apparent, it's actually quite obscure.

D.h., es werden über die phonologische Komponente zuvorderst beide Optionen und mit ihnen die jeweiligen syntaktischen und konzeptuellen Strukturen abgerufen und in den Arbeitsspeicher eingespeist. Erst durch den Abgleich der Nachfolgeinformation mit den zeitgleich präsent gehaltenen Repräsentationsoptionen kann es zur richtigen Auswahl und damit zur Disambiguierung kommen. Wie Csépe in Jackendoff (2003: 673) hervorhebt, weist die Parallele Architektur in ihrem Aufbau und ihrer illustrierten Arbeitsweise Ähnlichkeiten zu dem Arbeitsgedächtnismodell von Just & Carpenter (1992) auf. In diesem werden ebenfalls multiple Informationen parallel in einem ressourcen- oder kapazitätsbeschränkten Arbeitsgedächtnis verarbeitet. Dieses Modell fußt auf der Evidenz, dass unterschiedliche Arbeitsgedächtniskapazitäten direkte Effekte auf das Verstehen komplexer Sätze und das Auflösen von ambigen Sätzen hatten.

In seinem Bestreben die „Firewall“ der Kompetenz-Performanz-Unterscheidung aufzuheben, konzipiert Jackendoff Kompetenz also als unbewusste Kenntnis der Sprache, allerdings nicht derjenigen eines idealen Sprechers, sondern als kognitive Kenntnis, die sich in nachweisbaren Sprachproduktions- und Sprachverstehensprozessen des individuellen Sprechers zeigen lässt:

„The rules of the language, including the words, are posited to be precisely what the processing system uses in constructing mappings between sound and meaning.“
(Jackendoff 2007b: 69)

Allerdings bleiben in der Darstellung der Parallelen Architektur einige wichtige Punkte unklar. Sein Plädoyer für die Interaktion von Performanztheorien mit Kompetenztheorien wie

auch seine Modellierung der Sprachverarbeitung lassen die Frage offen, in welche Richtung Adjustierungen möglich sind. In Hinblick auf die Sprachverarbeitung ist m. W. nur die Rede von Beschränkungen durch die Strukturen der Integrativen Prozessoren. Es stellt sich natürlich die Frage, ob Langzeitgedächtnisrepräsentationen im Grunde nicht wiederum nur durch das Arbeits- und Kurzzeitgedächtnis konsolidiert werden können. Wichtig ist an dieser Stelle die Abgrenzung der Parallelen Architektur zu Ansätzen, die Kompetenz im Wesentlichen von der Performanz ableiten. Der Ansatz von Hawkins (1994, 2004) sei hier als eine prominente Variante zur Illustration herangezogen. Mit seiner Performance-Grammar Correspondence Hypothesis (PGCH) vertritt er genau diese Idee, dass in Grammatiken im Grunde Performanzpräferenzen konventionalisiert sind, die man anhand von Sprachkorpora und Messungen der Verarbeitung in psycholinguistischen Experimenten nachweisen könne.²⁷ Ausdrücklich weist Hawkins darauf hin, dass er das Verhältnis bisheriger Kompetenz-Performanz-Distinktion umkehren möchte:

„Chomsky claimed (and still claims) that grammar was autonomous and UG was innate [...]. The PGCH [...] is built on the opposite assumption that grammatical rules have incorporated properties that reflect memory limitations and other forms of complexity and efficiency that we observe in performance.“ (Hawkins 2014: 6)

Resultat solcher Ansätze ist m.E., dass hinter die Errungenschaften der Generativen Grammatik zurückgegangen wird und durch eine taxonomische Methode (vor allem in der Korpuslinguistik) im Grunde unter dem Etikett der Psycholinguistik eine E-Sprachen-Linguistik betrieben wird.²⁸ Die Unterscheidung zwischen E(xternalisierter)- und I(nternalisierter)-Sprache wurde von Chomsky (1986) hinsichtlich des Untersuchungsgegenstandes der Linguistik eingeführt. Eine E-Sprachen-Linguistik steht in der Tradition des amerikanischen Strukturalismus und geht strikt taxonomisch vor, d.h. sie beschreibt auf Basis von gesammelten Sprachdaten in diesen enthaltene Regularitäten und begreift insofern Sprache als ein externes Objekt unabhängig vom Sprecher. Aus der Forschungsperspektive einer I-Sprachen-Linguistik stellt sich Sprache hingegen als ein internes, mentales Vermögen eine potentiell unendliche Anzahl von Sätze zu generieren dar und der Forschungsfokus ist darauf gerichtet, wie Sprache in einem kognitiven System repräsentiert und sprachliche

²⁷ Performance-Grammar Correspondence Hypothesis (PGCH): Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (Hawkins 2004: 3)

²⁸ Besonders makaber wird diese Praxis, wenn auf schriftliche Daten aus Internet-Korpora zurückgegriffen wird und z.B. jahrelang (und eventuell von verschiedenen Personen) überformte juristische oder medizinische Textabschnitte völlig selbstverständlich analog zu oralen Performanzdaten analysiert werden, um anschließend Rückschlüsse auf Performanz-Kompetenz-Interaktionen zu ziehen (so z.B. in Trotzke et al. (2013), wo dieses Vorgehen u.a. auch noch als ein Beitrag zur Bilinguistik verstanden wird).

Information verarbeitet wird. Auch Jackendoff hebt für den Spracherwerb eindeutig eine I-Sprachen-Perspektive hervor:

„In any computational theory, 'learning' can consist only of creating novel combinations of primitives already innately available.“ (Jackendoff (1990: 40))

Genau eine solche Sicht einer I-Sprachen-Linguistik soll auch für den frühen Spracherwerb in dieser Arbeit stark gemacht werden. Besonders in Kapitel 2.1.2 wird mit der Strong Identity Hypothese von Hale & Reiss (2008) ein I-Phonologie-Modell zur Erfassung phonologischer Entwicklung vorgestellt, das problemlos mit der Parallelen Architektur zusammengedacht werden kann.

In Kapitel 1.2.3 wurde bereits darauf hingewiesen, dass auch Reduplikationen unter Zugrundelegung einer klassischen Architektur (d.h. der Trennung von Programm und Datenspeicher) verstanden werden können. Vor allem mit Raimy (2000) und Frampton (2009) findet man Ansätze vor, die auf sehr elegante Weise Reduplikationsprozesse erfassen und durch ihre postulierten Analyse-Instrumentarien für die Beschreibung von Spracherwerbsprozessen optimal verwertbar sind.

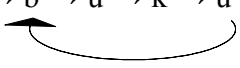
1.7 Reduplikation

1.7.1 Precedence-Based Phonology

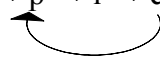
Der Precedence-Based Phonology (Raimy 2000, 2009) liegt die Annahme zugrunde, dass Reduplikation die Wiederholung einer Sequenz von Segmenten ist. Die Abfolge von Segmenten ist in diesem Ansatz explizit repräsentiert und zwar wie folgt: Der Beginn einer Segmentabfolge wird durch das Symbol #, das Ende durch das Symbol % gekennzeichnet. Die Abfolgeangaben, die normalerweise nur implizit in der graphischen Anordnung dargestellt werden, sind in der Precedence-Based Phonology mit dem Symbol \rightarrow spezifiziert. Die Anordnung $X \rightarrow Y$ bedeutet also X geht Y voraus.

Die Abfolgeangabe bzw. Linearisierung der phonologischen Repräsentation wird im Sinne von Chomsky (1995) aufgrund von *bare output conditions* benötigt. Die Beschaffenheit des phonetischen Moduls bzw. der sensomotorischen Schnittstelle verlangt, dass, um eine phonologische Repräsentation interpretieren zu können, diese asymmetrisch, transitiv und irreflexiv sein muss. Asymmetrisch bedeutet, dass wenn bei einmaligem Vorkommen der genannten Segmente in einer Form $\#k \rightarrow \text{æ} \rightarrow t\%$ „k geht æ voraus“ gegeben ist, „æ geht k


voraus“ notwendig falsch ist. Transitiv ist eine Abfolge deshalb, weil im Falle des Wahrseins von „k geht t voraus“, auch „k geht æ voraus“ und „æ geht t voraus“ wahr ist. Eine Abfolge ist irreflexiv, weil in einer solchen Repräsentation kein Segment sich selbst vorausgehen kann. Reduplikation geht in der Precedence-Based Phonology auf Schleifenbildungen (engl. *looping*) zurück Die morphologische Komponente fügt durch definierte Ankerpunkte (engl. *anchor points*)²⁹ Schleifen ein. Ankerpunkte geben an, ab welchem und bis zu welchem Segment eine Abfolge wiederholt wird. Im folgenden Beispiel aus dem Indonesischen findet in b) eine totale Reduplikation des Wortes [buku] statt, da die Schleife beginnend vom letzten Segment anzeigt, die Sequenz ab dem ersten Segment zu wiederholen (vgl. Raimy 2000: 12).

- a) # → b → u → k → u → % [buku]
 b) # → b → u → k → u → % [buku-buku]
- 

Auch andere Formen von Reduplikationen können mit Hilfe dieses Modells beschrieben werden. Die partielle Reduplikation in a) aus Nxa[?]amxcin kann wie in b) analysiert werden (nach Raimy 2009: 179):

- a) picpicx^w Basis: picx^w Bedeutung: 'abstoßend'
 b) Basis: # → p → i → c → x^w → %
 Morphologische Reduplikation: # → p → i → c → x^w → %
- 

Letztlich kann man auch sogenannte Out-of-control-Reduplikationen mit der Precedence-Based Phonology beschreiben (s. Beispiel a)). Die Reduplikation von nur einem Konsonanten wird dadurch erfasst, dass die Ankerpunkte für das gleiche Segment spezifiziert sind (b):

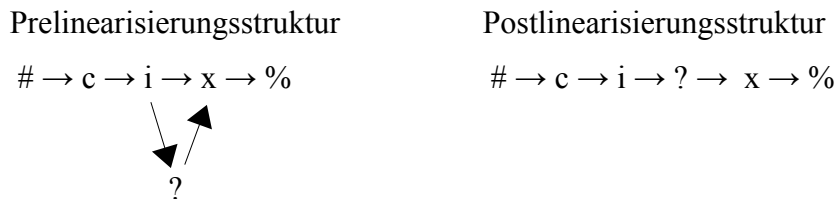
- a) k'ip'əp' Basis: k'ip' Bedeutung: 'gekniffen werden'
 b) Basis: # → k' → i → p' → %
 Morphologische Reduplikation³⁰: # → k' → i → p' → %
- 

²⁹ Raimy (2009: 180) weist darauf hin, dass die Anzahl der möglichen Ankerpunkte in Sprachen auf eine bestimmte Menge begrenzt ist. Diese Annahme ist identisch mit der Annahme von Yu (2003) von pivot points in der crosslinguistischen Analyse von Infigierung.

³⁰ Die Schwa-Epenthese erfolgt nach der Linearisierung und wird hier nicht besprochen (vgl. Raimy 2009: 181).

Raimy kann insofern sagen, dass „repetition is caused by loops in phonological representations because of a linearization process that is present within the phonology.“ (Raimy 2000: 12)

Auch andere morphologische Prozesse können mit der Precedence-Based Phonology analysiert werden. Infigierung in Nxaʔamxcin erfolgt dadurch, dass ein Segment inseriert wird, ohne dass eine Schleifenbildung gegeben ist (vgl. Raimy 2009: 183):



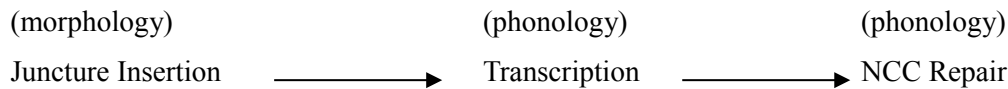
Ein Vorteil der Precedence-Based Phonology ist so, dass kein spezifischer Mechanismus für Reduplikation angenommen werden muss. Reduplikation kann im Gegenteil wie andere Prozesse durch die Schnittstellenbeschränkungen erfasst werden. In den Worten von Raimy (2000: 15):

„If reduplication is a loop, then reduplicative morphemes are reduced to a purely phonological representation without requiring any additional copying mechanism or correspondence function. This allows reduplication to be accounted by general architecture of grammar without resorting to any process specific mechanism. Linearization is not specific to reduplication because it ensures that the bare output condition of asymmetry is met by all phonological representations.“

1.7.2 Distributed Reduplication

Frampton (2009) versucht in seinem Modell der Distribuierten Reduplikation (DR) Evidenz für eine fundamentale Annahme generativer Phonologie zu präsentieren, die schon in Chomsky & Halle (1968) entwickelt wurde. Diese Annahme ist, dass Oberflächenrepräsentationen durch die sukzessive Modifikation von zugrundeliegenden Formen entstehen. Hier ist nicht der grundsätzliche Aspekt einer einfachen Derivation von Interesse, den auch andere phonologische Theorien, wie z.B. die Optimalitätstheorie, beinhalten, sondern die Idee, dass die Zwischenstufen im Ableitungsprozess Modifikationen einer einzigen Einheit repräsentieren.

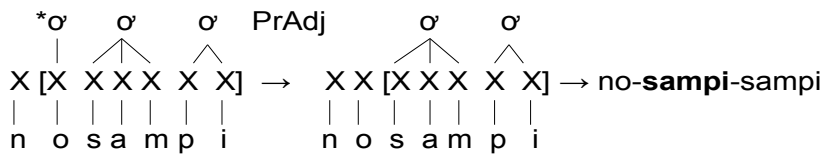
Reduplikation ist so für Frampton „the surface manifestation of nonconcatenative inflectional morphology“ (ebd.: 3), im Gegensatz zu verkettenden morphologischen Typen wie Präfigierung oder Suffigierung. Es wird hier kein zusätzliches Element affigiert, sondern es



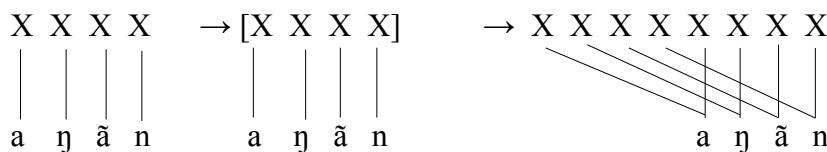
Rules can intervene here that adjust the prosodic shape of the projected reduplicant to a target prosodic condition.

Segmental rules that intervene here can be affected by conditions on applying rules to geminates.

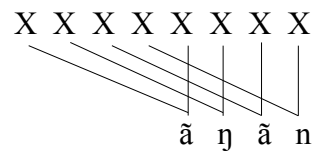
Zwei Beispiele sollen die intervenierenden Regelprozesse illustrieren. In der peruanischen Sprache Ashéninka Campa wird durch Reduplikation eine Form von Intensivierungspräfix gebildet, das mit „mehr und mehr“ übersetzt werden könnte. Die Markierung für die erste Person ist *n-*, *osampi* (*paddeln*) ist die Verbwurzel, die morphologisch redupliziert wird. Durch prosodische Adjustierung (PrAdj) wird in diesem Beispiel nun die ungrammatische onsetlose Silbe *-o* aus der Repräsentation der Basis und somit auch aus dem Reduplikanten entfernt. Das Ergebnis ist entsprechend *no-sampi-sampi* (vgl. ebd.: 83):



Als Beispiel für intervenierende Segmentregeln soll die Anwendung der Nasalisierungsregel im Malaiischen aus Kapitel 1.2.3 wiederholt werden. Hier ist die reduplizierte Form von *aŋān* (Träumerei) die Form *ãŋānãŋān* (Ambition). Die fehlende Erklärung dafür, warum die Nasalisierungsregel am ersten Vokal appliziert wird, obwohl kein hierfür notwendiger vorangehender Nasal gegeben ist, wird nach Frampton (ebd.: 8-9) eben durch die repräsentationalen Zwischenebenen geliefert. Bevor die gekreuzten Strukturen repariert werden, indem jedem X-Slot nur ein Phonem zugeordnet wird, wird zuvorderst die Nasalisierungsregel wirksam.



Auf dieser abstrakten Zwischenebene befindet sich qua Assoziationslinien der letzte Nasal vor dem initialen Vokal, so dass für die Nasalisierungsregel die entsprechenden lokalen Bedingungen gegeben sind:



Die Theorie der Distributed Reduplication bietet eine überaus elegante Beschreibung durch ihre Annahme dreidimensionaler Kopien der jeweiligen Basis.³¹ Frampton selbst gibt Hinweise darauf, dass Reduplikation in dieser Form für psycholinguistische Sprachproduktionsmodelle fruchtbar gemacht werden kann bzw. Reduplikation auf Mechanismen basiert, die diesen schon immanent sind:

„String copy is a mechanism that the language apparatus needs independently of reduplication. Strings must be copied from the lexicon to an active workspace, for example. I suppose that morphology and phonology co-opt this mechanism for reduplication. Morphology inserts junctures into the timing tier and phonology carries out copying [...]“ (ebd.: 51)

Im phonologischen Arbeitsgedächtnis werden also reduplizierte Ketten durch einfache Kopie erzeugt, wenn in der morphologischen Repräsentation eine solche Verdopplung symbolisch encodiert ist. Autosegmentale oder dreidimensionale Repräsentationen werden nach Frampton allerdings ebenfalls erst im phonologischen Arbeitsgedächtnis aufgebaut, wobei er auch im Sinne der Two-Root-Theory of Length (vgl. Kapitel 1.4.1) die Umcodierung segmentaler Längeninformation in eine dreidimensionale Repräsentation annimmt:

„Under the plausible assumption that the lexicon is optimized for space efficiency, it is implausible to suppose that exponents are stored as autosegmental representations. I therefore suppose that exponents are stored as simple phoneme/juncture sequences and that timing slots are supplied in the process of transcription from the lexicon to the phonological workspace. Gemimates are indicated by some diacritic that is removed in the process of transcription: perhaps a [+long] feature in the phoneme itself that is translated into autosegmental terms and removed in the process. Crucially, I assume that *autosegmental representation is the privilege of phonology*, used in its internal computation, but that autosegmental representation is unintelligible to both the input and output systems.“ (ebd.: 191)

Da Reduplikationen im frühen Spracherwerb ein so gut wie obligatorisches Phänomen sind, ist es m.E. unabdingbar, sie kohärent in eine kognitive Modellierung früher Wortproduktionen

³¹ Frampton (2009: 7) hebt in einer Zitation einer Stelle aus Kitagawa (1987) explizit auf die Dreidimensionalität der Repräsentationen ab: “[...] a morphological rule of reduplication *three-dimensionally copies* [my emphasis] the relevant CV skeleton and its association lines directly from the base.”

einzuordnen. Um eine solche Einordnung adäquat vornehmen zu können, gilt es allerdings zuvor, die allgemeine Entfaltung phonologischer Repräsentationen im Spracherwerb unter Beachtung bisheriger Ergebnisse zu beleuchten.

2 Der Erwerb phonologischer Strukturen

2.1 Der Erwerb in mentalistischen Ansätzen

Eine der zentralen Fragen innerhalb der Spracherwerbsforschung ist, welche Beziehung zwischen den sprachlichen Repräsentationen von Erwachsenen und denen von Kindern besteht. Eine weitere wichtige Thematik ist, ob innerhalb des Spracherwerbs selbst fließende Übergänge oder diskrete Stufen angenommen werden müssen. Einmal erfolgt also ein Vergleich der Repräsentationen von Sprachlerner (bzw. seinen unterhaltenen Zwischengrammatiken) und der Zielgrammatik, ein anderes Mal ein intrinsischer Vergleich von Repräsentationen des Sprachlerner in zeitlicher Abfolge. Für beide Fragestellungen gilt, dass sie das jeweils gegebene Phänomen nach Maßgabe von Kontinuität bzw. Diskontinuität untersuchen. Hinzu kommen Fragen des Geltungsbereiches. D.h., für welches repräsentationale Format (phonetisch vs. phonologisch) wird Kontinuität unter den obigen Fragestellungen untersucht und betrifft die Untersuchung die Sprachproduktion oder das Sprachverstehen? Es herrscht hinsichtlich der Fragestellungen und Geltungsbereiche einige Konfusion in den gängigen Forschungsansätzen. Es kann an dieser Stelle kein umfassender Überblick über die unterschiedlichen Ansätze gegeben werden. Vielmehr soll unter Zugrundelegung der theoretischen Aspekte der vorherigen Kapitel und grundsätzlichen Überlegungen von Hale & Reiss (2008) eine Exemplifikation der Problematik anhand der Besprechung prominenter Ansätze zum Erwerb der Phonologie durchgeführt werden.

In Bezug auf die Frage der Kontinuität von Lautrepräsentationen innerhalb des Sprachsystems von Kindern entwirft Jakobson (1941) ein Zwei-Phasen-Erwerbsmodell. Er unterscheidet strikt zwischen einer phonetischen Phase, in der vorsprachliche Lalllaute produziert werden, und einer eigentlichen Sprachstufe, in der es nach einer kurzen stillen Periode „um die Auslese der Sprachlaute und um den Aufbau des Phonemsystems geht“ (ebd.: 30). Dieses Modell von Jakobson wird häufig in Untersuchungen des Lauterwerbs herangezogen (vgl. z.B. Vihman et al. 1985), wobei allerdings der theoretische Unterbau meist nicht reflektiert wird. Als Erwerbsmodell eines strukturalistischen Ansatzes, das folglich ausschließlich in einer E-Sprachen-Linguistik begründet ist, fehlt natürlich die differenzierte Modellierung, die

ein auf I-Sprache fundierender, kognitionspsychologischer Ansatz zur Verfügung stellt. Dies gilt selbst für die simple Unterscheidung zwischen der Entwicklung von Perzeption und Produktion respektive der Unterscheidung zwischen einem Input- und Output-Lexikon. Die Vernachlässigung solcher modelltheoretischen Aspekte führt nicht selten zu obskuren Argumentationen. So unterstützt z.B. Vihman (1992) die Kontinuitätshypothese, da sie zwischen den Lauten und Silben, die in der Lallphase, und denjenigen, die in der Einwort-Phase produziert werden, signifikante Übereinstimmungen gefunden hat. Für die Repräsentationen von Wörtern wird andererseits angenommen, dass sich der Erwerb über holistische zu segmentierten Repräsentationen vollzieht:

„The assumption that the child begins at the middle level, which is immediately available in the speech the child hears, eliminates the need to posit the universal stock of linguistic units, segments, features, principles, rules, or word-learning constraints that some assume to be innately available.“ (Velleman & Vihman 2002: 10)

Wie Bernhardt & Stemberger (1998: 7) betonen, ist hierbei die Frage, ob es wirklich möglich ist, dass Repräsentationen sukzessiv und diskontinuierlich erworben werden, wohingegen sich die Produktion kontinuierlich vollzieht. Solch ein Widerspruch ergibt sich m.E. zwangsläufig aus einer rein phonetischen Betrachtungsweise, da nach dieser die phänomenologische Diskrepanz zwischen Perzeption und Produktion negiert werden muss und sprachliche Repräsentationen als Brückenglieder fehlen.

Wenn man mit einem mentalistischen Ansatz die These von universalgrammatischen Beschränkungen akzeptiert, so kann man, was die Frage des Verhältnisses von Zielgrammatik und Zwischengrammatiken betrifft, für Vertreter eines nativistischen Ansatzes grob folgende Hypothesen unterscheiden.

Befürworter der starken Kontinuitätshypothese (strong hypothesis of continuity) vertreten die Anschauung, dass UG-Prinzipien von Anfang an in Gänze vorhanden und aktiv sind und so jede Zwischengrammatik beschränken. Eine Variante, die Hypothese der vollen Kompetenz (hypothesis of full competence), nimmt an, dass alle vom Kind erzeugten sprachlichen Repräsentationen denjenigen, die in der zu erlernenden Zielsprache vorkommen, stark äquivalent sind.

Nach der schwachen Kontinuitätshypothese (hypothesis of weak continuity) hingegen ist die UG ebenfalls prinzipiell verfügbar. Befürworter dieser These nehmen an, dass strukturelle Repräsentationen nicht zieläquivalent von Anfang an vorhanden sind, jede Zwischengrammatik aber durch die Beschränkungen der UG definiert sind.

Vertreter der Reifungshypothese (hypothesis of maturation) hingegen postulieren, dass UG-Prinzipien im Laufe der sprachlichen Ontogenese heranreifen. Die Implikationen im Unterschied zu den anderen Ansätzen bestehen darin, dass das Kind zeitweise Hypothesen unterhält, die aus der UG herausfallen, die somit also nicht Teil einer Grammatik natürlicher Sprachen wären.

Als Ausgangspunkt für die Diskussion um den Erwerb von phonologischen Merkmalen soll das Structure-elaborating constituency Modell von Rice & Avery (1995) dargestellt werden, das hier als Beispiel für die schwache Kontinuitätshypothese gelten kann.

2.1.1 Structure-elaborating constituency model

Rice & Avery (ebd.) unterscheiden einen Full-specification-Ansatz von einem Minimal-specification-Ansatz. Ersterem liegt die Annahme zugrunde, dass ein Phonem wie z.B. /p/ in voll spezifizierter Form vom Kind erworben wird und bei Erwerb von weiteren Phonemen redundante Information der Merkmalsmenge gelöscht wird. Der von Rice & Avery vertretene Minimal-specification-Ansatz hingegen geht davon aus, dass nur soviel Information erworben wird, wie nötig ist, um ein Phonem von anderen Phonemen im Sprachsystem abzugrenzen. Insofern wäre die Feststellung, dass ein Kind das Phon [p] kontinuierlich gebraucht, für sich wenig aussagekräftig, da der phonematische Status nur über den Kontrast zu anderen Lauten ermittelt werden kann. Der Erwerb von phonologischen Repräsentationen wird von Rice & Avery also nicht als Eliminierung redundanter Information, sondern als sukzessiver Ausbau von phonologischen Kontrasten verstanden.

Phoneme sind nach Rice & Avery intern hierarchisch organisiert, mit vier Hauptknoten, Laryngeal (CG = Constricted Glottis; SG = Spread Glottis), Air Flow (Cont = Continuant), Sonorant Voice (SV) und Place (vgl. ebd.: 31):

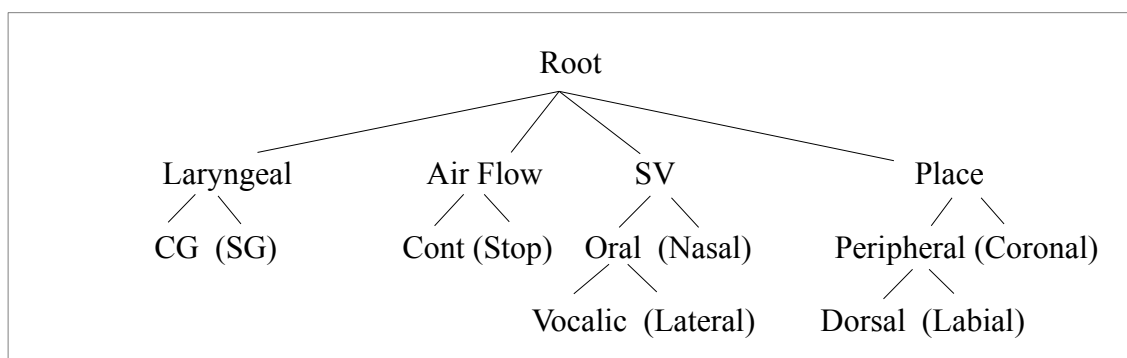


Abbildung 34: Erwerbshierarchie phonologischer Merkmale nach Rice & Avery (1995).

Innerhalb der Knoten bestehen Dependenz- und Markiertheitsrelationen. Auf jeder Stufe ist eine markierte und unmarkierte (mit Klammern gekennzeichnete) Option gegeben und Dependenz durch dominierende Knoten angezeigt. Wenn z.B. das Merkmal Dorsal vorhanden ist, beinhaltet dies das Vorhandensein der Merkmale Peripheral und Place. Die unmarkierten Merkmale sind Default-Merkmale, welche als redundante Information nicht angegeben werden müssen. So haben Koronale nur das Merkmal Place. Hier drei Repräsentationen für Plosive mit verschiedenen Artikulationsmerkmalen:

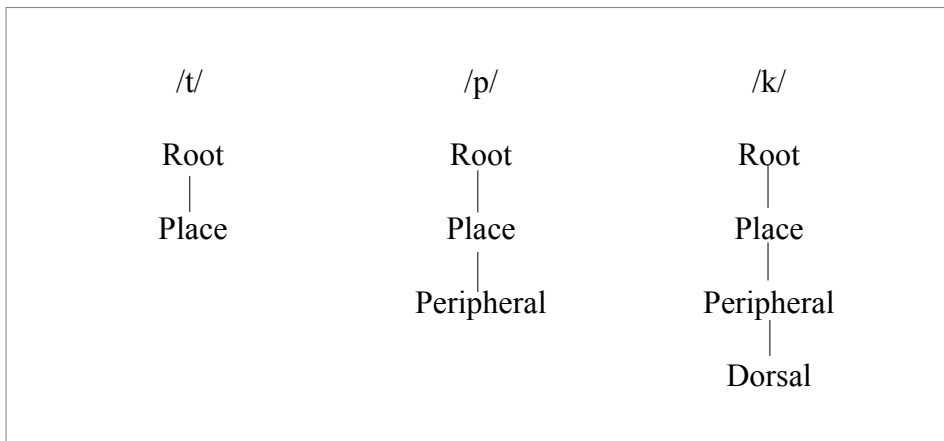


Abbildung 35: Ortsmerkmale bei Plosiven.

Differenzen zwischen Phonemen sind also auf Strukturweiterungen mit nachfolgender, kontrastiver Funktion zurückzuführen. Diese Strukturweiterungen finden, so die Annahme, Schritt für Schritt von der komplett unmarkierten Struktur ausgehend statt, so dass folgende Hypothesen gelten (ebd.: 35):

- a. *Minimality*: Initially the child has minimal structure.
- b. *Monotonicity*: Inventories are built up in a monotonic fashion.

Die erste Erwerbsstufe beinhaltet so noch keinen Kontrast, die zweite den Kontrast zweier Strukturen, die dritte den von dreien. Abbildung 36 gibt den Erwerb des Merkmals *Place* wieder, welcher über den Koronal-Nicht-Koronal-Kontrast zu einer Unterscheidung zwischen Nicht-Koronalen (Labiale vs. Dorsale) in deterministischer Form verläuft:

Zum Zeitpunkt 1 hat das Merkmal Continuant noch keine distinktive Funktion in wortfinaler Position, so dass auch der Frikativ als Plosiv realisiert wurde. Diese distinktive Funktion respektive der phonologische Kontrast ist aber zum Zeitpunkt 2 etabliert und folglich werden Plosiv und Frikativ wortfinal adäquat realisiert.

Strukturelle Variabilität ergibt sich durch eine differente zeitliche Abfolge im Ausbau der präeterminierten Strukturoptionen. So kann ein Kind zuerst den Kontrast zwischen Koronalen und anderen Placemerkmalen erwerben (Abb. 37) und hätte damit ein Inventar aus oralen und nasalen Phonemen, weil der oral/nasal-Kontrast noch nicht etabliert ist. Ein anderes Kind könnte aber auf dieser ersten Erwerbsstufe den Kontrast zwischen Obstruenten und Sonoranten etablieren, während Place nicht-contrastiv bleibt (Abb. 38).

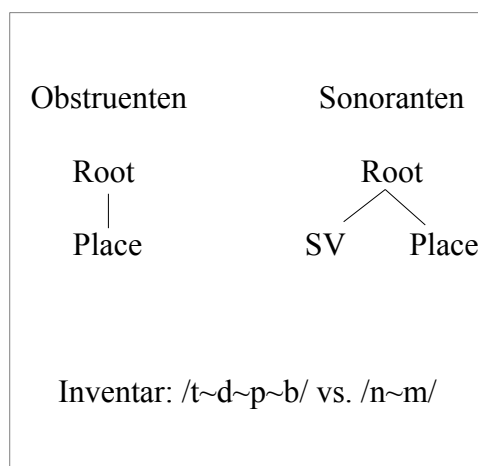


Abbildung 37: Phonologisches Segmentinventar bei Ausbau des Artikulationsartmerkmals.

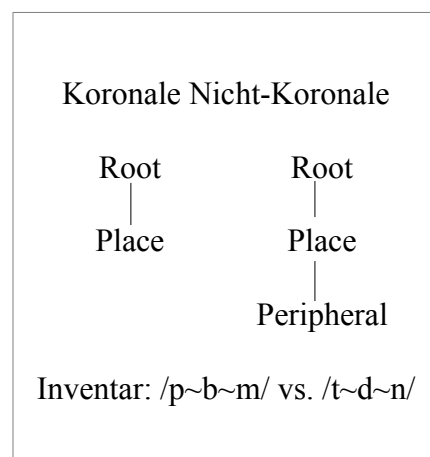


Abbildung 38: Phonologisches Segmentinventar bei Ausbau des Ortskontrastes.

Im weiteren Verlauf könnte Kind 1 nun z.B. weiter die Place-Merkmale ausbauen oder aber eventuell zuerst das SV-Merkmal.

Deutlich geworden sein sollte, dass das Modell von Rice & Avery wesentlich auf der Idee des strukturellen Merkmalausbaus fußt. Kinder beginnen mit einem kleinen phonologischen Inventar und beenden den Erwerb mit einem größeren Inventar, das alle Kontraste der vorherigen Erwerbsstufen beinhaltet und somit auch alle vorangegangenen Phoneme.

Hale & Reiss (2003) kritisieren diese Idee des Strukturausbaus scharf und führen gewichtige Argumente dagegen an. Im nächsten Kapitel sollen diese ausführlich dargestellt werden.

2.1.2 Strong Identity Hypothesis

Anhand einer eingehenden Diskussion des Subset Principles kritisieren Hale & Reiss (2003, 2008) prinzipiell die Idee des Phonemerwerbs als Strukturausbau. Das Subset Principle ist eine Lernbarkeitsbeschränkung, welche anzunehmen ist, da der Spracherwerb nur durch positive Evidenz getriggert wird. Die zugrundeliegende Frage ist, wie das Sprachsystem bei rein positiver Evidenz keine zu breiten Generalisierungen zulässt. Hale & Reiss (ebd.) reinterpreten das in der Syntaxtheorie häufig diskutierte Prinzip (vgl. Berwick 1986; Wexler Manzini 1987) in dem Sinne, dass Strukturausbaukonzeptionen der phonologischen Entwicklung als nicht lernbar klassifiziert werden müssen (vgl. Hale & Reiss 2003). Ausgangspunkt ihrer Kritik ist eine Definition von Jackendoff (1990: 40):

„In any computational theory, 'learning' can consist only of creating novel combinations of primitives already innately available.“

Für die Hypothesen der Minimality und Monotonicity (s.o.) ergeben sich aus einer strikt nativistischen Sicht grundlegende Probleme. Diese Hypothesen besagen letztlich im Einklang mit dem traditionellen Modell des Subset Principles, dass der Erwerbpfad von einem kleinen repräsentationalen (hier spezifisch phonologischen) Inventar sukzessive zu einem größeren ausgebaut wird:

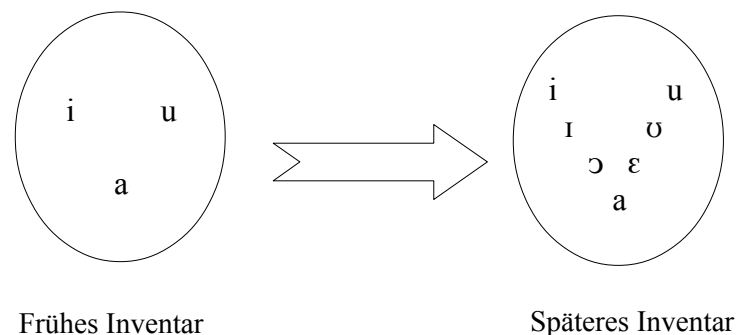


Abbildung 39: Das Subset Principle im traditionellen Modell.

Unklar ist, wie mit den gegebenen phonologischen Repräsentationen des früheren Inventars neue Phoneme überhaupt geparkt werden können. „To parse a string is to assign it a linguistic representation“, so Hale & Reiss (2008: 40), was bedeutet, dass nur die phonologischen Repräsentationen zugewiesen werden können, die bereits zugänglich sind, oder gar keine Zuordnung stattfinden kann. Das Problem lässt sich nur lösen, wenn man eine

merkmalsorientierte Interpretation des Subset Principles vornimmt. Auf eine Studie von Hyams & Sigurjónsdóttir (1990) rekurrend, in der ein Kontrast zwischen der Anapherbindung im Englischen und Isländischen im Rahmen von Parameterbelegung ausgearbeitet wird (im Englischen können Anaphern nur mit NPs im selben Satz koreferieren, im Isländischen können sie dies hingegen mit einem Antezedens in einer höheren Satzprojektion), zeigen Hale & Reiss, wie eine Spezifikation lexikalischer Merkmale im Rahmen des Minimalistischen Programms den Kontrast adäquater erfasst:

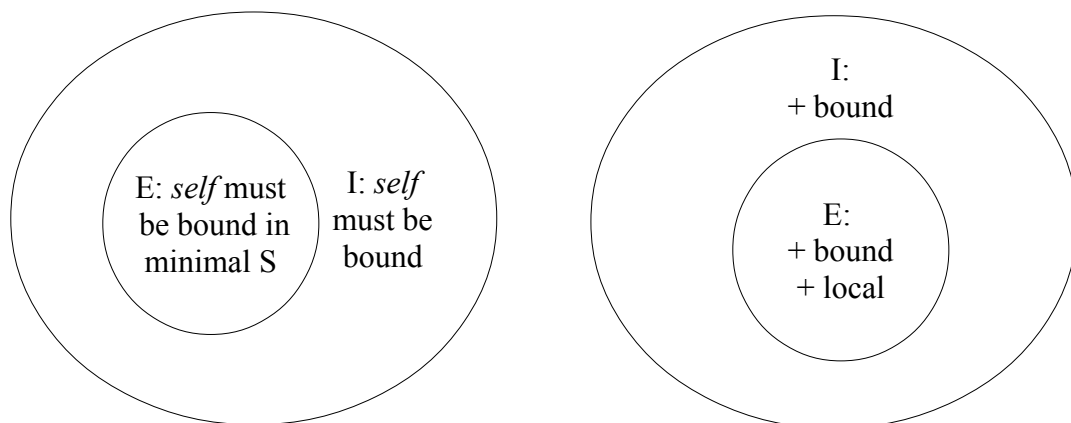
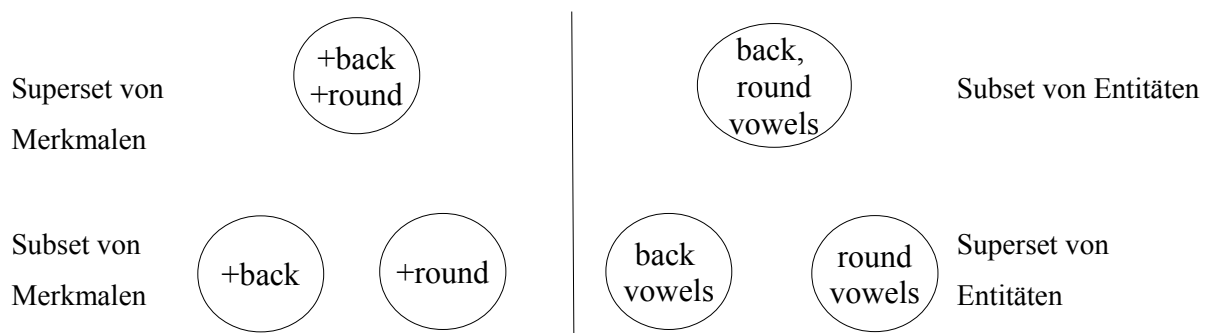


Abbildung 40: Traditionelle Interpretation des Subset Principles (links) vs. reinterpretierte Version.

Wie man erkennt, definiert das Merkmal +bound einen breiteren (alle Anaphern), die beiden Merkmale +bound, +local einen kleineren Geltungsbereich von Entitäten (lokal gebundene Anaphern). Weniger Merkmale implizieren so also mehr Entitäten. Auf phonologische Repräsentationen bezogen bedeutet das, dass der Erwerbpfad in einer Reduktion bzw. zunehmenden Spezifikation von Merkmalen bestehen muss und nicht in einem Ausbau derselben:



Diese konzeptuelle Notwendigkeit der Angeborenheit aller phonologischer Merkmale passt widerspruchsfrei zu den gegebenen Spracherwerbsdaten. Dass Kinder im Alter von 10 Monaten große Teile ihrer Diskriminationsfähigkeit von Sprachlauten eingeübt haben, zeigt den Spezifikations- und Selektionsprozess an. „We are „deafened“ by our experience“, so Hale & Reiss (2008: 52). Die Zusammenfassung der Argumentation von Hale & Reiss soll hier der Vollständigkeit halber in Gänze wiedergegeben werden (vgl. 2007: 16):

- „A. Subset Principle Argument
 - a. The Subset Principle reflects restrictiveness in the initial state.
 - b. Greater restrictiveness is encoded through fuller specification.
 - c. All features must be available for representations at S_0 .
- B. Learnability Argument
 - a. Linguistic representations contain features.
 - b. If a feature F is unavailable at stage L_j , then positive evidence of F cannot be evaluated by the learner since the learner cannot evaluate representations with respect to F.
 - c. All features must be available for representations at S_0 .
- C. Empirical Argument
 - a. Infants appear to be innately sensitive to any possible phonological contrast.
 - b. Phonological contrasts are parsed and represented in terms of features.
 - c. All features must be available for representations at S_0 .“

Mit ihrer Schlussfolgerung, dass alle Merkmale angeboren sind, vertreten Hale & Reiss eindeutig eine starke Kontinuitätshypothese. Sie bezeichnen die von ihnen vertretene Hypothese genau deshalb als Strong Identity Hypothesis, weil das phonologische System von Kindern durch die gleichen Prinzipien wie das der Erwachsenen beschränkt wird (Hale & Reiss 2008: 70). Während sie davon ausgehen, dass sich die Kompetenz in den perceptiven Leistungen widerspiegelt, sind die von Zielwörtern abweichenden Produktionen von Kindern ihrer Meinung nach zum größten Teil auf Performanzfaktoren zurückzuführen (ebd.: 56):

„We know that children's (cognitive and physiological) performance systems are largely underdeveloped. Therefore, even in the absence of logical arguments against such a discrepancy in representational capacity, we would be forced by Occam's Razor to assume as the null hypothesis that the discrepancy in output is due only to the immaturity of the performance systems. Of course, this hypothesis is subject to revision, but it represents the most constrained model available.“

An einem der vielfältigen Beispiele, die Hale & Reiss anführen, soll dieser Performanzansatz verdeutlicht werden. Smith (1973) macht darauf aufmerksam, dass in der Produktion seines Sohnes Amahl distributionelle Muster auftraten, die eine kompetenzbasierte Interpretation nahelegen. Bekannt geworden ist in diesem Zusammenhang das sogenannte Puzzle-Problem.

Amahl produzierte zu einem Entwicklungszeitpunkt *puzzle* als [pʌdəl], realisierte allerdings *puddle* als [pʌgəl]. Die Schlussfolgerung von Smith ist, dass angesichts der nachgewiesenen artikulatorischen Möglichkeit *puddel* zieladäquat zu realisieren, keine physiologischen Gründe als Erklärung für die Muster gelten können, sondern nur phonologische Kompetenz. Da dieses Beispiel zum gängigen Kanon phonologischer Untersuchungen zum Spracherwerb gehört, reanalysieren Hale & Reiss (2008: 65) es neu. Sie stellen im Grunde nur zwei basale Fragen, die das Argument obsolet machen:

1. Kann das Kind reliabel seine mentale Repräsentation von [d] in *puddle* realisieren?
2. Kann das Kind reliabel seine mentale Repräsentation von [z] in *puzzle* als ein [z] realisieren?

Auf beide Fragen lautet die Antwort: Nein. Die fehlerhafte Produktion von [z] als [d] lässt nicht den Schluss zu, dass Amahl fähig war, [d] reliabel in der entsprechenden Zielposition zu realisieren. Die Argumentationslinie, aus zwei performanzbasierten Fehlleistungen ein Kompetenzphänomen herzuleiten, ist unzulässig.

Sowohl Ansätze einer Struktur- und Merkmalerweiterung wie derjenige von Avery & Rice, als auch reine Kompetenzansätze sind nach der schlüssigen Argumentation von Hale & Reiss für eine Erforschung phonologischer Komponente in der Sprachproduktion ungeeignet. Darüber hinaus erscheint der Ansatz von Hale & Reiss in Hinblick auf Hagoorts Adäquatheitskriterien in Kapitel 1.2.2 angemessener. Die hier angeführten neuronalen Voraussetzungen zu Beginn der eigentlichen Sprachproduktion im Alter von ca. 10-12 Monaten sprechen mehr für ein in Gänze gegebenes Merkmalsrepertoire als für einen späteren Erwerb bzw. eine sukzessive Erweiterung von Merkmalen. In der Sprachproduktion zeigt sich somit vielmehr eine performanzbasierte Limitation in der Realisierbarkeit von schon vorhandenen Merkmalen, als eine (limitierte) Kompetenz. Performanzbeschränkungen des Kleinkindes in der Produktion grundsätzlich mit artikulatorischen Beschränkungen zu identifizieren, ist nicht zwingend notwendig, auch wenn dies unhinterfragter Usus scheint. Schon Jakobson (1941: 20) hat aber auf die Produktionen in der Lallphase verwiesen:

„Ein lallendes Kind kann Artikulationen anhäufen, die sich nie innerhalb einer Sprache oder sogar einer Sprachgruppe zusammenfinden: Konsonanten beliebiger Bildungsstelle, mouilliert und gerundet, Zischlaute, Affrikaten, Schnalze, komplizierte Vokale, Diphtonge usw.“

Da vor der phonologischen Produktion schon eine Fülle, geradezu ein Mehr an artikulatorisch-motorischen Optionen zur Verfügung steht, kann die Artikulationsfähigkeit bei Sprachgesunden nicht als hinreichend explanativer Performanzfaktor gelten. Es ist hilfreich,

sich in diesem Zusammenhang noch einmal die grundsätzliche duale Funktion von phonologischen Merkmalen zu verdeutlichen (vgl. Calabrese 2012). Sie bilden zum Einen ein optimales Hilfsmittel für die menschlichen Gedächtnissysteme, indem sie ein Phonem eineindeutig von einem anderen abgrenzen und mittels des merkmalsbasierten kombinatorischen Prinzips tausende Wörter aus wenigen immer gleichen Teilen zusammengesetzt werden. Zum Anderen stellen sie eine wichtige Steuerungsinformation für Sprachproduktion und –perzeption dar. Im Spracherkennungsmodell von Stevens (2002) kontrollieren phonologische Merkmale die Verbindung zwischen akustischen und artikulatorischen Informationen. So ist z.B. die Vor- und Rückwärtsbewegung des Zungenkörpers mit spezifischen Unterschieden in der Frequenz des zweiten Formanten korreliert. Genau diese Korrelation wird durch das Merkmal [hinten] gesteuert.

Es liegt insofern nahe, die Entwicklung der Diskriminierungs- und Steuerungsfunktionen von Phonemen weniger direkt mit artikulatorisch-motorischen Optionen, sondern vielmehr mit der Etablierung von Korrespondenzregeln von phonologischen Repräsentationen mit denselben zu identifizieren. Da eine solche Etablierung nach dem Modell der Parallelen Architektur (aber auch anderen kognitiven Modellen) nur über die Verarbeitung und Berechnung im Arbeitsgedächtnis erfolgen kann, sollte der Fokus auf diesem Performanzfaktor liegen.

Unter Zugrundelegung der Konzeption von Hale & Reiss sollen in den folgenden Kapiteln eine Untersuchung und weitergehende Reflexion aktueller Modelle und Studien zum Erwerb von Betonungsmustern, Silbenstrukturen und Reduplikationen durchgeführt werden. Wie zu zeigen sein wird, lassen sich viele Phänomene unter einer Forschungsperspektive, die genuin kognitive Performanzfaktoren in die Analyse von Produktionen einbezieht, willkürliche Postulate und Konstruktionen anderer Theorien durchaus vermeiden.

Da ihre Ausführungen vornehmlich um phonologische Merkmale kreisen, werden m. W. Parametermodelle zum Erwerb von Betonungsmustern und Silbenstrukturen von Hale & Reiss nicht erwähnt. In Hinblick auf obige Ausführungen kann man allerdings ähnliche „Verirrungen“ ausmachen, wie an dem Beispiel von Smith (1973) kenntlich gemacht wurde. In Kapitel 2.2.1 soll hierzu exemplarisch die Diskussion um den Ansatz und die Ergebnisse von Fikkert (1994) beleuchtet werden.

2.2 Der Erwerb auf einzelnen Repräsentationsebenen

2.2.1 Der Erwerb von Betonungsmustern

Archibald (1995: 81) weist darauf hin, dass der Fokus der Forschung zum Phonologieerwerb in den Anfängen der generativen Grammatik stark auf der Frage lag, wie segmentale Informationen verarbeitet und in der Produktion erworben werden. M.E. ist dieser Trend seit Beginn der 90'er Jahre zugunsten der Erforschung prosodischer Strukturen umgekehrt worden. Vor allem die Bedeutung der prosodischen Komponente für die frühe Markierung von Wort- und Phrasenrändern fand große Beachtung. Da Kinder in der frühen Produktion unbetonte finale Silben auslassen, argumentieren Echols und Newport (1992) für eine perzeptuelle Grundausstattung, die es ermöglicht, den unsegmentierten Lautstrom nach Maßgabe perzeptuell prominenter Informationen respektive betonter Elemente in Einheiten aufzuteilen. Die Nutzung dieser Grundausstattung wird von den meisten Autoren innerhalb eines theoretischen Rahmens beschrieben, der Baumstrukturen und somit die prosodische Hierarchie voraussetzt. So kann ein Kind nach Penner (2002) mit Hilfe dieser Grundausstattung drei sprach-rhythmische Parameter belegen:

- a) Die Grundbetonungsregel des zweisilbigen minimalen Wortes.
- b) Die Zeitstruktur des zweisilbigen minimalen Wortes.
- c) Die Mechanismen der Wortrandmarkierung.

Das Kind muss nach dieser Theorie also den Trochäus (betont-unbetont) als Grundbetonung des minimalen Wortes im Deutschen erfassen (a), der betonten Silbe innerhalb des trochäischen Musters gemäß der Grundregel der temporalen Organisation mehr Gewicht als der unbetonten zuweisen (Dauer in Moren) (b) und die kritische Rolle des Schwa für die Wortrandmarkierung des trochäischen Fußes erkennen (z.B. Schwa-Einfügung bei konsonantisch auslautenden Wurzeln wie /li:b-/ > li:bə) (c).

Das sprachunauffällige Kind hat die genannten Parameter schon vor der Produktion erster Wörter, genauer während der kanonischen Lallphase (ca. im Zeitraum vom 7. - 10. Monaten (Oller 1980)) korrekt belegt und kann diese Informationen zur Bildung des prototypischen minimalen Wortes nutzen (vgl. Ramus et al. (1999)). Es herrscht keine Einigkeit darüber, wie genau die Struktur des minimalen Wortes gestaltet ist (vgl. Archibald 1995: 95-96). Nach McCarthy und Prince (1990) konstituiert sich das minimale Wort aus einem zweimoraischen Fuß (nur Phoneme innerhalb der Reimkonstituente sind mit einer Mora assoziiert):

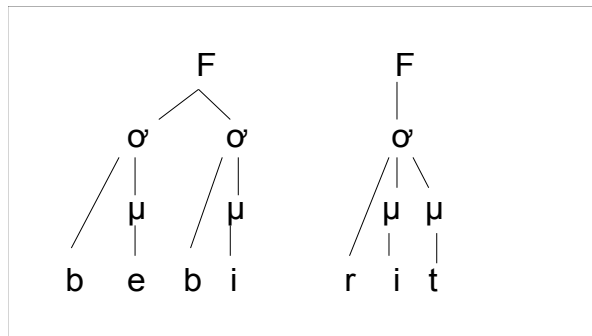
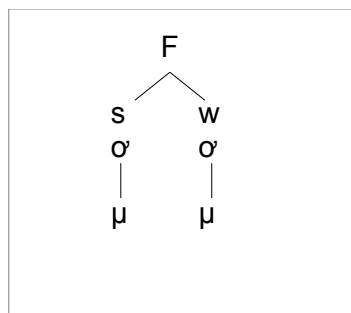


Abbildung 41: Die Struktur des minimalen Wortes.

Dies würde aber die Vorhersage machen, dass sowohl CVCV-Strukturen als auch CVC-Silben in den ersten Produktionen beobachtet werden könnten. Da dies nach Archibald (1995), Demuth (1995), Fikkert (1994) und Fee (1996) nicht der Fall ist, gehen sie davon aus, dass eine Silbe anfangs nur einmoraisch sein kann, das minimale Wort deshalb ein zweisilbiger zweimoraischer Fuß mit trochäischer Betonung ist:



Beginnend mit der Produktion von einfachen CV-Silben vollzieht sich der Erwerb nach Fikkert (Fikkert 1994; vgl. auch Fikkert, Penner & Wymann 1998: 87-88) über folgende Phasen hinweg:

- Phase 1: Nur der Kopf (die betonte Silbe) des prominenten Fußes wird als CV realisiert. (bis 1;7)
- Phase 2: Übergang zu Trochäen. (1;7-2;0)
- Phase 3: Übergeneralisierung des trochäischen Musters. (1;7-2;0)
- Phase 4: Übergang zu mehrsilbigen Wörtern, die allerdings gleichförmig betont werden (level stress: Jeder Fuss wird als prosodisches Wort behandelt). (2;2-2;3)
- Phase 5-6: Übergang von Komposita-betonten (Betonung auf dem linken Fuß) zu endbetonten mehrsilbigen Wörtern (zielsprachliche Wortbetonungsregel bei Wörtern mit einer superschweren Silbe). (Phase 5: 2;2-2;3 / Phase 6: 2;3-2;4)

Zu Beginn der Wortproduktion orientieren sich Kinder an der betonten Silbe des Zielwortes. Bei trochäischem Betonungsmuster der Zielsprache also hauptsächlich an der Anfangs-, bei jambischem Betonungsmuster an der Endsilbe eines zweisilbigen Zielwortes. Da der Reim

der Silbe das Hauptgewicht (in Moren) erhält, somit prosodisch die größte Prominenz repräsentiert, erfolgt der Erwerb der Silbenstruktur sukzessive vom Reim hin zur maximalen Silbenstruktur der Zielsprache.

Ähnlich wie Fikkert nimmt auch Demuth für die anfängliche Stufe I eine reine CV-Form an, die sie als Sub-Minimal Word bezeichnet. Im Unterschied zu Fikkert, die auf Stufe 2 ihres Modells nur von einem generalisierten Muster ausgeht, besteht Stufe II für Demuth aus drei Subphasen. Stufe IIa ist die Phase, in der minimale Wörter produziert werden. Dies muss für einige Sprachen präzisiert werden, da in Sprachen mit ausschließlich offenen Silben die zweimoraische CVCV-Struktur das einzige minimale Wort bleibt, während in Sprachen mit geschlossenen Silben eine solche Struktur nur übergangsweise gebildet wird, bis auf Stufe IIb Kodakonsonanten gebildet werden können (ebd.: 15).

Auf Stufe IIc ist der Vokallängenunterschied etabliert und auf Stufe III werden Prosodische Wörter produziert, die größer sind als ein binärer Fuß. Auf Stufe IV werden schließlich prosodische Wörter zieladäquat produziert:

- Stage I. Core Syllables - CV (No vowel length distinctions) (μ)
- Stage II. Minimal Words/Binary Feet
 - a. Core Syllables - CVCV ($\mu\mu$)
 - b. Closed Syllables - CVC ($\mu\mu$)
 - c. Vowel length distinctions – CVV ($\mu\mu$)
- Stage III. Prosodic Words - larger than a binary foot
- Stage IV. Prosodic Words - target form

Wichtige konzeptuelle Fragen wirft Archibald (1995) in seiner Besprechung verschiedener solcher Modellvorstellungen auf. Mit einer Theorie, die nur auf perzeptueller Salienz basiert (vgl. Echols & Newport 1992), kann nicht erklärt werden, warum unbetonte Silben, die der betonten Silbe vorangehen, häufiger ausgelassen werden als solche, die auf die betonte Silbe folgen. Obige Theorien erfassen zwar mit der Konzeption des minimalen Wortes diese Asymmetrie, setzen aber letztlich voraus, was das Kind aus dem Input herausfiltern muss. D.h., es muss im Grunde ein betonter Fuß aus dem Input extrahiert und dann wiederum auf die UG bzw. auf die Repräsentation des minimalen Wortes abgebildet werden. Archibald (1995: 99) schlägt basierend auf Parsing eine Synthese vor. Nach dieser Annahme werden prosodisch prominente Elemente in der nichtsprachlichen Wahrnehmung aufgrund ihrer Salienz aus dem Input herausgefiltert und auf die UG abgebildet (s. Abb. 41). Erst anschließend können entsprechende Repräsentationen aufgebaut werden.

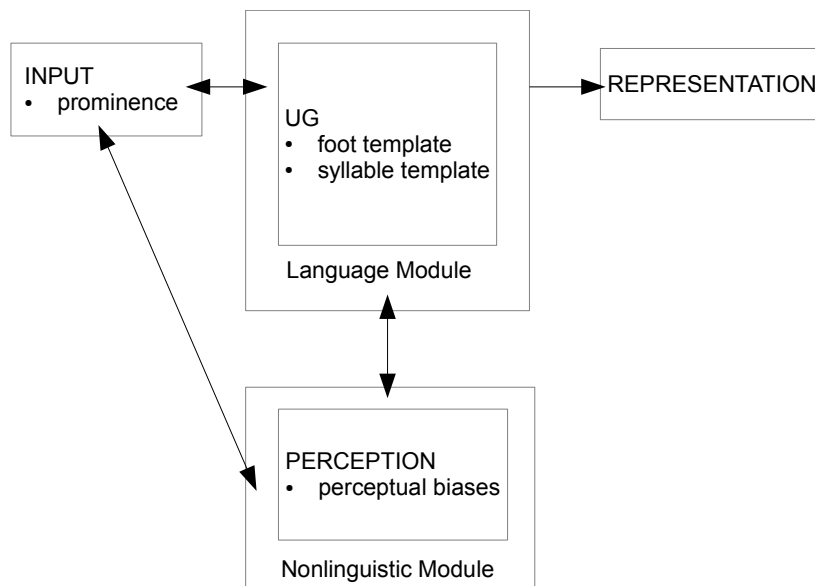


Abbildung 42: Prosodisches Parsing-Modell nach Archibald (1995).

Die Verarbeitung findet in folgenden Schritten statt (nach Archibald 1996: 101-103):

1. Speicherung der gesamten phonetischen Form eines Erwachsenenwortes:

ə n a ǝ ə r

2. Identifikation der betonten Silbe:

ə n 'a ǝ ə r

3. Projektion der betonten Silbe auf Silbenebene

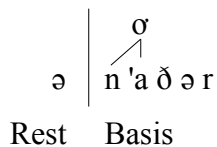
σ
|
ə n 'a ǝ ə r

4. Silbifizierung der betonten Silbe:

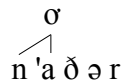
σ
/ \
ə n 'a ǝ ə r

5. Unterteilen der linearen Abfolge. Hier bezieht sich Archibald auf Idsardi (1992), wo eine Unterteilung in Basis (base) und Rest (residue) vorgenommen wird. Diese Unterteilung wird vorgenommen, weil bestimmte Verschmelzungsprozesse nur auf die Basis, eine Anzahl von Tilgungsprozessen allerdings nur auf den Rest appliziert werden. Die Regel ist folgende: Divide a form into two domains at the main stress (ebd.: 43).

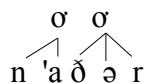
Man erhält nach Anwendung der Unterteilung diese Struktur:



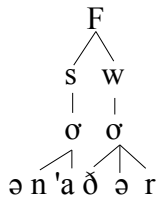
6. Löschen des Rests:



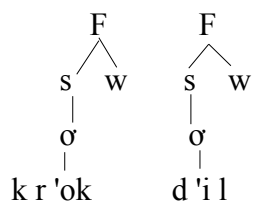
7. Silbifizierung der Basis:



8. Projektion der Basis auf Minimales Wort:



Diese Verarbeitungsprozedur setzt Archibald nun in Relation zu den Erwerbsstufen nach Fikkert. In Phase 1 wird nur die betonte Silbe auf das universelle Fußschema abgebildet, in Phase 2 hingegen die gesamte Basis, so dass ein trochäisches Muster produzierbar ist. Für Phase 3, die Phase, in der zwei Silben gleich betont werden, zieht er vergrößerte Parsingkapazitäten und mit ihnen eine Veränderung von Parsingstrategien in Betracht. Das leitende Prinzip bei der Parsingprozedur in Phase 3 scheint demnach zu sein, dass wenn zwei betonte Silben in der Form als vorhanden erkannt worden sind, sie jeweils auf einen Fuß abgebildet werden müssen. Somit würden folgende Strukturen aufgebaut werden:



Archibalds Analyse hat also den Vorteil, dass sie nicht nur erklärt, warum Silben vor der Hauptbetonung häufiger ausgelassen werden als Silben nach der Hauptbetonung, sondern auch durch die Hinzunahme einer Parsingperspektive das Phänomen der Gleichbetonung

erfassen kann. Dennoch bleibt kritisch zu fragen, ob nicht auch seine Analyse mehr Redundanz als nötig enthält. Denn wenn man eine Parsingprozedur annimmt, die sensitiv für prosodische Prominenz ist, inwiefern braucht man dann noch Fußstrukturen, auf die diese Prominenz abgebildet werden muss? In dem obigen Beispiel, in dem auf Idsardis Theorie³² rekurriert wird, ist das trochäische Muster qua Prominenz schon eingeschrieben, eine Abbildung auf einen Fuß ist somit nicht notwendig. Auch das Phänomen der Übergeneralisierung des trochäischen Musters und der Gleichbetonung lässt sich m.E. auch ohne Fußstrukturen beschreiben. Hierzu muss man sich die einzelnen, von Fikkert beschriebenen, Stufen noch einmal verdeutlichen. Die Leitfrage soll hier sein, wie viele Silben und wie viel an prosodischer Information berechnet werden kann.

In Phase 1 wird nach Fikkert (1994) nur die betonte Silbe des rechten, d.h. prominenten Fußes realisiert. Die einzige Evidenz, die hier Fikkerts Argumentation wirklich stützen könnte, sind die Produktionen von dreisilbigen Wörtern mit initialer Hauptbetonung. Hier wird aber nur die finale Silbe, die die sekundäre Betonung trägt, realisiert. Dies schließt allerdings nicht aus, dass sich die Berechnung dennoch an der grundsätzlichen Parametrisierung orientiert, wenn man Einflüsse von Frequenzeffekten nicht ausschließt. Im Niederländischen liegt die Auftretenshäufigkeit (Tokens) der Hauptbetonung bei dreisilbigen Nomen zu 53% medial, zu 35,3% final und nur zu 11,7% initial (vgl. Schiller et al. 2006: 126). Auch wenn zugrundeliegend also ein initiales Hauptbetonungsmuster vorhanden ist, orientiert sich die Berechnungsprozedur in diesem frühen Erwerbsstadium eventuell an diesen Frequenzmustern und bildet so auf dieser Stufe, in der nur eine Betonung in die Berechnung eingeht, die finale Silbe mit sekundärer Betonung ab.

In Phase 2 können Trochäen, also zwei Silben berechnet werden. Mehrsilber werden, wie oben beschrieben, ab der betonten Silbe analysiert, so dass der Eindruck entsteht, eine trochäische Schablone werde appliziert. Allerdings wird von Wörtern mit der Hauptbetonung am Ende, wie z.B. Kroko'dil oder Ele'fant, auch meist nur diese Silbe produziert. Die Produktion von zweisilbigen Formen wie z.B. Kamel als ['me:.lə] ist gerade nicht die Regel (vgl. Penner et al. 2006: 85), was man bei einer Abbildung auf einen trochäischen Fuß erwarten könnte. Es kann sich deshalb im Grunde nur um eine Füllersilbe handeln, die aufgrund der Links-nach-Rechts-Berechnung bei einer Kapazität von zwei Silben per Default inseriert werden kann oder eben nicht.

³² Diese basiert wesentlich auf Berechnungen mittels metrischer Gitter (vgl. Kap. 1.3.3).

In Phase 3 können zwei betonte Silben der zugrundeliegenden Form mit nur einer hauptbetonten Silbe produziert werden. Man kann in diesem Sinne nicht mehr von einer Übergeneralisierung sprechen. Der Berechnungsvorgang von links nach rechts identifiziert so wie zuvor die erste Silbe als diejenige mit Hauptbetonung. Da in jambischen Formen wie Kamel auch die erste Silbe mit Vollvokal eine relative Prominenz besitzt (zwei Grundschläge, vgl. Kap. 1.3.3), wird dieser hier die Hauptbetonung zugewiesen ['ka.me:l]. Es ist auch auf dieser Stufe, in der eine Interpretation der Daten mithilfe von Fußstrukturen besonders konstruiert wirkt. Nehmen wir die Realisierung von dreisilbigen Wörtern auf dieser Stufe. Nach Fikkert (1994: 219) vollzieht sich diese folgendermaßen:

„The prediction is that at the next stage an additional syllable from the adult target from the left edge of the word (or the most prominent remaining one after the rightmost foot is taken out) is circumscribed, the segmental content of that syllable added to the segmental content of the child's previously circumscribed monosyllabic foot, and the result mapped onto a trochaic template.“

Was von Fikkert in Klammern nur angedeutet wird, scheint der entscheidende Faktor zu sein: Es ist bereits die prosodische Struktur vorhanden und die Berechnung ist sensitiv für diese Struktur. Folgende Beispiele sollen dies noch einmal verdeutlichen.

Dreisilbige Wörter mit finaler Betonung (ebd.: 219):

Tom:
 /,te:le'fo:n/ → ['tefo:] (1;8.6)
 /,kro:ko:'dil/ → ['ko:wi] (1;5.14)

Dreisilbige Wörter mit initialer Betonung (ebd.: 225):

Elke:
 /'o:li:,fant/ → ['o:fant](2;3.27)
 /'ale,ma:l/ → ['amam] (2;3.27)

Dreisilbige Wörter mit medialer Betonung (ebd.: 230):

Leon:
 /pi:'ja:no:/ → ['pa:no:] (1;10.15)
 /spi:'na:zi:/ → ['na:si:] (2;3.4)

Das Systematische an den produzierten Mustern ist, dass eine Orientierung an Hauptbetonung und sekundärer Betonung vorhanden ist und regelhaft die unbetonte Silbe aller zugrundeliegenden Form nicht realisiert wird. Es erscheint nicht plausibel, dass bei einem Aufbau von rechts nach links ad hoc nur die initiale Silbe des Wortes in den Fuß integriert

werden soll. Und wenn Prominenz der entscheidende Faktor ist, dann fügt die Abbildung auf einen Fuß nur Redundanz in den Strukturaufbau hinzu.³³

In Phase 4 werden zwei Silben mit je einer Betonung produziert. Da nun mehrere Silben berechnet werden können und, wie Archibald betont, die Berechnungskapazität größer ist, werden mehrere Betonungen identifiziert und oberflächlich in einem Zweisilber realisiert, was zum relativen Phänomen der Gleichbetonung führt. Erst in Phase 5 können die Unterschiede in der Prominenz zieladäquat produziert werden.

In Tabelle 4 sind die Annahmen obiger Darstellung für die einzelnen Spracherwerbsphasen, die sukzessive Erweiterungen der Berechnungskapazität für silbische und metrische Information sowie der Startpunkt der Berechnung und die Oberflächenrealisierung zusammengefasst.

Phase (nach Fikkert)	Berechenbare Silbenanzahl	Berechenbare prosodische Prominenz	Zuweisungsdomäne (von links nach rechts)	Realisierung
1	Eine Silbe	eine Betonung	nur hauptbetonte Silbe	Einsilber
2	Potentiell zwei Silben	eine Betonung	ab hauptbetonter Silbe	Trochäus
3	Zwei Silben	eine Betonung	ab relativ betonter Silbe	Übeneralisierung
4	Mehrere Silben	mehrere Betonungen	ab relativ betonter Silbe	Level stress
5	Mehrere Silben	Hauptbetonung + Nebenbetonung	Gesamte zugrundeliegende Repräsentation	Zieladäquat

Tabelle 4: Metrisches Erwerbsmodell nach Maßgabe von Prominenz und Berechnungskapazität.

Nach der hier dargestellten Modellierung ist also die metrische Prominenzrepräsentation auf der Basis der Parameter von Halle & Idsardi (1995) schon gegeben und die Produktionen von Kindern spiegeln die sukzessive Erweiterung der Berechnungskapazität wider. Berechnungskapazität bemisst sich hierbei im Rahmen eines 3D-Modells der Phonologie nach der Interaktion von metrischer Ebene (prosodische Prominenzrepräsentation) und

³³ Problematisch sind allerdings die dreisilbigen Muster mit medialer Betonung. Hier wird teilweise die sekundäre betonte Silbe nicht realisiert (vgl. Fikkert 1994: 230). Allerdings scheint, wie die Beispiele zeigen, sekundäre Betonung nicht regelhaft bei diesen Formen gegeben zu sein. Insofern gilt hier die Orientierung an der regelhaft gegebenen Hauptbetonung.

Silbenebene. Sukzessive wird in der Produktion vom Beginn bis zum Ende einer linear geordneten Segmentkette ein Mehr an Silben mit einem einhergehenden Mehr an prosodischen Prominenzrepräsentationen assoziiert. Genau diese Interaktionen erzeugen, wie oben dargestellt, je nach Berechnungskapazität die spezifischen Realisierungsformen, so dass in der Produktion auf eine Parameterbelegung oder Abbildungen auf Repräsentationen der prosodischen Hierarchie (wie ein prosodischer Fuß) verzichtet werden kann. Vorausgesetzt ist bei der Interaktion von metrischer Ebene und Silbenebene im Spracherwerb entsprechend die Integration von Phonemen in Silben. D.h., vor der Produktion von spezifischen Betonungsmustern nimmt Sonorität, die die Assoziation zwischen segmentaler Struktur und Silbenstruktur steuert (vgl. Kap. 1.4.2), eine immens wichtige Funktion ein und muss als Vorläufer angemessen berücksichtigt werden.

Tsakosta (2004: 51) bemängelt am Ansatz von Archibald, dass unklar sei, wie links stehende Elemente (Rest) integriert und Silben auf den Fuß abgebildet werden könnten. Diese Probleme ergeben sich im Grunde daraus, dass Archibald ein hybrides 3D-Modell verwendet, das Cairns (2009) kritisiert. In diesem wird zusätzlich zur metrischen Ebene eine prosodische Ebene postuliert, die ein Set von Repräsentationen der prosodischen Hierarchie beinhaltet. Setzt man allerdings die schon vorhandene Erwachsenenrepräsentation als zugrundeliegend voraus, so kann man auch davon ausgehen, dass perzeptiv schon alle wesentlichen metrischen Parameter auf der Wortebene belegt wurden und (in der Produktion) nicht mehr auf Füße abgebildet werden müssen. Dies zeigen nicht zuletzt die mannigfaltigen Studien zum frühen Spracherwerb (z.B. Höhle 2002). Es ist insofern legitim, für den Erwerb von Betonungsmustern eine starke Kontinuitätshypothese (vgl. z.B. Penner & Weissenborn 1996) zu vertreten. Alle repräsentationale Information ist im ungestörten Spracherwerb in einem dreidimensionalen Format vorhanden und die sukzessive Entfaltung in der Produktion folgt dem Prinzip der minimalen Struktur (Weissenborn 2000). Dieses besagt,

„dass das Kind, solange es nicht die Anwendungsbedingungen einer zielsprachlichen Regel herausgefunden hat, vorzugsweise eine Äußerungsstruktur wählt, die unter Anwendung des zu diesem Zeitpunkt schon erworbenen zielsprachlichen grammatischen Wissens, den geringsten Aufwand an morphosyntaktischen Prozessen erfordert.“ (ebd.: 158)

Es geht nach diesem Prinzip also im Wesentlichen um die Verringerung der kognitiver Komplexität jeweiliger Sprachproduktionsaufgaben (vgl. ebd.: 159) und es spricht m.E. nichts dagegen, es auch auf die phonologische Domäne zu übertragen. Bildlich gesprochen, kann man sagen, dass die Produktion mit den Krücken der Performanz langsam und in sehr kleinen

Schritten dem repräsentationalen Pfad nachgeht, der durch die wesentlich schnelleren Staffelläufer Universalgrammatik und Perzeption schon lange bewältigt wurde.

Verwendet man nur die Repräsentationen metrischer Gitter, so lässt sich einerseits die an Prominenz orientierte Berechnung zeigen und andererseits die Interaktion mit einer unabhängigen Silbenebene darstellen. Nur bei einer Kapazität zur Berechnung von mehr als zwei Silben ist es möglich, primäre und sekundäre Betonung zu differenzieren.

Eine sehr ähnliche Kritik von Fikkerts Ansatz, wenn auch unter modelltheoretischer Perspektive nicht sehr ausgeformt, findet sich in Taelman & Gillis (im Druck). Sie fanden in einer Reanalyse von Fikkerts Datenkorpus, dass die Annahme einer Abbildung auf einen trochäischen Fuß selbst durch ihre Daten nicht gestützt wird. Dies wird an den beobachtbaren Tilgungsmustern deutlich. Wenn auch Tilgung in Wörtern mit der Betonung wS häufiger stattfindet als in Wörtern mit der Betonung Sw, so findet man entgegen einer zu erwartenden uniformen Tilgungsrate in dreisilbigen Wörtern (Betonung: wSw, Ssw, Sws, swS) sehr ungleiche Verteilungen. So findet man z.B. für das Muster wSw eine sehr hohe Tilgungsrate, während bei dem Muster Ssw extrem wenige Tilgungen vorkommen (vgl. ebd.: 24). Zudem ergab sich bei den Tilgungen eine signifikante Rate an jambischen Formen (bei zweisilbigen Formen 166 der 791 Tilgungen; bei dreisilbigen Formen 93 der 436 Tilgungen (ebd.: 33)), die nach Fikkert gar nicht vorkommen dürften. Taelman & Gillis interpretieren die Daten anhand der Acoustic Prominence Hypothesis (Lohuis-Weber & Zonneveld 1996), die besagt, dass Silben mit geringer akustischer Prominenz, d.h. schwache und nicht-finale Silben, getilgt werden. Sie folgen also den selben Prinzipien wie denen von Echols & Newport (1992, s.o), zeigt dennoch aber die Schwächen von Fikkerts Analysen auf.

Abschließend soll an dieser Stelle noch einmal explizit gemacht werden, was in obiger Darstellung implizit enthalten war. In Fikkerts Ansatz ist im Grunde die kompetenztheoretische Vorstellung einer Parameterbelegung enthalten, die sich unmittelbar in der Performanz zeigt. Dies ist umso bemerkenswerter, als dass die Unterscheidung von Performanz und Kompetenz von Fikkert diskutiert wird. So macht sie folgendes deutlich (Fikkert 1994: 15):

„[...] I assume that the child's (production) representations are initially very limited, and gradually expand along parametric dimensions. In other words, I assume continuity in the course of development guided by UG. Production is constrained by performance limitations, and by incomplete representations. The former reflects performance, the latter competence. I am mainly concerned with the latter, although the former cannot completely be ignored.“

Performanzfaktoren sind für Fikkert aber anscheinend nur direkt beobachtbare Phänomene, wie z.B. Imitationen. Außer solchen (recht marginalen) Phänomenen reflektiert Performanz für sie direkt Kompetenz (ebd.: 14):

„I therefore assume that performance reflects competence, unless performance factors can be identified, as is, for example, the case with immediate imitations.“

Diese theoretische Herangehensweise hat zur Folge, dass Perzeption und Produktion in ihrem Modell nicht adäquat gesondert werden. Obwohl sie die Differenzierung im obigen Zitat vorgenommen hat, werden die untersuchten Produktionen der Kinder als direkte Reflexion der Kompetenz gewertet werden (ebd.: 14):

„I therefore assume that the child has a single representation underlying both perception and production.“

Im Sinne dieser theoretischen Vorentscheidung (oder etwas provokanter formuliert: dieser rhetorischen Figur) können nun alle Produktionen bei Bedarf direkt auf die Belegung von Parametern der UG zurückgeführt werden. Untersuchungen zum perzeptiven Erwerb von Betonung werden nicht korrektiv in diese Modellierung integriert, obwohl diese zu den frühesten Studien zur Entwicklung der kindlichen Sprachwahrnehmung gehören (vgl. auch Höhle 2002). So haben z.B. schon Spring & Dale (1977) die Fähigkeit von 1-4 Monate alten Säuglingen festgestellt, bei Zweisilbern trochäische und jambische Betonung grundsätzlich zu unterscheiden. Auch in späteren Publikationen hält Fikkert die kompetenztheoretische Perspektive durch (z.B. Fikkert 2007: 540), obwohl die Erforschung des perzeptiven Prosodieerwerbs in der Zwischenzeit einen regelrechten Boom erlebt hat (vgl. Schröder & Höhle 2011). Als Beispiele mögen zwei Studien gelten. Jusczyk et al. (1999) fanden, dass 7-8 Monate alten Englisch lernende Säuglinge trochäische Zweisilber in einem Satzkontext wiedererkennen, jambische Zweisilber aber nicht. Höhle et al. (2009) fanden bei Deutsch lernenden Säuglingen im Alter von 6 Monaten eine klare Präferenz für trochäische Nichtwörter, die bei 4 Monate alten Säuglingen noch nicht gegeben ist. D.h., ungefähr 15 Monate, bevor Kinder die ersten trochäischen Wörter produzieren, ist perzeptiv die Fähigkeit zur phonologischen Differenzierung schon entwickelt. Ein Ignorieren dieser empirischen Befunde führt letztlich konzeptuell zu einer groben Fehlinterpretation der Produktionsdaten, indem Einflüsse von Performanzfaktoren in unkritischer Weise als Kompetenzphänomene klassifiziert werden.

In Anbetracht aller dargestellten Widersprüche und Redundanzen bietet hier also ein auf der dreidimensionalen Phonologie basierendes, psycholinguistisches Sprachproduktionsmodell, das eine klare Trennung von Performanz (im Sinne von Berechnungskapazität) und Kompetenz beinhaltet, für die modelltheoretische Differenzierung als auch für die eigentliche Dateninterpretation wesentliche Analysevorteile und mit dieser auch eine deutlich größere explanative Reichweite. Wurde in diesem Kapitel die metrische Ebene der dreidimensionalen Phonologie für den Erwerb beleuchtet, so soll nun weiter untersucht werden, inwiefern auch der Erwerb auf der Silbenebene unter der Anwendung einer solchen Perspektive feinkörniger bestimmt werden kann.

2.2.2 Erwerb der Silbenstruktur

Es gibt eine anhaltende Debatte darüber, in welcher Reihenfolge CV- oder X-Repräsentationen erworben werden. Das zur Zeit wohl prominenteste Erwerbsmodell der Silbe ist wiederum das von Fikkert (1994). Dieses Modell greift, wie in Kapitel 2.2.1 angedeutet, auf die Idee der Parameterbelegung zurück, welche Chomsky (1981) im Rahmen seiner Principles and Parameters Theory entworfen hat. Die Idee der Universalgrammatik (UG) ist zusammengefasst, dass dem Kind im Spracherwerb ein angeborenes Repertoire von Prinzipien und offenen Parametern zur Verfügung steht und diese unbewusste Kenntnis dem Kind bei unzureichender Datenmenge und Datenqualität im Input und der daraus folgenden Vielzahl deskriptiv möglicher Grammatiken verhilft, „mit schlafwandlerischer Sicherheit in das System der Muttersprache“ (Weissenborn 1996: 96) einzusteigen.

Das Fortschreiten im Erwerb erfolgt zum einen nur aufgrund von positiver Evidenz, welche die Parameter triggert (wobei die Frequenz von parameterrelevanten Inputinformationen irrelevant ist (vgl. Meisel 1995: 20-21)). Zum anderen besteht ein uniformer Verlauf des Erwerbs, in welchem eine Kontinuität in dem Sinne vorhanden ist, als dass jede Erwerbsstufe eine mögliche Grammatik definiert und diese Zwischengrammatiken hinsichtlich der Repräsentationen mit der Zielgrammatik übereinstimmen. Unterschiede zwischen den Grammatiken sind auf noch nicht belegte Parameter oder eben Einflüsse von Performanzfaktoren (z.B. Gedächtnis, Artikulation, Aufmerksamkeit) zurückzuführen.

Das Modell von Fikkert (1994) ist auf der Basis einer Longitudinalstudie mit 12 Holländisch sprechenden Kindern entstanden. Nachdem die drei sprach-rhythmischen Parameter belegt wurden, startet die Wortproduktion mit der Realisierung der prosodisch prominentesten

Elemente auf Silbenebene, d.h., mit den betonten Silben der zielsprachlichen Wörter (vgl. Kap. 2.2.1). Vier Stufen des Silbenerwerbs werden von Fikkert für den Ausbau des Silbenreims beschrieben, welche sich mit der Belegung von drei Parametern assoziieren lassen.

Stufe 1 (bis 1;6.13) ist gekennzeichnet durch die Realisierung einfacher CV-Kernsilben. Koda-Konsonanten werden auf dieser Stufe nicht realisiert. Ein Zusammenhang dieser Auslassung von Koda-Konsonanten und Vokallängenrepräsentation kann auf dieser Stufe nicht attestiert werden. Langvokale wie auch Kurzvokale werden in unsystematischer Weise kurz und lang realisiert. Fikkert folgert daraus, dass die phonologische Vokallängenunterscheidung auf dieser Stufe noch nicht etabliert ist und Vokale weiterhin mit nur einem V auf der CV-Ebene assoziiert sind.

Auf **Stufe 2** erscheinen silbenfinale Frikative (1;6.13) und Plosive (1;7.15). Frikative werden einerseits zieladäquat als auch substituierend für andere Lautklassen verwendet. Sonoranten werden noch nicht produziert, sondern im Normalfall ausgelassen oder, weitaus seltener, durch Frikative ersetzt. Die Rate der Vokallängenfehler in der Produktion (39%) lässt weiterhin darauf schließen, dass auf dieser Stufe ebenfalls noch keine systematische Vokallängenunterscheidung vorhanden ist.

Charakteristisch für **Stufe 3** ist das Auftauchen von silbenfinalen Nasalen (1;7.29), welche aber immer noch häufig ausgelassen werden. Das Gleiche trifft auch auf Liquide zu, welche nun ab und an realisiert (1;11.20), aber ebenfalls mit großer Häufigkeit noch durch Laute anderer Merkmalsklassen ersetzt werden. In dieser Phase ist eine wichtige Tendenz in der Auslassung silbenfinaler Konsonanten zu beobachten. Wenn geschlossene Silben mit finalen Sonoranten produziert werden, bleibt die Vokallänge der zieladäquaten Form entweder ohne Sonoranten erhalten (bei langem Vokal) oder erhält eine Kürzung mit Sonoranten bzw. eine Längung ohne Sonoranten. Es scheint also eine Verbindung zwischen Sonorantenproduktion und Vokallänge auf dieser Stufe zu geben. Für Obstruenten gilt dies nicht. Da die Vokallänge immer noch unsystematisch realisiert wird, scheint diese Lautklasse die Vokallängenunterscheidung auf dieser Stufe nicht zu triggern. Fikkert folgert aus diesen Beobachtungen (ebd.: 6):

„We may therefore conclude that vowel length is non-distinctive before obstruents: all vowels are monopositional in this position. If this is correct, than the child has a (minimal and maximal) bipositional rhyme template at this stage.”

Dieses Muster der silbfinalen Obstruentenproduktion ändert sich in **Stufe 4** (2;2.6). Man kann nun Ähnlichkeiten zu der Sonorantenproduktion auf Stufe 3 feststellen.

Es ist nun eine Stufe erreicht, auf welcher das Kind eine Verbindung zwischen Vokallänge und finalen Obstruenten herstellt. Vokallängenfehler sind nun in jeder lautlichen Umgebung selten, was darauf schließen lässt, dass die Vokallängenunterscheidung etabliert ist. Weiterhin ist nun die Form VVC_{son} lizenziert, was auf den Erwerb einer extra-reimischen Position schließen lässt.

Auf Basis der beschriebenen Erwerbsmuster bezieht Fikkert (vgl. ebd.: 8-11) die Silbenentwicklung auf die Belegung von spezifischen Parametern:

1. CV-Core-syllable (stage 1)
2. Branching rhyme parameter (stage 2):
Rhymes can branch into a nucleus and a coda [No/Yes]
3. Branching nucleus parameter (stage 3):
The nucleus can be branching [No/Yes]
4. Extrarhymal parameter (stage 4):
A (final) bipositional rhyme can be followed by an extra consonant.
[No/Yes]

Das Kind erwirbt in dieser Reihenfolge sukzessive Silben mit größerem Gewicht. Auf Stufe 1 werden universal CV-Silben, auf Stufe 2 und 3 schwere Silben und auf Stufe 4 superschwere Silben gebildet (vgl. Penner 2000).

Der Erwerb des Silbenonsets vollzieht sich anfänglich in überschaubarer Form. Auch hier unterscheidet Fikkert wiederum Stufen des Erwerbs, die hier allerdings nur für monopositionale Onsets beschrieben werden sollen.³⁴

Auf **Stufe 1** produzieren Kinder einen obligatorischen Onset, der invariabel mit Plosiven realisiert wird. Auf **Stufe 2** produzieren Kinder optional Zielwörter ohne Onset. Zielwörter mit Onset-Konsonanten werden weiterhin nur mit einem Plosiv in dieser Position produziert. Später auf dieser Stufe wird /h/ als Substitut für einen leeren Onset oder noch nicht produzierbare Frikative und Gleitlaute verwendet. Auf **Stufe 3** schließlich werden auch Konsonanten mit anderen Artikulationsartmerkmalen realisiert. Hier unterscheidet Fikkert noch einmal zwischen Stufe 3a, auf der als erste Kontrastgruppe Nasale und Plosive realisiert werden, und Stufe 3b, auf der variabel Frikative, Liquide und Gleitlaute in der Produktion auftreten. Die Abfolge ist in Tabelle 5 zusammengefasst.

³⁴ Clusterbildungen treten erst später im Erwerb auf und sind für den frühen Spracherwerb insofern nicht von Relevanz.

Stufe 1	(-sonorant)	(-kontinuierlich)	Plosive
Stufe 3a			Nasale
Stufe 3b	+sonorant	+kontinuierlich	Andere Artikulationsartmerkmale

Tabelle 5: Merkmalswerb im Onset nach Fikkert (1994).

Eine häufig diskutierte Frage, auch in Bezug auf Fikkerts Modell, ist, ob die frühe Wortproduktion mit CV-Silben beginnt oder schon mehr Positionen angenommen werden können. Für die Beantwortung dieser Frage müssen auch Daten aus anderen Sprachen hinzugenommen werden. Einen ausführlichen Einblick in die Reimentwicklung im Deutschen gibt z.B. eine Studie von Kehoe & Lleó (2003). Durch eine Spontansprachuntersuchung von drei deutschsprachigen Kindern (1;3-2;6), in welcher sowohl ein- wie zweisilbige Wörter analysiert wurden, konnten Einzelheiten von Fikkerts Modell nicht bestätigt werden. Schon die ersten Produktionen des Kindes Marion entsprachen eher Stufe 3 des Fikkert-Modells, mit dem Unterschied, dass sie Konsonanten nach langen Vokalen ausließ und nach kurzen Vokalen Konsonanten produzierte, unabhängig davon, ob die Koda durch einen Obstruenten oder Sonoranten besetzt war. Daraus schlossen Kehoe & Lleó, dass in der ersten Stufe eine Beschränkung in Form eines bipositionalen Reimes vorliegt. Es ist unklar, ob die Vokallängenunterscheidung auf dieser Stufe schon etabliert ist oder ob die bipositionale Beschränkung die adäquate Reproduktion der phonologisch repräsentierten Distinktion verunmöglicht. Auch wenn auf der frühesten Stufe die Vokallänge korrekt erhalten bleibt, ist auf der folgenden Stufe die Akkuratheit in diesem Sinne nicht mehr vorhanden, so dass die Vokallängenunterscheidung nicht stabil scheint.

Zusammengefasst stellt sich der Silbenerwerb nach Kehoe & Lleó so dar (ebd.: 549):

Stages of vowel length acquisition

Early	Bipositional rhymes			
	Target VVC	Target VC		
	Child VV	Child VC		
Middle	Beginning of opposition V vs. V (V) or Opposition V vs. VV			
	Target VVC	Target VC	Target VVC	Target VC
	Child VC or VVC	Child VC	Child VVC	Child VC
Later	Adult-like opposition V vs. VV			
	Target VVC	Target VC		
	Child VVC	Child VC		

Aus den Ergebnissen kann auch geschlossen werden, dass die bei Fikkert postulierte universelle CV-Kernsilbe im Deutschen nicht das initiale Stadium darstellt, sondern schon zu Beginn der Wortproduktion die Bipositionalität des Reims vorgegeben und somit CVV oder CVC obligatorisch wäre. Es scheint in Anbetracht der Daten interindividuelle Unterschiede in der Ausgangsstruktur als auch in der Ausbaurichtung derselben zu geben. Für das Englische hat Ingram (1978) den frühen Erwerb von Silbentypen anhand eigener Daten als auch in einer Reanalyse des Datenkorpus von Weir (1962) und der Gruppenstudie von Winitz & Irwin (1958) untersucht. Letztere zeigt, dass im Alter von 1;0 – 1;1 großteils einfache CV-Silben (18%) als auch zweisilbige CV-Muster produziert werden (insgesamt 49%). CVC-Silben werden erst ab 1;2 – 1;3 vermehrt produziert.

Silbentypen	Alter	Altersgruppen		
		1;0 – 1;1	1;2 – 1;3	1;4 – 1;5
	Anzahl Kinder	23	35	35
	Anzahl Wörter	137	224	268
CV		.18	.44	.34
CVC		.05	.20	.16
C ₁ V ₁ C ₁ V ₁		.18	.05	.09
C ₁ VC ₁ V ₂		.03	0	0
C ₁ VC ₁ V oder CV ₁ CV ₁		.31	.16	.28
C ₁ VC ₁ VC ₁ oder CV ₁ CV ₁ C		0	.01	.01
Total		.76	.87	.89

Tabelle 6: Silbentypen im Datenkorpus von Weir (nach Ingram 1978).

In der Untersuchung der Sprachproduktionen seiner eigenen Tochter, Jennika, im Alter von 1;3 – 2;3 konnte Ingram diese Muster ebenfalls bestätigen. 89% der einsilbigen Wörter waren im Alter von 1;3 CV-Silben und die einzigen Formen zweisilbiger Wörter CVCV-Strukturen. CVC-Silben werden schon ab 1;3 produziert (11%) und werden ab 1;5 häufiger produziert (30%), so dass sie bereits mit 1;6 die am häufigsten produzierte Silbenstruktur darstellen (66%). Ab 1;6 werden CVCC-Strukturen produziert, so dass der Ausbau des Silbenreims vor dem Ausbau des Onsets erfolgt.

Silbentyp	Alter													Häufigkeit Total
	1;3	1;4	1;5	1;6	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2	2;3	
Einsilbige Wörter														
CV	.89	.67	.42	.14	.11	.14	.14	.14	.14	.13	.10	.05	.08	.16
CVC	.11	.11	.30	.66	.66	.66	.68	.71	.70	.79	.74	.70	.77	.65
VC		.11	.15	.15	.11	.06	.10	.02	.01	.03	.03	0	0	.07
CVCC				.02	.05	.06	.05	.02	.10	.03	.10	.15	.12	.05
Andere		.11	.13	.03	.06	.06	.05	.10	.05	.03	.03	.10	.04	.07
Total	9	9	53	59	94	71	111	59	74	71	39	20	26	695
Zweiselbige Wörter														
CVCV	.87	.63	.62	.47	.50	.48	.39	.25	.21	.36	.46	.40 ³⁵	.14	.40
CVCVC			.05	.26	.17	.11	.28	.33	.34	.26	.24	.20	.21	.29
VCV		.37	.16	.08	.07	.07	.05	.05	.01	.02	.03	0	.14	.06
VCVC			.08	.08	.10	.16	.09	0	.01	.09	.03	.07	0	.06
CVCC				.04	.08	.11	.07	.10	.09	.05	.11	.14	0	.07
CVCCVC								.03	.10	.11	.03	0	.38	.04
Andere	.13	0	.09	.07	.08	.07	.12	.24	.24	.11	.10	.19	.13	.08
Gesamt	8	16	37	53	60	44	67	61	70	66	37	15	14	548
Gesamtzahl Wörter														1243

Tabelle 7: Silbentypen im Erwerb von Jennika (nach Ingram 1978).

Dies deckt sich mit den Daten von Weir (1962; nach Ingram 1978: 145-146), in deren Datenkorpus aus 5.130 Wörtern ihres Sohns Anthony im Alter von 2;4-2;6 sich einsilbige Wörter vorwiegend aus CVC-Silben zusammensetzten (36%) und eine maximale CVCC-Struktur (10%) zeigten. Allerdings sind hier zusätzlich in zweiselbigen Wörtern komplexe Onsets mit zwei C-Positionen gegeben (9%) (ebd.: 146):

Einsilbige Wörter	Häufigkeit	Zweiselbige Wörter	Häufigkeit	Dreiselbige Wörter	Häufigkeit
CVC	.36	CVCV	.35	VCCVCV	.52
CV	.22	CVCVC	.22		
VC	.18	CCVCCVC	.09		
CVCC	.10				

Tabelle 8: Silbentypen im Erwerb von Anthony (nach Ingram 1978).

³⁵ In Ingram (1978) befindet sich an dieser Stelle ein Eintragungsfehler mit der Angabe 4.0. Dieser wurde hier nach Maßgabe der anderen Werte zu .40 korrigiert.

Dass Variationen im strukturellen Ausbau interindividuell eher die Regel als die Ausnahme sind, zeigt besonders eine Studie von Levelt u.a. (2000). Sie reanalysierten die Daten der zwölf von Fikkert (1994) untersuchten Kinder mit Fokus auf die Erwerbsreihenfolge der CV-Positionen. Das Holländische weist eine Silbenschablone der Form (C) (C) V (C) (C) auf. Dies erwerben die Kinder bis zur Form VC uniform, danach spaltet sich die Erwerbsreihenfolge derart auf (ebd. 293):

Gruppe A: CVCC → VCC → CCV → CCVC
 CV → CVC → V → VC CCVCC
 Gruppe B: CCV → CCVC → CVCC → VCC

Wir haben also Gruppe A, die Koda-Cluster vor Onset-Cluster erwirbt und Gruppe B, die in umgekehrter Reihenfolge vorgeht. Hinsichtlich der Reihenfolge des Erwerbs von Silbenpositionen gibt es also anscheinend auch eine gewisse interindividuelle Variation.

Elsen (1991) hat in einer detaillierten Tagebuchstudie die phonologische Entwicklung ihrer Tochter Annalena im Deutschen beschrieben. Wir finden hier Muster wie in Gruppe A der Studie von Levelt et al. (2000). Annalena fing mit CV-Reduplikationen an (nur eine Form zeigte eine CVC-Struktur) und das Silbenskelett baute sich schrittweise zum Endrand hin aus. Nachdem der Reim strukturell erweitert war, wurde nun der Onset vergrößert. Die einzelnen Erwerbsschritte sind in Tabelle 5 mit den jeweiligen Produktionen zusammengefasst (vgl. Dümig & Leuninger 2013):

Monat	Struktur	Beispiele
8	CVC (erste Form; verschwand wieder, nur bei einem Wort gegeben)	nein
9	CV(CV)	[mama], Mama; [da], da
10	CVCV	[dabe], bitte; [data], das da; tedə, Teddy
11	CVC CVCV	[bem], Baum; [vava], Hund
14	CVCC CCVC	[momp], [momf], plumps; [aiz], [ainz], eins; [bRaw], Frau
15 (nach Elsen mehr Endränder und Abnahme der Protowortproduktion)	CVCC	[mons], Mond; [balç], falsch; [finth], Kind
18	CVCC(X)	Plumps; brauchst
19	CCVCC(X)	Stinkst; [bRoRst] (Brust); [kvitst] (quitscht); trägt
21	CCVVC(X) (Diphthonge)	[kroits] Kreuz; [SRaip], schreibt
22	(X)CCVCC	Strikt

Tabelle 9: Zusammenfassung von Annalenas Erwerb der Silbenstruktur.

Im Erwerb des Deutschen werden keine onsetlosen V- oder VC-Silben gebildet (für eine andere Darstellung vgl. Grijzenhout & Joppen (2002), so dass hier direkt nach dem Erwerb der CVC-Struktur die Optionen des Reim- oder Onsetausbaus zur Verfügung stehen (vgl. Dümig & Leuninger 2013):

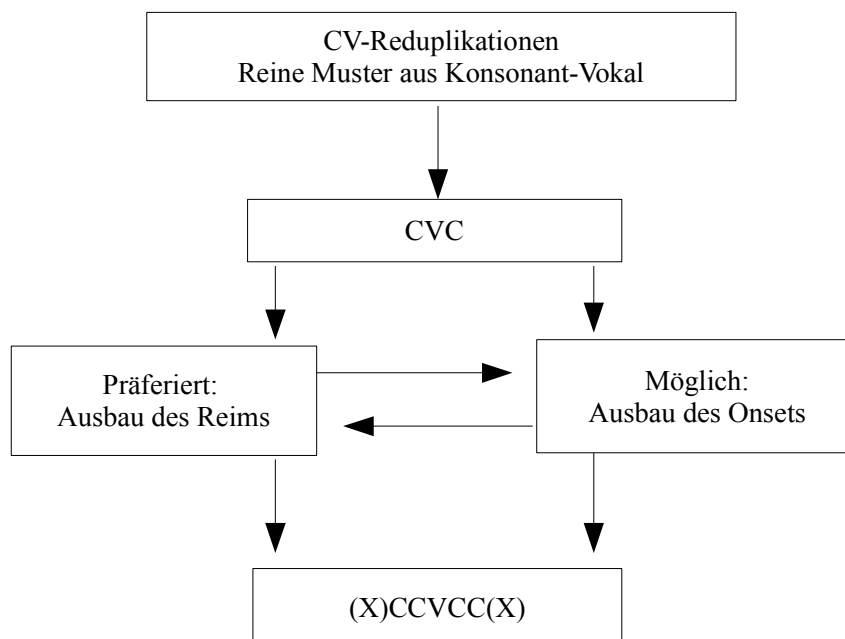


Abbildung 43: Erwerbsreihenfolge und Optionen im Silbenstrukturausbau des Deutschen und Englischen.

Am Ende dieses Kapitels stellt sich wie in Kapitel 2.2.1 die Frage, wieso für die Produktion von Silbentypen eine Parameterbelegung angenommen werden sollte, deren frühe Belegung anhand von Studien diskriminativer Fähigkeiten von Säuglingen und Kleinkindern in der Sprachwahrnehmung relativ sicher scheint. Acht Monate alte Kinder z.B. nutzen schon die Auftretenshäufigkeit von Silben, um Wortgrenzen zu identifizieren (Saffran et al. 1996)³⁶ und bereits mit neun Monaten werden schon phonotaktische Regularitäten der Zielsprache erkannt und ebenfalls für die Identifikation von Wortgrenzen genutzt (Jusczyk et al. 1994). Wenn man nun in Betracht zieht, dass das Kind Jarmo (Fikkert 1994) bis zum Alter von 1;6.13 Jahren nur CV-Silben produzierte und somit erst ab diesem Alter eine mit Frikativen gefüllte Kodaposition realisierte, erscheint die Annahme einer Parameterbelegung zu diesem Zeitpunkt sehr abwegig. Fikkert müsste gemäß ihrer eigenen Argumentation, nach der die Kompetenz des Kindes durch noch unvollständige Repräsentationen beschränkt ist (vgl. Kap. 2.2.1: 112), im Grunde eine doppelte Parameterbelegung, eine für die Sprachproduktion und eine für die Sprachperzeption, annehmen, um im Rahmen ihres Modells kongruent zu bleiben. Da eine solch doppelte Parameterbelegung und die mit ihr einhergehende unökonomische Redundanz wenig Sinn macht, sollte man m.E. die Produktionsdaten konsequent im Rahmen der Strong Identity Hypothesis interpretieren. Die Produktionsdaten sind so hauptsächlich als eine Oberflächenrealisierung der durch Performanz (im Sinne von Berechnungskapazität) beschränkten vollen Kompetenz zu analysieren. Wenn man in diesem Kontext Weissenborns Prinzip der minimalen Struktur (vgl. Kap. 2.2.1), in dem es um die Verringerung der kognitiven Komplexität jeweiliger Sprachproduktionsaufgaben geht, konsequent folgt, dann stellt sich weiter die Frage, wie eine solche Verringerung im Sprachsystem durch die spezifische Erwerbsreihenfolge von Silbenpositionen bewerkstelligt wird. Im folgenden Kapitel soll gezeigt werden, dass mithilfe der Demisilbentheorie von Clements (1990) der Silbenpositionserwerbs anders modelliert werden kann. Die Berechnung von (kognitiver)

³⁶ Saffran et al. (1996) interpretieren diese Ergebnisse in Bezug auf statistisches Lernen (SL), das bei der Wortsegmentierung operieren könnte. Übergangswahrscheinlichkeiten (ÜW) zwischen zwei adjazenten Silben A und B bilden den Indikator dafür, wo ein Wort anfängt, wobei die Übergangswahrscheinlichkeit $\dot{U}W(A \rightarrow B) = W(AB) / W(A)$. $W(AB)$ ist die Häufigkeit der Folge von der Silbe B auf Silbe A und $W(A)$ die Gesamthäufigkeit von Silbe A. So kann z.B. der Lerner der englischen Sprache in einer viersilbigen Sequenz *prettybaby* anhand $\dot{U}W(\text{pre} \rightarrow \text{tty})$ und $\dot{U}W(\text{ba} \rightarrow \text{by})$ ausmachen, dass beide höher sind als $\dot{U}W(\text{tty} \rightarrow \text{ba})$. Eine Wortgrenze, welche im SL ein lokales Minimum von $\dot{U}W$ darstellt (Das SL-Modell, welches lokale Minima verwenden, wird weiterhin mit SLM bezeichnet), ist deshalb korrekter Weise zwischen *pretty* und *baby* anzunehmen. Yang allerdings sieht in der erfolgreichen Anwendung von SLM kein Argument gegen angeborene sprachliche Kenntnis: „The fact that infants can use SLM at all entails that, at minimum, they know the relevant unit of information over which correlative statistics is gathered; in this case, it is the syllable, rather than segments, or front vowels, or labial consonants.“ (2004: 452)

Komplexität kann als maßgeblich einbezogen und so Widersprüche des Fikkert-Modells grundlegend vermieden werden.

2.2.2.1 Demisilbentheorie und der Erwerb von Silbenpositionen

Da im Niederländischen eine minimale und maximale Silbe der Form (C) (C) V (C) (C) gegeben ist, sprich alle positionalen Optionen von Konsonant-Vokal-Folgen möglich sind, lassen sich die Einflüsse von phonologischer Komplexität anhand der detaillierten Studie von Levelt et al. (2000) ohne größere Formalisierung elegant erfassen und neu interpretieren.

Es ist anzunehmen, dass jede neu zu erwerbende Silbenposition neue Verarbeitungsanforderungen an das Sprachproduktionssystem stellt, d.h., mehr kognitive Ressourcen erfordert als schon etablierte Positionen. Vokale sind als Nuklei von Silben primär gegeben und zeigen nach der Demisilbentheorie von Clements (1990) unveränderliche Komplexitätswerte, d.h., diese sind unabhängig vom jeweils realisierten Vokal:

Initial: 5
Final: 0

Finale Vokale werden also immer mit geringer Komplexität wenig Ressourcen beanspruchen, während sie initial, d.h., ohne vorangehenden Konsonanten, mit hoher Komplexität auch hohe Systemanforderungen stellen. Wir können also folgendes Schema aus positionaler Ressourcenbeanspruchung und Vokalkomplexität ableiten (+ steht für hohe, - für niedrige Systemanforderung):

	Initialer Vokal	Finaler Vokal
Neue Position	++	+-
Alte Position	-+	--

Tabelle 10: Systemanforderung nach Maßgabe von Positionen und Vokalkomplexität.

Wichtig ist, dass in diesem Schema die neuen Positionen und initiale Vokale größere Anforderungen stellen, für den Erwerb neuer Gesamtstrukturen aber die Interaktion von Neuheit der Position und Vokal relevant ist. Wenn wir dieses Bewertungsschema auf die Ergebnisse von Levelt et al. (2000) anwenden (vgl. Kap. 2.2.2), erhalten wir die Muster in Abbildung 44.

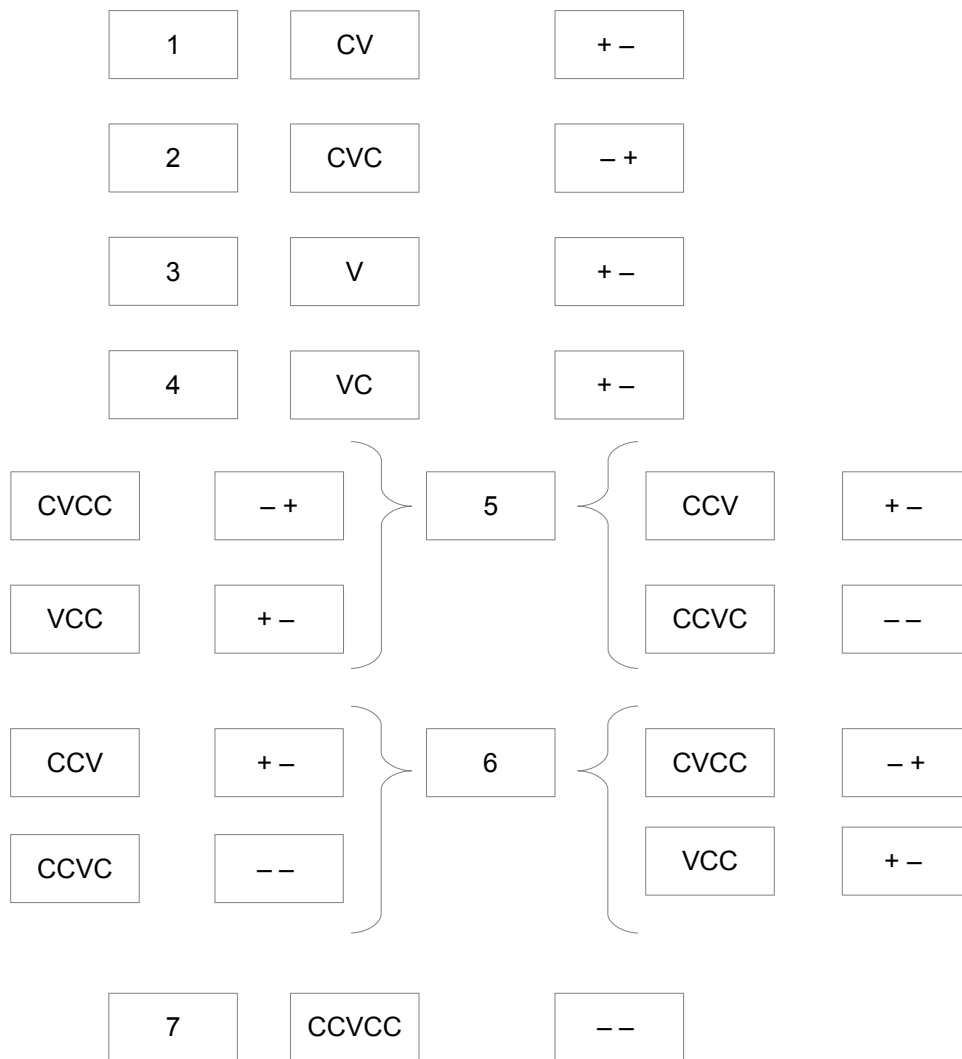


Abbildung 44: Erwerbsreihenfolge von Silbenpositionen mit Kennzeichnung von Ressourcenbeanspruchung.

Eindeutig ist zu erkennen, dass auf jeder Erwerbsstufe ein Maximum an Komplexität vermieden wird. Neue Position werden unter Einbezug der Komplexität von finalen Demisilben, spezifischer der Komplexität von Vokalen erworben. Besonders eindeutig erkennt man diese Gesetzmäßigkeit auf Stufe 5 in der linken Spalte bzw. Stufe 6 in der rechten. Nicht die reine Länge von Silben machen Formen komplex. Wäre dem so, müsste VCC vor CVCC erworben werden. Nach oben dargestellter Regel würde bei vorherigem Erwerb von VCC hierbei aber der Vokal als komplex, wie auch die neue Silbenposition als komplex gewertet werden, was zu einem Komplexitätsmaximum (++) führen würde. Durch den vorherigen Erwerb von CVCC wird initial Komplexität vermieden und genügend Ressourcen für die neue Silbenposition sind gegeben. Auch wird mit diesem Erwerbsschema ersichtlich, warum auf Stufe 2 CVC-Silben vor anderen Silbentypen erworben werden muss.

Da die Kodaposition schon integriert wurde, kann auf Stufe 4 ein Komplexitätsmaximum vermieden werden. Die o.g. Daten von Weir (1962) und Ingram (1978) legen nahe, dass die Erwerbsreihenfolge CVC vor VC auch für das Englische gegeben ist.

Levelt & Van de Vijver (2004) erörtern die Frage der Erwerbsreihenfolge unter Rekurs auf andere Faktoren wie Phonemerwerb und Salienz. So soll z.B. eine CCV-Struktur nicht zuerst wegen fehlender Phoneme zur Clusterbildung gebildet werden und V weniger salient als CVC sein. Letzteres wird damit begründet, dass kein Prozess im Niederländischen obligatorisch onsetlose Silben verlangt, während ungespannte Vokale obligatorisch geschlossen sind. Zusammen mit dem überbordenden Katalog an Constraints, die hier für den Silbenaufbau angenommen werden, scheint die Erklärung des Positionsausbaus wenig ökonomisch. Auch typologisch ist es für eine Argumentation auf Demisilbenbasis interessant, dass es wohl keine Sprache gibt, die neben der CV-Silbe einzig eine VC- oder VCC-Silbe zeigen, obwohl V- und CCV-Silben als einzig zusätzliche Silbentypen vorkommen (in Tab. 11 Cayuvava, Arabela und Mazateco). Alle Sprachen mit VC- und VCC-Silben setzen somit das Vorkommen von CVC-Silben voraus, ein Phänomen, das man mit der hier ausgeführten Analyse auf Basis von Demisilben und positionaler Ressourcenbeanspruchung hinreichend erfassen kann.

Language example	Type(s)				
Ilua	CV				
Thargari	CV	CVC			
Cayuvava	CV	V			
Arabela	CV	CCV			
Sedang	CV	CVC	CCV(C)		
Mazateco	CV	V	CCV		
Klamath	CV	CVC	CVCC		
Mokilese	CV	CVC	V(C)		
Totonac	CV	CVC	CVCC	CCV(C)(C)	
Finish	CV	CVC	V(C)	(C)VCC	
Spanish	CV	CVC	V(C)	CCV(C)	
Dutch	CV	CVC	V(C)	CCV(C)	(C)(C)VCC

Tabelle 11: Silbentypen in verschiedenen Sprachen (nach Levelt & Van de Vijver 2004).

Zusammenfassend kann man nach der hier vorgestellten Theorie die beobachteten Muster durch drei Leitprinzipien des Silbenaufbaus in der Sprachproduktion erklären:

1. Besteht eine Demisilbe nur aus einem Vokal, wird in die Verarbeitung immer die Komplexität derselben einbezogen.
2. Besteht eine Demisilbe aus einer Verbindung aus Konsonant und Vokal, wird in die Verarbeitung immer einbezogen, ob der Demisilbentyp neu produziert wird oder bereits vorhanden ist.
3. Optimierte nach Maßgabe der zielsprachlichen Maximalstruktur den Ausbau von Positionen bei Vermeidung maximaler Komplexität.

In einer strikt formalen Darstellung stellt sich der Erwerbsmechanismus wie folgt dar³⁷ :

$$\begin{array}{l} \# (-) \quad (\dots) \quad V \quad (\dots) \quad (+/-) \# \rightarrow \quad \# Q \quad (\dots) \quad V \quad (\dots) \quad (+/-) \# \\ \# (+/-) (\dots) \quad V \quad (\dots) \quad (-) \# \rightarrow \quad \# (+/-) (\dots) \quad V \quad (\dots) \quad Q \# \end{array}$$

Wir haben so unter Zugrundelegung des Prinzips der minimalen Struktur und der Demisilbentheorie eine Modellierung des Silbenpositionsausbaus gewonnen, die in ihrer strikten Distinktion von Performanz (in der Produktion) und Kompetenz (im Sprachverstehen) den chronologisch beobachtbaren Spracherwerbsdaten gerecht wird, zudem aber auch die notwendigen und optionalen Schritte des sukzessiven Silbenpositionsausbaus systematisch erfasst.

Nach der Besprechung der metrischen und silbenpositionalen Ebene soll nun im folgenden Kapitel der segmentale Erwerb unter Bezug auf Sonorität im Vordergrund stehen.

2.2.3 Sonorität im Erwerb

Das überzeugendste Beispiel für die Wirksamkeit von Sonoritätsbeschränkungen im Spracherwerb sind Clusterreduktionen (für eine Übersicht s. Yavaş 2003). Hier werden zwei C-Positionen in Onset oder Coda zu einer Position reduziert, so dass auch nur noch ein Phonem in der verbleibenden Silbenposition eines entsprechenden Zielwortes realisiert wird. Ohala (1999) hat Clusterreduktionen von 16 Kindern im Alter von 1;9 bis 3;2 in einem

³⁷ # = Begrenzung durch minimale und maximale Silbenschablone;
 = wenn X, dann Y;
 V = Vokal;
 Q = neue Position oder nur Übernahme einer neuen Struktur;
 (-) = nicht komplex bzw. letzte Silbenposition schon erworben;
 (+) = komplex bzw. letzte Silbenposition noch nicht erworben;
 (...) = Zwischen Vokal und letzter Position schon erworbene Positionen

Nachsprechtest von Nonsense-Wörtern untersucht. Hier fand sie, dass bei initialen Clusterreduktionen der weniger sonorante Laut (z.B. Frikativ + Nasal [snuf] zu Frikativ [suf]), final aber der sonorantere Laut erhalten blieb (Frikativ + Plosiv [skub] zu Plosiv [kub]). In den Produktionen der Kinder konnte also die Tendenz beobachtet werden, dass Konsonantencluster in einer Systematik reduziert wurden, die sich mit dem Sonority Dispersion Principle beschreiben lässt.³⁸ D.h. initial werden Phoneme mit geringer Sonorität präferiert, während final Phoneme mit hoher Sonorität bevorzugt werden. Es ist allerdings fraglich, inwiefern diese Tendenz auf den Erwerb aller Zielsprachen abbildbar ist. So fanden z.B. Lleó & Prinz (1996) Unterschiede zwischen den Clusterreduktionen von Spanisch und Deutsch lernenden Kindern im Alter von 0;9-2;1 Jahren.

Besonders interessant ist im Zusammenhang mit Clusterreduktionen eine Studie von Ott et al. (2006). Hier wurde mit zwölf Deutsch sprechenden Kindern im Alter von ca. 5;1 Jahren ein Nachsprechtest durchgeführt. Alle Kinder hatten eine phonologische Störung und zeigten in ihrer Produktion systematisch entweder Vorverlagerungen oder/und Plosivierungen. Die Items waren derart konzipiert, dass die genannten phonologischen Prozesse innerhalb von Onset-Clustern zu einer Verletzung der Phonotaktik führen würden (z.B. Kleid → Tleid). Das erstaunliche Ergebnis dieser Studie war, dass die Phoneme innerhalb der Cluster signifikant häufiger korrekt produziert wurden (ansonsten wurde das Cluster reduziert) als Phoneme in Wörtern ohne initiale Cluster. Selbst bei Kindern mit gestörtem Spracherwerb sind also Sonoritätsbeschränkungen weiterhin grundlegend und bilden den Rahmen für die Phonemproduktion.

Unabhängig von Clusterreduktionen scheint es im Spracherwerb allerdings keine einfache Abbildung des Sonority Sequencing Principles und Sonority Dispersion Principles auf Produktionen zu geben. Fikkert (1994) nimmt für die Kodaposition im Niederländischen eine Erwerbsreihenfolge an, in der Obstruenten vor Sonoranten erworben werden, was dem Sonority Dispersion Principle klar widerspricht. Auch Stites, Demuth und Kirk (2004) fanden in ihrer Untersuchung der Spontansprachdaten von zwei englischsprechenden Kindern zwei unterschiedliche Orientierungen vor. Kind N (1;1.10—1;4.25) produzierte schon im Alter von 1;2.23 80% der zieladäquaten Kodaplosive.³⁹ Diese Produktionsrate erreichte es für Frikative erst im Alter von 1;3.12, aber für Kodanasale selbst zwei Monate später noch nicht. Kind W hingegen konnte im Alter von 1;7.17 Nasale und Frikative mit 80 % Akkuratheit produzieren,

³⁸ Ohala (ebd.) bezieht sich auf Clements Sonoritätskonzeption, unterscheidet aber dennoch Plosive und Frikative hinsichtlich ihrer Sonorität. Diese Unterscheidung findet man in vielen Studien vor. Man beachte, dass Clements nur die Klasse der Obstruenten annimmt.

³⁹ Die Lautklasse der Plosive zeigt im Input der Kinder die höchste Frequenz in der Kodaposition.

während Plosive weiterhin schwierig blieben. Die Autoren interpretierten die Daten dahingehend, dass Kind N sich an der Frequenz der Laute im Englischen, Kind W sich an Markiertheit orientierte.

Wie sind solche differenten Ergebnisse zu erklären? M.E. liegt das Problem an den unterschiedlichen Methoden, ein Aspekt, der nicht berücksichtigt wurde (so z.B. auch nicht in Yavaş 2003). Während Ohala (1999) einen Nachsprechtest mit Nicht-Wörtern durchführte, untersuchten die anderen o.g. Studien zum ungestörten Erwerb die Spontansprachproduktionen von Kindern. Die Vermutung liegt nahe, dass es spezifische Effekte von Sonorität gibt, je nachdem, ob die Produktion über die lexikalische oder nicht-lexikalische Route vonstattengeht. Auf diesen Unterschied wurde schon in Kapitel 1.5.2 in Bezug auf die Produktionen von Jargon-Aphasikern hingewiesen. Im Zweitspracherwerb finden wir zusätzlich ein Phänomen vor, wo zwar keine Blockierung des Lexikons gegeben ist, allerdings erst einmal ein zusätzliches Lexikon neu aufgebaut werden muss. Und eben hier findet man interessanterweise Evidenz für einen spezifischen Einfluss des Sonority Dispersion Principle. Eckman & Iverson (1994) analysierten Fehler in der Kodaposition von erwachsenen nativen Sprechern des Japanischen, Kantonesischen und Koreanischen in der Produktion von Englisch. Die Fehlerrate bei der Produktion von Obstruenten in der Kodaposition war mit Abstand die höchste, obwohl Obstruenten im Kantonesischen und Koreanischen ebenfalls in der Kodaposition vorkommen. Dass Obstruenten nach dem Sonority Dispersion Principle in der Kodaposition die höchste Komplexität besitzen, scheint ein entsprechender Einfluss von Markiertheit bei dem Aufbau eines neuen Ziellexikons augenscheinlich. Baptista and da Silva Filho (1997) stellten analoge Schwierigkeiten bei Sprechern von brasilianischem Portugiesisch mit der Zweitsprache Englisch fest. Auch bei Clusterreduktionen existiert offensichtlich ein ähnlicher Einfluss des Sonority Sequencing Principles wie Ohala ihn bei Kindern festgestellt hat. Tropf (1987) fand anhand der Produktionen von Sprechern des Spanischen mit Zweitsprache Deutsch, dass Plosiv-Frikativ-Cluster in der Kodaposition nach Maßgabe des Sonority Dispersion Principles am häufigsten umgestellt wurden. Bei Reduktionen in Kodaposition wurde entsprechend das Phonem mit geringerer Sonorität getilgt. Was bedeuten diese Ergebnisse für den Erstspracherwerb?

Das Sonority Dispersion Principle ist wahrscheinlich vornehmlich auf der nicht-lexikalischen Route als Beschränkung wirksam. Dies gilt sowohl für Produktionen von Jargon-Aphasikern als auch (in einem anderen Grade) für diejenigen im Fremdspracherwerb und der Produktion von Nicht-Wörtern im Erstspracherwerb. In der Produktion haben diese Beschränkungen

insofern einen Einfluss, als dass solange keine Spezifikation einer zielsprachlichen Repräsentation gegeben oder diese nicht mehr zugänglich ist, eine phonologische Struktur für die Produktion gewählt wird, die den geringsten kognitiven Aufwand erfordert, vorzugsweise also Demisilben mit geringer Komplexität (vgl. Kap. 1.4.2). Bei Clustertilgungen wird so schlicht die grundsätzliche Tendenz zur optimalen Silbe deutlich. Die Planung und Berechnung in der Produktion folgt hier wiederum dem Prinzip der minimalen Struktur (vgl. Kap.2.2.1). Der Vorteil liegt auf der Hand, wie Dittmann (2002: 57) im Zusammenhang mit diesem Prinzip betont:

„Damit verringert sich die kognitive Komplexität der Sprachproduktionsaufgabe solange, bis die entsprechenden grammatischen Prozesse stärker automatisiert sind.“

Da Protowörter kein zielsprachliches Pendant besitzen, insofern definitiv nicht über die lexikalische Route produziert werden, liegt die Vermutung nahe, dass auch sie in ihrem Aufbau dem Prinzip der minimalen Struktur folgen und Demisilben mit geringer phonologischer Komplexität in der Produktion präferiert werden.

Bis hierher konnte für den Erwerb durch die Anwendung der 3D-Phonologie und der klassischen Performanz-Kompetenz-Distinktion auf die prosodische Hierarchie verzichtet werden. Ein wichtiges Phänomen im frühen Spracherwerb, das häufig mit letzterer erklärt wird, ist die Produktion von Reduplikationen. Im folgenden Kapitel soll deshalb untersucht werden, inwiefern sich auch dieses Phänomen neu fassen lässt.

2.2.4 Reduplikation im Erwerb

In der wissenschaftlichen Untersuchung von Reduplikationen im Spracherwerb ist man bei gegenwärtiger Forschungslage mit gewichtigen Problemen konfrontiert. Zum einen besteht die Problematik, dass gerade aus Sprachen, in denen Reduplikationen zur regelhaften morphologischen Markierung, wie z.B. nominalem oder verbalem Plural, Tempus oder Aspekt gebraucht werden, kaum Erwerbsdaten vorliegen (vgl. Dressler et al. 2005). Dementsprechend existiert keine umfassende Theorie von Reduplikation im Spracherwerb, welche die Interaktion von Phonologie und Morphologie adäquat beschreibt und in ein Modell integriert. Die Folge ist, dass Reduplikationen separiert jeweils unter phonologischer oder morphologischer Perspektive betrachtet werden. So auch in der Natürlichen Morphologie, in der Reduplikation als rein morphologischer Prozess verstanden wird.

Unterschieden werden grammatische bzw. modularisierte und extragrammatische, nicht-modularisierte Reduplikationen. Letztere sind vor allem onomatopoetische Lautimitationen, so dass extragrammatische Reduplikation nach dieser Theorie in einem größeren Grade ikonisch ist als grammatische Reduplikation. Dies spiegelt sich auch darin wider, dass hier häufiger totale als partielle Reduplikation vorkommt, d.h., es wird häufiger ein Wortstamm als kleinere Phonemketten verdoppelt. Auch sogenannte Echo-Wörter fallen unter extragrammatische Reduplikation. Wörter wie Ticktack oder Zickzack sind hier zu nennen. Während in Ticktack das Verb tick-en verdoppelt wird, ist in Zickzack das Nomen Zack-e das verdoppelte Element. Da beide syntagmatische Richtungen bei der Verdopplung möglich sind, wird angenommen, dass keine Regel bei Echo-Wörtern wirksam ist (vgl. Dressler 2005).

Im Spracherwerb unterscheidet dieser Ansatz drei Phasen:

1. Prämorphologie
2. Protomorphologie
3. Modularisierte Morphologie

In die Phase der Prämorphologie fallen alle phonologische Typen der Reduplikation. Hierunter werden sowohl reimende Sprachspiele (Beispiel aus dem Französischen: Emma 1;9 photo ptiptoptip = petite photo (ebd.: 459)) wie auch Fälle von Konsonantenharmonie (Beispiel aus dem Deutschen: Jan 1;3 bam = Banane (ebd.) subsumiert.

In der protomorphologischen Phase werden Reduplikationen zum Teil für morphologische Markierungen verwendet. So ersetzen zwei Mädchen aus Litauen Verbpräfixe mit reduplizierter CV- oder V-Silbe, hier das Reflexiv si (vgl. Wójcik & Smoczyńska 1997):

iš-si-maude (ein Bad genommen) → amamaude

at-si-kels (wird aufstehen) → akakels

Eine zieladäquate morphologische Markierung ist schließlich in der Phase der modularisierten Morphologie gegeben. Frühe Reduplikationen werden also nach diesem Ansatz als extragrammatische Operationen verstanden, mit denen Kinder in einem gewissen Sinne Vorübungen für spätere morphologische Prozesse vornehmen (vgl. Dressler 2005: 469).

Aus rein phonologischer Perspektive gibt es einen Dissens darüber, ob Reduplikation als universelles Phänomen (Moskowitz 1973) oder als individuelle Strategie (eventuell von langsamen Sprachlernern) (Ferguson 1983) angesehen werden muss. In einigen frühen Studien (Fee & Ingram 1982; Ingram 1974; Schwartz et al. 1980) wurde darauf hingewiesen,

dass Reduplikation eventuell als Vermeidung komplexer Formen angewendet wird. So interpretieren Schwartz et al. (1980) das Ergebnis aus einer Studie mit zwölf Englisch sprechenden Kindern im Alter von 1;3 bis 1;10 Jahren, in der sechs Kinder mit größerer Reduplikationsrate vor allem bei intendierten mehrsilbigen Zielsprachwörtern Reduplikationen produzierten, als einen Hinweis darauf, dass der Erhalt der mehrsilbigen Struktur eine entscheidende Beschränkung darstellt. Einen weiteren Faktor stellte das Vorkommen von Kodakonsonanten in der Zielform dar (vgl. Ingram 1974). Auf dieser Studie aufbauend untersuchten Fee & Ingram (1982) Daten von 24 Kindern im Alter von 1;1 bis 2;8 Jahren (ebenfalls mit Englisch als Erstsprache) und konnten die Ergebnisse von Schwartz et al. (1980) im Wesentlichen bestätigen. Allerdings fanden sie, dass die Kinder mit hoher Reduplikationsrate auch jüngere Kinder waren, die noch keine Kodakonsonanten produzierten. Insofern scheint Reduplikation dem Erwerb von Kodakonsonanten voranzugehen. Von Fee und Demuth (1995) wurden diese und weitere Ergebnisse schließlich im Rahmen einer Theorie prosodischer Lizenzierung interpretiert. Kinder sind nach ihrem Ansatz anfänglich nur in der Lage eine minimales Wort, d.h. einen Fuß mit zwei Moren zu bilden (s. Kap. 2.2.1). Dies erreichen sie durch Reduplikation einer offenen Silbe oder durch die Produktion einer geschlossenen. Können von Kindern noch keine Kodakonsonanten produziert werden, sind dementsprechend auch mehr Reduplikationen in ihren Produktionen vorhanden. Allerdings sprechen selbst Daten von Befürwortern einer Phase des minimalen Wortes nicht eindeutig für eine solche. Salidis & Johnson (1997) analysierten die Produktionen des Kindes Kyle im Alter 0;11-1;7 Jahren. Am häufigsten wurden minimale Wörter produziert, allerdings bestanden fast alle aus nur einer Silbe (ebd.: 16). Die häufigsten zweisilbigen Produktionen bildeten mit zunehmenden Alter supraminimale Wörter (Formen mit mindestens einem Kodakonsonanten) (ebd.: 17). Die Annahme, dass hier einem zweimoraischem Prinzip gefolgt wird, wirkt angesichts der Daten doch recht konstruiert. Dass ein Sprachlerner des Englischen, in dem CVC-Silben mit einer Häufigkeit von 30% seltener auftreten als CV-Silben mit 34% (vgl. Dauer 1983) (und diese zudem universal als unmarkierte Form gelten), kaum minimale Wörter der Form CVCV produziert, steht klar im Widerspruch zur Theorie des minimalen Wortes. In der Studie von Salidis & Johnson wurden Reduplikationen nicht explizit analysiert, aber der Verweis darauf, dass zweisilbige minimale Wörter bei Kindern, welche die Bantu-Sprachen Sesotho und Swahili sprechen, häufiger vorkommen, lässt m.E. mehr auf einen Einfluss häufiger Reduplikationen in diesen Sprachen

als auf Quantitätssensitivität schließen. Reduplikation ist in Sesotho wie auch in Swahili ein produktiver morphologischer Prozess. (vgl. z.B. McNally 1990; Mohammed 2001).

Auch Daten von Lleó (1990) sprechen gegen die Annahme eines engen Zeitfensters für Reduplikationen basierend auf prosodischer Lizenzierung. Lleó sieht Reduplikation mit Homonymie zusammen als eine Form kognitiver Strategie an. Strategien werden hier verstanden als „ability to create phonological rules, are universally available and the child can apply them when she needs to.“ (ebd.: 274). Reduplikation hat demnach die Funktion, ein innerlexikalisches Muster zu erzeugen, um die Struktur und die artikulatorische Anstrengung zu minimieren. Die Strategie der Homonymie hingegen führt dazu, dass ein einziges Muster mit identischen Konsonanten und identischen Vokalen für verschiedene Wörter, d.h. interlexikal vorhanden ist. Beide Strategien vereinfachen nach Lleó die Speicherung in einem schnell wachsenden Lexikon. Dies sieht sie darin bestätigt, dass Reduplikationen nach dem Alter von 1;7 bis 2;5 kontinuierlich abnehmen, bis so gut wie keine mehr produziert werden. Dies gilt aber nur für die Produktion zweisilbiger Wörter. In der Produktion dreisilbiger Wörter steigt die Reduplikationsrate ab ca. 2;7 graduell an, bis Reduplikationen im Alter von 2;11 ungefähr 20% aller produzierten dreisilbigen Wörter (und nur dieser) ausmachen. Da in diesem Alter auch Kodakonsonanten schon erworben sind, kann man diese späteren Reduplikationen natürlich nicht auf das Fehlen derselben zurückführen. Eine strikte Interpretation der prosodischen Lizenzierung sagt natürlich auch voraus, dass es keine totale CVC- oder CV-CVC-Reduplikationen im frühen Erwerbssalter geben dürfte. Diese sind allerdings attestiert, wie z.B. in Dressler (1995: 463):

Jan (Deutsch), 1;3 Jahre: baubau und babau = Bauch
Vanja (Russisch), 1;2 Jahre: kapkap = kápat (tropfen)
Emma (Französisch), 1;5 Jahre: lalam = lampe (Lampe)

Zusammenfassend kann man sagen, dass sowohl die Natürliche Morphologie als auch rein phonologische Ansätze Reduplikationen strikt in phonologische und morphologische unterteilen und damit unterschiedliche Mechanismen für deren Zustandekommen postulieren müssen. Ansätze, die auf prosodischer Lizenzierung aufbauen, können nicht erfassen, wieso auch schwere Silben redupliziert werden und wieso Reduplikationen auch in einem späteren Erwerbssalter bei dreisilbigen Wörtern noch einmal zunehmen. Im folgenden Kapitel soll gezeigt werden, wie ein Modell, das auf der Precedence Based Phonology aufbaut, diese Lücken schließen kann.

2.2.4.1 Precedence-Based Phonology im Spracherwerb

Raimy (2009) wie auch Nevins (2004), deren Ausarbeitungen zum Spracherwerb hier vorgestellt werden sollen, beziehen sich explizit auf das *Variational Model* von Yang (2002, 2004). In diesem Modell wird angenommen, dass Parameter einen Hypothesenraum definieren, in dem ein quasi-darwinistischer Selektionsprozess durch Frequenz zur Auswahl der richtigen Hypothese führt. Nevins (2004) verwendet dieses Modell zur Erklärung dafür, wie Ambiguitäten im Spracherwerb aufgelöst werden können. Ambiguitäten werden von ihm in einem weiten Sinne verstanden. Die syntaktische Konstruktion „I saw the man with the telescope“ kann auf zwei verschieden Arten geparst werden. Zum Einen als VP NP PP, zum Anderen als VP NP, in der die NP nochmal NP PP dominiert. Wichtig ist, dass Nevins Ambiguität nicht notwendig mit einer doppelten Bedeutungszuweisung assoziiert, sondern eine Anschauung vertritt „in which an observed phenomenon can correspond to two or more distinct mental structures, regardless of their particular representational vocabulary.“ (ebd.: 116) Wenn nun ein Sprecher des Kaigan die Reduplikation *va-va* erstmals vernimmt, steht er vor einer Vielzahl universalgrammatischer Optionen, was den Typus der Reduplikation anbelangt:

- (1) total reduplication: Chinese *tiantian*
- (2) first-syllable reduplication: Ilokano *kal-kaldin, tak-takder*
- (3) last-syllable reduplication: Kaingang: *va-va, jemi-mi, vasan-san*
- (4) iterative reduplication: Chinese *gaogaoxingxing*
- (5) stressed-syllable reduplication: Chamorro *hu-ga-gándo*

Die Reduplikation *va-va* im Kaigan ist also insofern ambig, als dem Lerner fünf mögliche strukturelle Repräsentationen zur Verfügung stehen. Wenn Reduplikation die Wiederholung einer Sequenz von Phonemen innerhalb eines Wortes ist, muss der Lerner die Start- und Endpunkte, spezifischer o.g. Ankerpunkte, der kopierten Sequenz ausmachen und generalisieren. Die Ankerpunkte sind nicht zufällig, sondern lassen sich in Form von Parameterwerten wiedergeben:

Parameters of Anchor Points

The segment **defined-by**, **immediately-preceding**, or **immediately-following**: {0,1,2}

The **Segment**, **Vowel**, or **consonant** {0,1,2}

That is the **first**, **last**, **stressed**, or **every** in the word {0,1,2,3}

Für die totale Reduplikation im Chinesischen *tiantian* bedeutet dies also:

Ankerpunkt des Starts: The segment **defined-by** {0}; The **Segment** {0}; That is the **first** in the word {0}

Ankerpunkts des Endes: The segment **defined-by** {0};The **Segment** {0}, That is the **last** in the word {1}

In dieser Form lassen sich mit den Parameterwerten für die Start- und Endpunkte nun alle Reduplikationstypen definieren:

Total reduplication:	<0,0,0> to <0,0,1>
Chamorro stressed CV:	<1,1,2> to <0,1,2>
Lushootseed first CVC:	<1,1,0> to <2,1,0>
Lushootseed first VC:	<0,1,0> to <2,1,0>
Kaingang last CV:	<1,1,1> to <0,1,1>
Iterative reduplication:	<1,1,3> to <2,1,3>

Für den Lerner ist mit Hilfe dieser universalgrammatischen Parameterwerte der Suchraum für den adäquaten Reduplikationstypus auf 24 Wertoptionen pro Ankerpunkt beschränkt. Der Lerner des Kaingan steht nun also vor folgenden Optionen:

<0,0,0> to <0,0,1>
<1,1,2> to <0,1,2>
<1,1,0> to <0,1,0>
<1,1,1> to <0,1,1>
<1,1,3> to <0,1,3>

Diese Optionen beschränken aber nur den Suchraum, nichts in der UG weist den Lerner des Kaingan darauf hin, welche Option die zielsprachliche ist. Vielmehr ist die Annahme, dass eine parallele Suche stattfindet, indem multiple Grammatiken simultan unterhalten werden. Variation entsteht so dadurch, dass simultan multiple Parameter belegt werden. Die Entscheidung für eine Parameterbelegung wird letztlich, wie im Variational Modell von Yang, durch Frequenz im Input geregelt. Nevins (ebd.: 125) fasst seine Grundannahmen bezüglich des Erwerbs von Reduplikationen wie folgt zusammen:

„At least in reduplication, there need not be any special learning path, or any principled means of ignoring certain inputs until other parameters have stabilized. [...] UG provides no bias towards „marked“ or „unmarked“ types of reduplication, and contains no provisions for identifying and ignoring ambiguous input. But it excludes so many irrelevant hypothesis to allow convergence often within six to ten steps. It seems that the secret to successful

acquisition is not in a clever learning-algorithm, but rather a careful delimitation of the parameters of variation.“

Raimy (2009) nimmt hingegen an, dass in der UG Markiertheit durchaus eine Rolle spielt. So nimmt er im Sinne der Elsewhere Condition (Kirparsky 1973) die universalgrammatisch implementierte Tendenz an, dass Sprachlerner von allgemeineren Hypothesen zu spezifischeren voranschreiten (in a: *be conservative*). Zudem wird von ihm in Anlehnung an das Obligatory Contour Principle angenommen, dass ähnliche Segmentsequenzen vermieden werden sollen (in b: *be different*).

a) Be conservative: favor anchor points with more general parameter settings.

Placement: at > before/after

Plane: x-tier > metrical/syllable/consonantal

Target: plain > stressed(head)/consonant

b) Be different: if the anchor points, X and Y, that define a precedence link are not identical (X=Y), the favor pairs of anchor points that are distinct with respect to (a).

Mithilfe dieser Präferenzen lässt sich eine Markiertheithierarchie konstruieren, in der totale Reduplikationen den unmarkierten Fall darstellen, partielle Reduplikationen allerdings auch relativ unmarkiert sind. Da im Erwerb nur reduzierte Silbenstrukturen produziert werden, ist nicht zu entscheiden, welche Hypothesen im Ganzen aktiv sind.

Unmarkierte Hypothese
Letztes Segment → erstes Segment
Nächste Markiertheitsstufe
Letztes Segment → erster Konsonant
Letzter Vokal → erstes Segment
Nächste Markiertheitsstufe
Letzter Vokal → erster Konsonant
Letztes Segment → vor ersten Vokal
Nach letzten Konsonanten → erstes Segment
Markiertesten Fälle
Letzter Vokal → vor ersten Vokal
nach letztem Konsonanten → erster Konsonant
Nach letztem Konsonanten → vor ersten Vokal

Tabelle 12: Hypothesen im Erwerb (nach Raimy 2009).

Der entscheidende Punkt ist hier nun, dass eine auf der UG basierende Theorie wie die Precedence-Based Phonology eine spezifische Vorhersage macht: Auch in Sprachen, in denen der Prozess der Reduplikation keine grammatische Funktion einnimmt, müssen unmarkierte Hypothesen im frühen Erwerbsstadium notwendig erst einmal unterhalten werden. Was oberflächlich also als eine prosodische Beschränkung erscheint, ist vielmehr als graduelle Reduzierung der Hypothesen über morphologische Markierung mittels Reduplikation zu verstehen.

Aus einer Reanalyse der Daten von Hilde (Stern & Stern 1928) kann man erkennen, dass im frühen Erwerbsstadium (Alter 1;2 bis 1;6) nicht die Frage danach relevant ist, ob die minimale Anzahl von Moren produziert werden, sondern ob es sich um totale oder partielle Reduplikation handelt. Wir finden hier auf Typen bezogen sowohl CVC-Reduplikationen wie auch einen Anteil an partiellen Reduplikationen der Form CVCVC.

Silbenstruktur	Reduplikationstyp und Häufigkeit	Häufigkeit aller produzierter Reduplikationstypen (in %; n= 23)	Anteil der Reduplikationstypen im Gesamtwortschatz (in %; n = 44 (Typen))
CV	Total (n = 8)	36	18
CVC	Total (n = 7)	30	16
CVCV	Total (n = 4)	17	9
CVCVC	Partiell (n = 3)	13	7
C	Total (?) (n = 1)	4	2

Tabelle 13: Typenanalyse von Reduplikationen in Stern & Stern (1928).

Auch die Daten von Lleó passen hier mit Daten aus dem morphologischen Erwerb zusammen. Sobald sich ab ca. 1;6 Jahren die produktive Markierung durch Suffigierung etabliert, nehmen die Reduplikationen von zweisilbigen Formen kontinuierlich ab. Kann man nun aber mithilfe der Precedence-Based Phonology auch die zweite Phase der Reduplikation erklären? Dies ist m.E. wiederum im Rekurs auf den morphologischen Erwerb möglich. Ist der morphologische Prozess der Suffigierung schon erworben, so beginnen Kinder erst ab ca. 2;0 erste Partizip Perfekt-Formen zu verwenden. Dies geschieht zunächst ohne die Markierung durch das Affix *ge-* und wird erst nach und nach morphologisch etabliert. D.h., Reduplikationen fungieren auch in diesem späteren Erwerbsalter wieder als optionale Markierung morphologischer Information. Dass in den von Lleó angegebenen Beispielen so gut wie ausschließlich die erste

Silbe redupliziert wird, zeigt hier gewissermaßen die Domäne bzw. die Richtung morphologischer Markierung an.

Eine auf der Precedence-Based Phonology fußende Interpretation der Erwerbsdaten integriert auch die Beobachtung von Fee & Ingram (1982) und Demuth & Fee (1995), dass Reduplikationen mit dem Erwerb von Kodakonsonanten abnehmen. Allerdings treten CV-Reduplikationen nach der hier vertretenen Theorie nur deshalb häufiger auf, weil die Silbenstruktur noch nicht konstant zu CVC erweitert wurde. Das minimale Wort kann so als ein Artefakt der Interaktion von Silbenstruktur und unmarkierten Reduplikationsprozessen angesehen werden. Da Wortstämme im Deutschen minimal eine CVC-Struktur haben müssen, besteht mit der stabilen Etablierung dieses Silbenformats um den Zeitraum von ungefähr 1;6 Jahren (Ingram 1978) nun erst die Applikationsmöglichkeit der Suffigierungsregel (nach Raimy 2000: 67; vgl. Kap. 1.7.1):

Anfang: { __ → % }

Ende: → %

Dass diese sehr frühe Koinzidenz zwischen phonologischer und morphologischer Entwicklung den Vertretern der Theorie des minimalen Wortes wie auch in den früheren Untersuchungen von Fee & Ingram (1982) und Schwartz et al. (1980) entgehen musste, liegt wohl nicht zuletzt daran, dass die Erstsprache der beobachteten Kinder das Englische war. Hier fehlen im frühen Spracherwerb in der Produktion morphologisch overte Markierungen aufgrund der phonologischen Entwicklung. Die im Alter von 19-33 Monaten erstmals auftretenden Morpheme *-ing* und *-s* (vgl. Brown 1973) sind aus silbenstrukturellen wie auch aus Gründen des segmentalen Erwerbs erst spät in der Produktion realisierbar (vgl. z.B. Song et al. 2009). So tritt im Erwerb des Englischen der Nasal /ŋ/ bereits im Alter von ca. 2;0 Jahren auf, während der Frikativ /s/ erst ab ca. 3;0 Jahren regelmäßig produziert wird (vgl. Stoel-Gammon & Dunn 1985). Finale Konsonanten werden großteils noch bis zum Alter von ca. 2;6 Jahren ausgelassen (vgl. Grunwell 1987). Da das Morphem *-ing* einen Vokal enthält, wird es in der Resilbifizierung mit einer eigenen Silbe assoziiert. Das Morphem *-s* hingegen benötigt eine weitere Silbenposition als Landeplatz. Es ist somit in der Produktion zu erwarten, dass es als morphologische Markierung spät zu beobachten ist. Ganz anders verhält es sich allerdings in der Perzeption. Im Rahmen einer Untersuchung mittels der Headturn Preference Procedure konnten z.B. Soderstrom et al. (2007) zeigen, dass schon sehr junge Kinder sensitiv für morphologische Kontraste waren (vgl. auch Soderstrom et al. 2002).

Bereits im Alter von 1;4 Jahren wurden von den Kindern dieser Studie Sätze bevorzugt, welche die nominale oder verbale Flexion *-s* zielsprachlich adäquat markierten (*Chairs are good to stand on when you sing a lot* vs. *Chair are good to stand on when you sings a lot*; vgl. Soderstrom 2008: 672). D.h., die morphologische Komponente muss bereits noch früher aktiv und die jeweiligen distributionellen Muster der Zielsprache analysiert sein. Für das Deutsche haben Höhle & Weissenborn (1998, 1999) gezeigt, dass Kinder bereits im Alter von 7-9 Monaten funktionale Elemente (Artikel, Pronomina und Präpositionen) in einem Text wahrnehmen können und mit 10 Monaten Artikel als distinkte Einheiten analysieren (Weissenborn 2000: 148). Wie Weissenborn betont, bedeutet dies, „dass spätestens in diesem Alter Kinder damit beginnen können, die Verteilungsmuster von funktionalen Elementen zur Ableitung grammatischer Regelmäßigkeiten zu benutzen.“ (ebd.: 148) Auch in Anbetracht der mannigfaltigen Erkenntnisse aus Perzeptionsstudien mit sehr jungen Kindern erscheinen die natürliche Morphologie wie auch die Theorie des minimalen Wortes überholt.

Letztlich findet man im Deutschen also entwicklungsmäßig eine zeitlich nähere Passung von Produktion und Perzeption durch die frühe (phonologische) Option overter Realisierung morphologischer Information. Die Endung *-t* tritt als erstes Suffix auf, wobei der Plosiv /t/ in der Kodaposition früh erworben ist und nur eine Silbenposition einnimmt. Hier offenbart sich also in Hinblick auf Reduplikationen im Gegensatz zum Englischen die modulare Interaktionen durch die in der Produktion beobachtbare Koinzidenz.

Solch modulare Interaktionen von Silbenstruktur und morphologischer Information sind auch in der Deutschen Gebärdensprache gegeben (vgl. Happ & Vorköper 2006). Hier wird z.B. der Durativaspekt bei Verben mit der Silbenstruktur HM und HMH (vgl. Kap. 1.4.1) durch Reduplikation des Verbstamms realisiert⁴⁰, während der Iterativ- und Habitativaspekt durch die Reduplikation des gesamten Verbs ausgeführt wird.

Ein im Hinblick auf obige Ausführungen weiterer interessanter Aspekt ist die Produktion von Reduplikationen in Pidgin- und Kreol-Sprachen. Bakker & Parkvall (2005) machen hier den Unterschied zwischen diesen beiden Sprachformen besonders stark. Mehrere Punkte sind hervorzuheben:

1. In Pidgin-Sprachen sind Reduplikationen äußerst selten, während sie in Kreolsprachen die Regel sind.

⁴⁰ Es gibt noch zwei andere Möglichkeiten, wie der Durativ in der Deutschen Gebärdensprache realisiert wird. Bei Verben mit der Silbenstruktur H oder MH wird das Verb einmal gebärdet und bleibt dann stehen, wohingegen bei Verben mit der Silbenstruktur M die Verbbewegung gedehnt wird.

2. Selbst wenn die Substratsprachen reich an Reduplikationen sind (vgl. Chinesisches Pidgin Englisch), fehlen sie in den Pidgin-Formen.
3. In Kreol-Sprachen sind totale Reduplikationen häufiger als partielle, wobei diese nicht einfach als ikonisch klassifiziert werden können.

Dass totale Reduplikationen in Kreolsprachen häufiger sind als partielle, stimmt mit den Annahmen von Raimy überein. Aus diesen Fakten kann weiter abgeleitet werden, dass der Prozess der Reduplikation als obligatorischer Parameter der UG angesehen werden muss. Denn wenn der Input keine Daten mit Reduplikationen liefert und die produzierten totalen Reduplikationen im Kreol nicht als ikonisch klassifiziert werden können, muss Reduplikation in der UG angelegt sein. Dieses lässt im weiteren Sinne keinen anderen Schluss zu, als dass man Reduplikation im frühen Spracherwerb ab ovo als eine Interaktion zwischen einer morphologischen und phonologischen Komponente verstehen muss.

In einer psycholinguistischen Abwandlung von einem Marxschen Ausspruch könnte man zum Abschluss dieses Kapitel durchaus sagen, dass der Erwerb in Anbetracht obiger Ausführungen wieder vom prosodischen Fuß auf seinen phonologischen Kopf gestellt werden muss. Auch Reduplikationen sind Bestandteil einer universalgrammatischen Grundausstattung, die vom Erwerbsbeginn an in komplexen Interaktionen modular organisierter Subkomponenten denselben steuert und beschränkt. Dass dies wirklich ab den frühesten Wortproduktionen der Fall ist, soll in den folgenden Kapiteln anhand der Untersuchung von Protowörtern weiter gestützt werden.

3 Protowörter

Protowörter werfen ähnlich wie Kreolsprachen die Frage auf, woher die systematische Verwendung sprachlicher Elemente und Regeln stammt, wenn kein Analogon für diese im Input existiert. Schon Ende des 19. Jahrhunderts wurde über diese Fragestellung ein Nativismusstreit geführt (vgl. Knobloch 2008). Die genaue Frage war, ob im frühen Spracherwerb Wörter vorkommen, die bezüglich ihrer Lautstruktur kein Vorbild in der Erwachsenensprache haben, in diesem Sinne also Erfindungen, Neubildungen oder Urschöpfungen des Kindes sind. Ein Instrumentarium phonologischer Analyse war zu dieser Zeit noch nicht gegeben, so dass die Diskussionen wesentlich auf die Herkunft und Funktion solcher Wörter beschränkt blieb.

Was die genuin phonologische Erforschung von Kindersprache betrifft, so schuf der Prager Strukturalismus in der Folgezeit Mitte des 20. Jahrhunderts die systematischen Grundlagen,

prägte allerdings ein Bild der frühen phonologischen Entwicklung, das auch in der nachfolgenden Generativen Grammatik bis heute persistiert. Eine Unterscheidung zwischen rein phonetischen und phonologischen Formen im Spracherwerb wie auch eine Einengung auf Phoneme und ihre Ermittlung durch ihre bedeutungsdifferenzierende Funktion innerhalb einer Ziel-, d.h. Erwachsenensprache, sind hier als Grundcharakteristika zu nennen.

Die folgende Darstellung der bisherigen Forschung zu Protowörtern folgt deshalb einer groben, aber durchaus hinreichenden Gliederung, in der eine prästrukturalistische, strukturalistische und nachstrukturalistische Phase der Protowortforschung unterschieden wird.

3.1 Phasen der Protowortforschung

3.1.1 Prästrukturalistische Phase

In den Anfängen der systematischen Spracherwerbsforschung, die mit den Anfängen der Entwicklungspsychologie zusammenfiel, nahm Preyer (1841-1897) die Haltung ein, dass das Kind keine eigenen Wörter erfinde oder präge, sondern alle produzierten Formen des Kindes auf den Verkehr mit der sprechenden Umgebung zurückgeführt werden könne (Preyer 1882: 295). Wundt schließt sich dieser Meinung an und baut die empiristischen Annahmen noch weiter aus. Erwachsene Sprecher erzeugten kindgemäße Ausdrücke, um die eigene Sprache dem kindlichen Laut- und Silbenrepertoire anzupassen (vgl. Wundt 1900: 311; Knobloch 2008: 1709). Alle übrigen Formen, so Preyer und Wundt, seien Nachahmungen von Erwachsenenwörtern, die bis zur Unkenntlichkeit verzerrt seien. Wichtig war im Rahmen dieser Diskussion die Arbeit von Carl Stumpf mit dem Titel „Eigenartige sprachliche Entwicklung eines Kindes“ (1901). Stumpf lehnt eine nativistische Erfindungstheorie ab, weil diese für ihn eine Mittel-Zweck-Relation impliziert. Das Sprachverhalten seines Sohnes, welches er in dieser Studie von 1;1 bis 3;3 beschreibt, ist für ihn vielmehr deshalb abweichend, weil es eine Mischung von misslungenen Nachahmungen, Onomatopoeika und spielerischen Umformungen des gegebenen Lautmaterials darstellt. Gerade dieser kreative Umgang mit gegebenem Lautmaterial wird aber von nativistisch orientierten Autoren als angeboren gedeutet. Ament (1902: 209) unterscheidet von der Lallphase ausgehend folgende Formen der Worterfindung:

- „1. Das Kind liefert von selbst den Laut oder das Wort, aber die Eltern geben den ohne besondere Absicht artikulierten Silben einen Sinn.
2. Das Kind erfindet das Wort und fixiert zugleich dessen Bedeutung. Dies ist der merkwürdigste, seltenste und zugleich am meisten umstrittene Fall.

3. In andern, sehr häufigen Fällen liefern die Eltern die Wörter; aber das Kind, das sie wiederholt, deutet sie in seiner Weise und benutzt sie in anderem Sinne.”
(Ament 1902: 209)

Die Tagebuchstudie von Stern & Stern (1928) bildete den vorläufigen Abschluss einer Diskussion, die heute in einem anderen Gewande fortgeführt wird. Zum Thema sogenannter Urschöpfungen kommen sie zu folgendem Urteil:

„Die völlig freie Schaffung von Wörtern und ihre Aufnahme in den Wortschatz kommt in den ersten Jahren der normalen kindlichen Sprachentwicklung so gut wie gar nicht vor [...].“

Aber auch wenn sie eine solch völlig freie Schaffung von Wörtern bestreiten, so stimmen sie doch mit Stumpf in dem Sinne überein, als dass die wahre Spontaneität der kindlichen Wortbildung „sich nicht im Schaffen aus dem Nichts, sondern im freien Schalten und Walten mit dem gegebenen Material“ (ebd.: 385) äußere. So ist denn bei ihnen das einzige positive Motiv zum Erfinden von Wörtern der Spieltrieb, der sich im Spiel mit Lauten äußere. Dass aber solch einer Lautspielerei eine dauerhafte Bedeutung zugewiesen werde, ist für Stern & Stern sehr unwahrscheinlich (ebd.: 391).

3.1.2 Strukturalistische Phase

Roman Jakobsons „Kindersprache, Aphasie und allgemeine Lautgesetze“ (1941) gilt allgemein als der Beginn phonologischer Untersuchungen zum Spracherwerb (vgl. auch Fikkert 2007). Auch wenn Jakobson seine grundsätzlichen Annahmen zur phonologischen Entwicklung nicht anhand eigener Studien, sondern durch die Verwertung bisheriger Forschungsarbeiten entwickelte⁴¹, sollten diese dennoch bis heute Gegenstand linguistischer Diskussionen sein (vgl. z.B. Drescher 2003; Menn & Vihman 2011). Dreh- und Angelpunkt von Jakobsons Spracherwerbstheorie ist die strukturalistische Phonemdefinition:

„Während jedem Wort bzw. jeder grammatischen Form eine eigene bestimmte Bedeutung innewohnt, übt das Phonem als solches einzig eine bedeutungsunterscheidende Funktion aus, ohne eine eigene positive Bedeutung zu besitzen: es unterscheidet jedes Wort, in dem es vorkommt, von allen Wörtern, welche ceteris paribus ein anderes Phonem enthalten.“
(Jakobson 1941: 41)

Die bedeutungsunterscheidende Funktion vermittelt einer paradigmatischen Opposition operiert im strukturalistischen Modell auf einer linearen Achse, d.h. auf Ketten von Phonemen. Die Frage ist also, in welchen lautlichen Umgebungen stehen spezifische Phone in Opposition zueinander und in welchen nicht. Hierfür werden Minimalpaarbildung, freie Variation und

⁴¹ Besonders die Untersuchung von Grégoires (1937) sind hier zu nennen.

komplementäre Distribution von Phonemen in der jeweiligen Sprache untersucht. Die strukturalistische Analyse ist insofern immer auf die jeweilige Zielsprache hin orientiert und die phonologische Opposition über Bedeutungsunterscheidung definiert:

„Unter phonologischem Gegensatz ist jeder lautliche Gegensatz zu verstehen, der in der betreffenden Sprache zur Differenzierung der intellektuellen Bedeutung verwendet wird.“
(Trubetzkoy 1935: 7)

Dieser Definition des Phonems gemäß entwickelte Jakobson sein Zwei-Phasen-Erwerbsmodell des kindlichen Lautsystems, mit einer ersten, rein phonetischen Phase, in der vorsprachliche Laute produziert werden, und einer eigentlichen Sprachstufe, in der es „um die Auslese der Sprachlaute und um den Aufbau des Phonemsystems geht“ (ebd. 1941: 30). Die beiden Phasen sind nach Jakobson durch eine stille Phase in ihrem zeitlichen Auftreten klar voneinander geschieden. Aufgrund dieser Annahme wird das Erwerbsmodell von Jakobson als ein diskontinuierliches klassifiziert. Unter Zugrundelegung dieser strukturalistischen Definition ist es verständlich, dass *alle* Produktionen des Kindes, in denen keine zielsprachlichen Bedeutungsunterscheidungen mittels realisierter Phone beobachtet werden kann, auch keinen phonematischen Status erhalten können und somit als vorsprachliche „wilde Laute“ oder „Lautgebärden“ gewertet werden (vgl. Jakobson 1941: 26). Es ist diese Orientierung an zielsprachlichen Phonemen, was man m.E. als das strukturalistische Residuum in der Spracherwerbsforschung bezeichnen kann und was sich, wie wir sehen werden, bis heute größtenteils als methodische Präsupposition in Spracherwerbsstudien verfestigt hat.

3.1.3 Nachstrukturalistische Phase

Es existiert in der modernen Spracherwerbsforschung ein unhinterfragtes Einverständnis darüber, ab wann Produktionen von Kindern als phonologisch gelten können. So findet sich z.B. in der jüngsten Veröffentlichung von Smith (2010) nach der Beschreibung des Zeitraumes vom Lallen zum Beginn der Produktion erster zielsprachlicher Wörter, der bezeichnende Satz: „His real phonology was beginning“ (ebd.: 53). In der Produktion findet man demnach echte oder wirkliche Phonologie erst dann, wenn eine Relation zu den Wörtern der Zielsprache hergestellt werden kann. Mit wenigen Ausnahmen werden dementsprechend frühe Sprachproduktionen in Studien zur phonologischen Entwicklung nach Maßgabe einer phonologischen Opposition analysiert, wie sie im Strukturalismus definiert wurde (s.o.). Dies bedeutet, dass man innerhalb des Forschungsparadigmas der Generativen Grammatik vom

diskontinuierlichen Modell Jakobsons ausgeht. Nicht-generative Ansätze betonen dagegen die Kontinuität von phonetischen und phonologischen Formen. Sie geben hierbei die grundsätzliche Unterscheidung von Phonetik und Phonologie auf und nehmen an, dass phonologische Formen durch Lernen direkt aus phonetisch-artikulatorischen Mustern (vgl. Vihman 1996; Piske 2001) emergieren.

Gemäß dieser Forschungstendenzen gibt es keine einzige Studie zu Protowörtern innerhalb der generativen Tradition, die diese Formen mithilfe eines phonologischen Instrumentariums analysiert. Hingegen gibt es mehrere Studien, die Protowörter als artikulatorische Routinen oder phonetische Muster klassifizieren, die quasi als Behelfsformen zum Bedeutungstransfer fungieren.

Der Terminus „Protowort“ geht aller Wahrscheinlichkeit auf Menn (1976) zurück und hat sich in der Forschungsliteratur gegen andere Termini durchgesetzt. Diese sind (vgl. Vihman 1996: 130):

- phonetically consistent forms (Dore et al. 1976)
- acts of meaning (Halliday 1975, 1979)
- vocable (Ferguson 1978)
- prewords (Fletcher & Garman 1979)
- sensorimotor morphemes (Carter 1979)
- call sounds (Werner & Kaplan 1984)
- quasi-words (Stoel-Gammon & Cooper 1984)

Alle oben genannten Termini bezeichnen das Phänomen des ersten konsistenten Wortgebrauchs, das gegen Ende des ersten Lebensjahres nach den Phasen des kanonischen und bunten Lallens auftritt. Während allerdings Menn (1976) auch Wörter mit zielsprachlichem Bezug als Protowörter klassifiziert, soll hier zuvorderst die Definition von Vihman (1996: 130) als Ausgangspunkt gelten⁴²:

„We will use the term [...] to refer to relatively stable child forms with relatively consistent use which lack any clear connection with the form + meaning unit of a conventional adult model.“

Wie oben genannt existieren nur wenige Studien, in denen die phonologische Struktur von Protowörtern untersucht wurde. Dore et al. (1976) z.B. erfassen sogenannte Phonetically Consistent Forms (PCFs) nach folgenden Kriterien (ebd.: 15-16):

1. Im Gegensatz zum Lallen sind PCFs isolierbare, durch Pausen getrennte, Einheiten.
2. PCFs werden wiederkehrend als Teil des Lautrepertoires produziert.

⁴² Diese Definition ist auch die in Studien zum Lexikonerwerb usuell zugrundeliegende (z.B. Kauschke 2000; Piske 2001).

3. PCFs können teilweise mit wiederkehrenden Bedingungen korreliert werden, d.h., sie werden nicht so zufällig wie Lallproduktionen, aber auch nicht so regelhaft wie Wörter der Zielsprache verwendet.
4. PCFs zeigen eine „protophonemische“ Struktur in dem Sinne, als dass ihre phonetischen Elemente stabiler als in Lallproduktionen, aber weniger stabil als in Wörtern der Zielsprache sind.

In ihrer Studie mit vier Kindern ab dem Alter von 0;11 bis 1;4 Jahren, in der acht Monate lang einmal im Monat einstündige Aufnahmen gemacht wurden, ermittelten sie die phonetische Ähnlichkeit von Wörtern nach der Identität der Silbenstruktur, der Ähnlichkeit der Konsonanten nach Artikulationsart und -ort und vokalische Ähnlichkeit nach Zungenstellung und -höhe. Solchermaßen phonetisch ähnliche Wörter mussten mindestens fünf Mal in einer Aufnahmesitzung auftreten. Auf Basis dieser Klassifikation identifizierten sie folgende vier Funktionstypen von PCFs (ebd.: 16-17):

1. Affektausdrücke (affect expressions): Hierbei handelt es sich um PCFs, die zum Ausdruck von Freude, Befriedigung, Ärger, Protest usw. verwendet werden, wobei sie nicht an eine Person gerichtet sein müssen. So äußerte ein Kind z.B. [æ::] in verschiedenen Spielkontexten als Ausdruck der Freude, [i::] hingegen verwendete es als Ausdruck des Protestes gegenüber der Mutter. Ein anderes Kind verwendete [gægi], [gagi] und [əgagi] als Ausdruck der Freude, [e::] hingegen als Ausdruck der Zufriedenheit und Befriedigung.
2. Instrumentelle Ausdrücke (instrumental expressions): Diese werden von Kindern gebraucht, wenn sie einen Gegenstand haben möchte und/oder einen Erwachsenen zu einer Handlung bewegen möchte. So verwendete z.B. ein Kind [m:::], wenn es versuchte an seine Flasche zu kommen, ein anderes [ʌ] und [ʌ:], wenn es in die Küche kam und sein Essen suchte.
3. Aufmerksamkeitsausdrücke (indicate expressions): Im Gegensatz zu den instrumentellen Ausdrücken möchte das Kind hier kein Objekt erhalten, so dass auch keine Frustration durch eine fehlende Reaktion entsteht. Ein Kind benutzte z.B. abwechselnd [bæ] und [dæ], wenn es sich selbst im Spiegel oder aus dem Fenster sah. Einen Monat später verwendete es [ʌh] mit Zeigegeste. Ein anderes Kind äußerte [dædi] mit immer gleicher Intonation, wenn es in Richtung der Tür zeigte oder Schritte hörte, die Form [didi] trat hingegen beim Zeigen auf Bilder in Büchern auf.
4. Gruppierungsausdrücke (grouping expressions): PCFs mit dieser Funktion beziehen sich auf spezifische Objekte und Gefühlsausdrücke. So wird zum Einen von einem Kind eine Gruppe verschiedener Spielzeuge einmal mit milder Frustration durch [ubiba], [ubibu], [upibu], [əbiba] benannt, während bei Zufriedenheit die selbe Gruppe mit [ədæ] und [dæ] bezeichnet wird. Zum Anderen benennt ein anderes Kind zwei verschiedene Objekte, die beide positiv besetzt sind, mit [babi].

Was die Relation zu Wörtern der Zielsprache angeht, sind für zwei Kinder, die innerhalb des Untersuchungszeitraums mehr als Einwortproduktionen äußerten, folgende Entwicklungen festzuhalten⁴³:

⁴³ In den Gesamtergebnissen der Studie auch sog. präsyntaktische Formen und zielkonforme syntaktische Formen enthalten. Diese wurden aus Gründen der Relevanz in der vorliegenden Arbeit nicht aufgenommen.

			PCFs			Wörter		
Kind	Sitzung	Alter	Type	Token	% Total	Type	Token	% Total
A	1	1;0	8	62	65,9	7	32	34,1
	4	1;3	8	41	24,6	24	97	58,0
	8	1;7	0	0	0	41	122	67,4
B	1	1;4	13	83	58,9	11	58	41,1
	4	1;8	1	10	3,5	90	199	68,8
	8	2;0	0	0	0	28	62	33,9

Tabelle 14: Produzierte PCFs und Zielwörter (modifiziert nach Dore et al. 1976).

Die Ergebnisse deuten Dore et al. als eine Progression, die mit PCFs beginnend zu komplexeren Formen führt. Ab der vierten Sitzung nehmen PCFs stark ab (Bei Kind A von 65,9% zu 24,6%; bei Kind B von 58,9% zu 3,5%), bei einer starken Zunahme von Zielwörtern (Bei Kind A von 34,1% zu 58%; bei Kind B von 41,1% zu 68,8%). PCFs und Zielwörter korrelieren demnach negativ miteinander. PCFs nehmen, so die Interpretation von Dore et al., im Entwicklungspfad in dem Sinne eine Brückenfunktion ein, als dass durch sie erstmals stabile phonematische Formen etabliert werden. Letztere beinhalten allerdings keine systematische Assoziationen mit konzeptuellen Strukturen, wie sie spätere Zielwörter zeigen. Dieses Herausbilden von zieladäquaten Verbindungen von Phonologie und dem konzeptuellen System sehen Dore et al. letztlich als die zentrale Aufgabe des Spracherwerbs an.

Die von Dore et al. gefundene Korrelation zwischen der Abnahme von PCFs und der Zunahme von Zielwörtern wurde von Robb et al. (1994) gestützt. In dieser Studie wurden sechs sprachunauffällige Kinder im Alter von 8 bis 25 Monaten mit Kaukasisch als Muttersprache in zwölf Sitzungen einmal im Monat aufgenommen. Zentral für die Auswertung der gegebenen Daten ist die Unterscheidung zwischen Wörtern (W) und Nichtwörtern (NW), wobei zu letzteren nicht nur PCFs, sondern auch alle anderen, phonetisch transkribierbaren, Lalläußerungen gezählt wurden. Der lexikalische Fortschritt wird mit einer Wort-Nicht-Wort-Ratio (W:NW-Ratio) gemessen. Hierbei wird die Anzahl der (zielsprachlichen) Wörter durch die Anzahl der Nichtwörter dividiert (in der Studie sollten mindestens zwei Wörter und Nichtwörter in einer Sitzung gegeben sein). Insgesamt produzierten die Kinder 4415 Einwortäußerungen (W=2336; NW=2079), wobei die Analyse mittels der Pearson-Produkt-Korrelation ergab, dass für jedes Kind bei einer stetig ansteigenden Entwicklungskurve der W:NW-Ratio eine signifikante Korrelation von Wort- und Nichtwort-Produktion vorhanden war ($p < 0.01$). Diese Korrelation respektive das stetige

Ansteigen der W:NW-Ratio bei jedem Kind zeigt an, dass die Zielwortproduktion mit Abnahme der Nichtwortproduktion sukzessive zunimmt.

In Tabelle 15 sind die Ergebnisse nach Maßgabe des chronologischen Alters dargestellt.

S1	NW	W	W: NW	S2	NW	W	W:N W	S3	NW	W	W:N W	S4	N	W	W: NW	S5	NW	W	W:N W	S6	NW	W	W:N W
8	58	0	0.0	8	66	0	0.0	8	41	0	0.0	11	0	0	0.0	13	45	0	0.0	14	64	0	0.0
9	43	0	0.0	9	60	0	0.0	9	45	0	0.0	12	52	0	0.0	14	59	0	0.0	15	22	16	0.72
10	51	0	0.0	10	54	0	0.0	10	50	0	0.0	13	30	10	0.33	15	37	27	0.72	16	22	28	1.27
11	35	4	0.11	11	32	9	0.28	11	56	0	0.0	14	64	11	0.17	16	27	19	0.70	17	19	29	1.52
12	41	30	0.73	12	67	11	0.16	12	42	8	0.19	15	45	28	0.62	17	45	3	0.06	18	15	23	1.53
13	31	36	1.16	13	19	20	1.05	14	20	37	1.85	16	35	23	0.65	18	38	25	0.65	19	45	53	1.17
14	27	31	1.14	14	35	40	1.14	15	31	19	0.61	17	25	42	1.68	20	32	20	0.62	20	29	23	0.79
15	16	31	1.93	16	14	49	3.50	16	4	39	9.75	18	27	18	0.66	21	19	38	2.00	21	15	47	3.13
16	32	56	1.75	17	9	85	9.44	17	17	62	3.64	20	25	51	2.04	22	20	53	2.65	22	8	31	3.87
17	17	26	1.52	18	9	45	5.00	18	19	55	2.89	21	20	68	3.40	23	12	71	5.91	23	12	42	3.50
18	23	62	2.69	19	13	71	5.46	19	12	70	5.83	22	8	73	9.12	24	9	80	8.88	24	13	57	4.38
19	12	43	3.58	20	5	69	13.80	20	5	85	17.00	23	18	102	5.45	25	6	74	12.33	25	6	58	9.66

Tabelle 15: Entwicklung der Wort-Nichtwort-Ratio in der Studie von Robb et al. (1994).

Wie Robb et al. (ebd.: 46) feststellen, sprechen ihre Daten gegen eine diskrete Phase zu Beginn der Wortproduktion, in der nur zielsprachliche Formen verwendet werden. Vielmehr betonen sie die Überlappung von Wort- und Nicht-Wortproduktion und befürworten entsprechend ein kontinuierliches Entwicklungsmodell.

Eine Brücken- oder Vorläuferfunktion von Protowörtern, wie sie Dore et al. (1976) annehmen, wird ihnen von Vihman & Miller (1988) nicht zugeordnet. In einer Longitudinalstudie mit 10 Kindern (Alter: 9-16 Monate) fanden sie, dass sich der Gebrauch erster konventioneller Wörter und die Lall- wie Protowortphase über einen längeren Zeitraum überschneiden. Da Protowörter in dieser Studie zwar benannt und in ihrer kommunikativen Funktion bestimmt werden, allerdings keine gesonderte phonologische Analyse erfolgt, kann eine phonologische Brückenfunktion aufgrund der methodischen Vorgehensweise m.E. auch nicht ermittelt werden. Dies gilt ebenfalls für die Studie von Piske (2001), in der die Protowörter von Kind LEO⁴⁴ unter Verweis auf die zielwortgleichen kommunikativen Funktionen zwar in die Datenanalyse eingehen und mit anderen Formen unter der Annahme artikulatorischer Muster analysiert werden, eine diskrete Beurteilung dieser Formen aus einem phonologischen Blickwinkel aber ausbleibt (ebd.: 78).

⁴⁴ Das andere Kind der Studie von Piske (2001), ISA, produzierte keine Protowörter.

Eine spezifische phonologische Analyse wie auch eine Untersuchung des Zusammenhangs von phonologischer Formen mit dem Lexikonerwerb wurde hingegen in einer Spontansprachstudie von Stoel-Gammon & Cooper (1984) durchgeführt. Die Daten stammten von drei Kindern, Daniel (Alter zu Beginn der Studie: 0;10), Sarah (Alter zu Beginn der Studie: 0;11) und Will (Alter zu Beginn der Studie: 0;11) und wurden bis zum Zeitpunkt des vollzogenen Erwerbs von 50 Wörtern erhoben (Alter von Sarah und Daniel: 1;4; Alter von Will: 1;7). Protowörter (in dieser Studie als *quasi-words* bezeichnet) werden hier zeitgleich mit den ersten zieladäquaten Wörtern für acht bis zehn Wochen bei Sarah und Daniel dokumentiert (Sarah [di:] und Daniel [na ~ ma ~ ã] für Aufforderungen). Will produzierte über den gesamten Zeitraum der Studie ein Dutzend solcher Formen und bildete auch im siebten Untersuchungsmonat noch neue Formen. Er produzierte sie ab dem ersten Monat, als auch die ersten zieladäquaten Wörter erschienen. Die Protowörter unterschieden sich von Zielwörtern dadurch, dass die meisten initial mit [d] gebildet wurden und aus reduplizierten CV-Silben bestanden. Deshalb unterschieden vor allem suprasegmentale Merkmale (Betonung, Intensität und Vokallänge) die Formen voneinander:

	Monat						
Zielwörter	1	2	3	4	5	6	7
Auto, Fahrrad	ŋŋ	ř	→	‡			
verlangen		dædæ	dado ~ dada	→	dædæ	didi	
Old MacDonald (Lied)			dadada ~ dʌdʌdʌ	→	→	→	→
Flugzeug, Vogel			dʌda	dida ~ didʌ	didi	→	→
Leute rufen			gɑ:jə:	→	‡		
Hundebellen			dida	→	→	‡	
nett			dʌ:	→	→	‡	
Hör auf zu bellen				dɪduba ~ dadaba	→	→	→
Allgemeiner Name					dɪdi	didi	→
Big Brown Buzzard (Spiel)					di:da:ba:	→	→
ja							di:
Bad							i:
Nackt, Schwimmen							aha

Tabelle 16: Protowörter von Kind Will.

Will zeigte unter den drei Kindern die langsamste lexikalische Erwerbsrate (50 Wörter mit 0;19,0). Stoel-Gammon und Cooper führen dies einerseits auf ein zu allgemeines CVCV-Muster ohne Beschränkung auf eine spezifische Konsonantenklasse, andererseits auf die vielen Protowörter zurück. Bei Daniel konnte hingegen die schnellste Erwerbsrate festgestellt werden (50 Wörter mit 0;16,2), da er vor allem Zielwörter mit CVC(V)-Struktur und einer spezifischen Konsonantenklasse (velare Plosive) produzierte und so neue Wörter diesem stabilen Muster anpassen konnte. Die dominanten Erwerbsmuster werden von Stoel-Gammon & Cooper in dem Sinne zusammengefasst, als dass velare Plosive und geschlossene Silben bei Daniel, offene Silben und phonetische Akkuratheit bei Sarah und reduplizierte Silben bei Will als hervorstechende Charakteristika der Produktionen identifiziert werden.

Daniel	Sarah	Will
Viele Phone in finaler Position	Wenige Phone in finaler Position	Wenige Phone in finaler Position
Velare Plosive am häufigsten in allen Positionen	Kein Phon besonders häufig produziert	Kein Phon besonders häufig produziert
CVC(C)-Struktur mit velaren Plosiven für viele Wörter verwendet	Kein häufiges Silbenstrukturmuster	CV-Reduplikationen häufig verwendet
Viele Zielwörter mit medialem oder finalelem velarem Plosiv	Kein klares Zielwortmuster	Viele Zielwörter und Onomatopoetika, die reduplikativ sind
Viele Homophone	Wenige Homophone	Wenige Homophone
Schnelle lexikalische Erwerbsrate	Lexikalische Erwerbsrate im Vergleich langsamer als diejenige von Daniel	Langsamste lexikalische Erwerbsrate
Ein Protowort	Ein Protowort	Viele Protowörter

Tabelle 17: Zusammenfassung der Produktionsmuster in Stoel-Gammon & Cooper (1984).

Wie oben dargestellt, wird der phonologisch strukturelle Aspekt von Stoel-Gammon und Cooper leider nicht quantitativ mit der Protowortproduktion korreliert. D.h., die Protowortproduktion wird zwar als Ursache der langsamen Erwerbsrate von Will vermutet, dies jedoch unabhängig von ihrem spezifischem phonologischen Status. Dass auch die von Will produzierten Protowörter als CV-Reduplikationen realisiert werden, bleibt interpretativ unbestimmt.

Ein solcher Zusammenhang zwischen Reduplikationen bzw. Silbenstruktur und Lexikonerwerb ist z.B. in Elsen (1991) dokumentiert. In einer Tagebuchstudie ihrer Tochter Annalena in einem Zeitraum von 0;8,23 bis 2;5 findet sie eine Korrelation von Wortschatzausbau und phonologischem Erwerb (Elsen 1999). Es ist hier der Ausbau des zielsprachlichen Frikativsystems in den Silbenrändern im Alter von 1;2, durch den die Produktion von Baby-Talk-Wörtern⁴⁵ rapide abnimmt, somit Homonyme abgebaut werden und ein Anstieg neuer produktiver Lexeme zu beobachten ist.

Menyuk und Menn (1979: 62) stellen bei Kind Jakob (vgl. Menn 1976) zwei Protowörter fest, die ebenfalls mit einer zugrundeliegenden CVCV-Struktur auftreten. Sie betonen allerdings die Variabilität auf der segmentalen Ebene. So wird [ioio] auch als [wiɔwio], [wejaweja], [iɔiɔiɔ], [ajaj] oder [ajɔajɔ] realisiert. Das Protowort für *thank you* wurde als [geika], [dejdɔ], [gigu][gaita] oder [dɛgɔ] produziert.

⁴⁵ Baby-Talk-Wörter sind nach Elsen (1991, 1999) definiert durch ein begrenztes Lautinventar, einfache Phonotaktik und Konsonant-Vokal-Reduplikationen, z.B. *pipi, dada, wawa*.

Ferguson (1978: 280) sieht insgesamt vier phonetische Typen in der Protowortproduktion gegeben:

1. einfacher oder wiederholter Vokal
2. silbischer Nasal (z.B. Robertos [m] (von Raffler-Engel 1965; zit. nach ebd.: 281))
3. silbischer Frikativ (z.B. Hildegards [ʃ] (Leopold 1947))
4. einfache oder wiederholte CV-Silbe mit Plosiv oder Nasal als Konsonant

In der Produktion von Vokalen, silbischen Nasalen und Frikativen, die in der Zielsprache nicht existent bzw. nur marginal vorhanden sind, sieht Ferguson eine beachtliche Asymmetrie zwischen Perzeption und Produktion. Dies vor allem darum, weil verschiedene Zielwörter von Kindern schon ab dem sechsten Lebensmonat an identifiziert werden. Die Annahme von zwei getrennten Lexika erscheint ihr an diesem Punkt nicht unwahrscheinlich. Ferguson et al. (1973) haben diese Möglichkeit schon aufgeworfen, bevor sie vor allem durch Menn (1983) systematisiert wurde. Sie erklären die Möglichkeit von Laut- bzw. Wortproduktionen, die in der Zielsprache nicht vorkommen, in ihrer sogenannten model-and-replica phonological grammar (MRPG) dadurch, dass nicht Phonemmerkmale, sondern Silbenmerkmale von der Grammatik berechnet werden. Das von ihnen untersuchte Kind Leslie produzierte im Alter von 0;11 nur totale CV-Reduplikationen, so dass diese Annahme wahrscheinlich ist.⁴⁶ Leslie produzierte fünf Wörter [mama], [dædæ], [gaga], [bæbæ] und [gægæ], die allesamt mit folgenden Silbenmerkmalspezifikationen erzeugbar sind:

Mermale	Silben					
	ma	bæ	dæ	ga	gæ	
labial	+	+	-	-	-	
velar	-	-	-	+	+	
front	-	+	+	-	+	

Tabelle 18: Silbenmerkmale nach Ferguson et al (1973).

Darüber hinaus macht diese (mentale) Merkmalsmatrix der MRPG die Vorhersage, dass potentiell zwölf verschiedene Silben [ba bæ da dæ ga gæ ma mæ na næ ɲa ɲæ] generiert werden und wiederum in den Reduplikationsprozess eingehen können. Leslie produzierte einige Monate später dann die im Englischen nicht existierende und phonotaktisch illegitime Form [ɲaɲə] für *song*, was Ferguson et al. als Bestätigung ihrer Annahmen ansehen (ebd.: 63-

⁴⁶ Interessanterweise wird auch Leslie ein langsamer Lexikonerwerb attestiert. Auch wird die Vermutung geäußert, dass dies mit ihrem langen Verwenden reduplikativer Formen zusammenhängt (vgl. Ferguson et al. 1973: 64).

64). Obwohl also das Zielmodell eines Wortes gegeben ist, wird es durch die spezifische Grammatik des kindlichen Sprachsystems dermaßen repliziert, dass durchaus neue, d.h. nicht in der Grammatik der Zielsprache gegebene Formen entstehen können. In diesem Modell wird somit die potentielle interne Kreativität der kindlichen Grammatik hervorgehoben. Dieser Aspekt findet sich in späteren Arbeiten zum Protowortlexikon, in der nicht mehr die Generativität der Grammatik, sondern Assoziationen in einem neuronalen Netzwerk im Mittelpunkt stehen, nicht mehr.

So beschreibt Levelt (1998) den Lexikonerwerb in der Produktion als eine sukzessiv stattfindende Assoziationsbildung. Vor Beginn der Protowortproduktion im Alter bis ca. 12 Monaten verläuft die Ausbildung von der konzeptuellen und der Silbenebene parallel, die Ebenen sind aber noch nicht miteinander gekoppelt. D.h., das Kind hat schon Konzepte für Zeit und Raum, Intentionalität und Kausalität erworben, produziert aber nur CV-Silbenfolgen oder einfache CV-Silben ohne Bedeutung. Zu dieser Zeit wird nach Levelt das mentale Silbenlexikon aufgebaut. Protowörter entstehen nach Levelt nun dadurch, dass aus dem mentalen Silbenlexikon akustisch-motorische Sprechsilben ausgewählt werden und mit einem konzeptuellen Referenten assoziiert werden (vgl. ebd.: 175). Sie sind nach dieser Theorie also sprechmotorische Ganzheiten assoziiert mit konzeptuellem Gehalt und eine zwangsläufige Erscheinung im frühen Spracherwerb. Soweit es aus den Darstellungen von Levelt nachvollziehbar ist, sind Protowörter demnach notwendiger Bestandteil eines einheitlichen Ziellexikons und phonologisch unanalysierte Silben der Zielsprache ohne genuin syntaktischen Status. Bis zur Mitte des zweiten Lebensjahres wird so ein Protolexikon aufgebaut, welches bis zu 50 Wörtern und mehr umfassen kann. Der entscheidende Prozess, der nun stattfindet, ist die Phonologisierung dieses Protolexikons. Auch Levelt (1994) geht davon aus, dass eine Phonologisierung des Lexikons die Ursache für den Übergang von einem kleinen Protolexikon zu dem zielsprachlichen Lexikonerwerb ist. In ihrem Ansatz ist es die Verteilung des Merkmales Artikulationsort, welches für lexikalischen Kontrast kritisch ist. Der Erwerb vollzieht sich nach ihr in vier Schritten:

1. Nur ein [Artikulationsort]-Merkmal (koronal oder labial) wird auf das ganz Wort verteilt.
2. Der Onset des Wortes wird für ein weiteres [Artikulationsort]-Merkmal spezifiziert.
3. Die Koda wird für ein weiteres [Artikulationsort]-Merkmal spezifiziert.
4. Der Nukleus wird für ein weiteres [Artikulationsort]-Merkmal spezifiziert.

Sind Protowörter zunächst also nur für einen einen Artikulationsort spezifiziert wie in [tin] für koronal und [pim] für labial, so ist zunehmend ein freierer Gebrauch des Merkmals Artikulationsort zu beobachten. Dieser Phonologisierungsprozess ist dann mit ca. 2;6 Jahren weitestgehend beendet und es werden anschließend die syntaktischen Informationen, d.h. die entsprechenden Lemma-Knoten assoziiert. Der hier beschriebene Prozess der Ebenenassoziation ist in Abbildung 45 dargestellt.

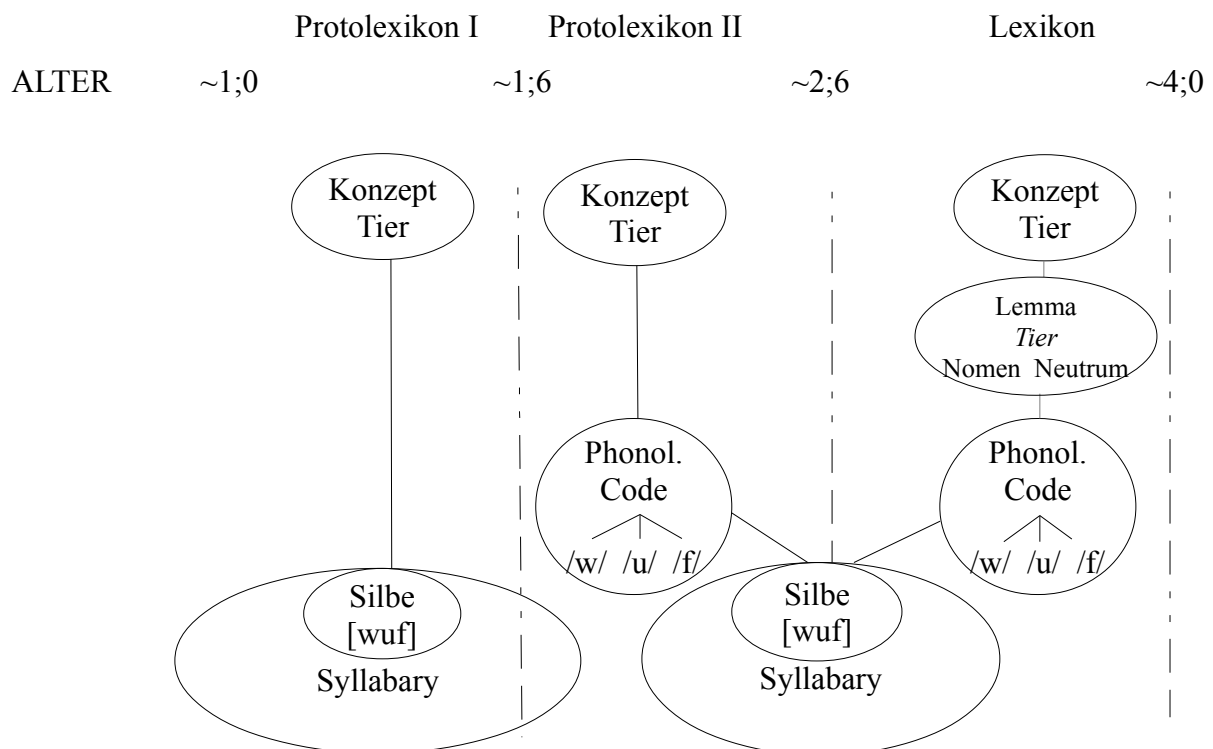


Abbildung 45: Aufbau der Lexikons nach Levelt (1998).

Zum Ende des Abrisses zum Stand der Erforschung von Protowörtern soll die neueste Veröffentlichung von Vihman (2014) genannt werden, die ein eigenes Kapitel über Protowörter enthält. Bezeichnend für die karge Forschungslage ist hier, dass so gut wie keine neuen Studien oder Erkenntnisse im Vergleich zu Vihman (1996) verzeichnet sind. Selbst die eingangs angeführte Definition von Protowörtern nach Vihman ist unverändert geblieben.

Wir können nun abschließend aus der Besprechung der obigen Untersuchungen zu Protowörtern und frühen Wortproduktionen allgemein folgende Hauptaspekte und Fragestellungen herausfiltern:

Lexikon:

1. Die Produktion von Protowörtern scheint nach Dore et al. (1976), Robb et al. (1994) und Stoel-Gammon & Cooper (1984) mit dem Aufbau eines Zielwortlexikons negativ zu

korrelieren, d.h., erst durch die sukzessive Abnahme der Protowortproduktion werden vermehrt Zielwörter realisiert.

2. Der lexikalische Status von Protowörtern ist in Bezug auf (1.) unklar. Nehmen alle Studien eine Überschneidung von Proto- und Zielwortproduktion an, so stellt sich die Frage, ob die Protowortproduktion eine notwendige Voraussetzung für den Aufbau des zielsprachlichen Lexikons ist, wie Dore et al. (1976) und Levelt (1998) sie annehmen. Wenn dies der Fall ist, wieso produzieren sie nicht alle Kinder in gleichem Maße und wieso begünstigt eine häufige Protowortproduktion dann eine negative Korrelation wie im Falle von Will in der Studie von Stoel-Gammon & Cooper (1984)?
3. Eine eindeutige modelltheoretische Einordnung von Protowörtern bleibt aus. Nur in Ferguson (1978) findet man den Hinweis auf die Dissoziation von Input- und Outputformen und damit auf die Option zweier unterschiedlicher Lexika.

Struktur:

4. Protowörter werden von vielen Autoren als nicht-phonologisch klassifiziert und somit konzeptuell als phonetische Ganzheiten konzipiert. Erst im Zuge eines Phonologisierungsprozesses werden diese als segmentale Einheiten analysiert (Dore et al. 1976, Levelt 1998, Piske 2001, Vihman & Miller 1988). Eine Ausnahme bilden Ferguson et al. (1973), die allerdings auch silbische statt segmentaler Merkmale in der frühen Produktion annehmen.
5. Dadurch, dass Protowörter häufig als (phonetische) Ganzheiten konzipiert werden, ist die Struktur dieser Formen insgesamt unklar. Folgende Charakteristika sind für diese Formen aus der Literatur zu entnehmen:
 - einheitlicher Artikulationsort
 - CV-Reduplikationen mit Plosiv oder Nasal als Konsonant
 - einfacher oder wiederholter Vokal
 - silbischer Nasal
 - silbischer Frikativ

Insgesamt bietet sich somit ein recht uneinheitliches Bild davon, was unter einer phonetisch-phonologischen Perspektive als Protowort zu verstehen ist. Nur die Beobachtung, dass Protowörter meist aus CV-Reduplikationen bestehen, ist bei allen Autoren die definitive Schnittmenge. Man kann sich entsprechend nicht des Eindrucks erwehren, dass das konzeptuelle Präjudiz, frühe Wortproduktionen mit Jakobson (1941) weiterhin als eine Form von Lautgebärden zu behandeln, die Frage danach, was ein Protowort denn nun sei, um kein Jota weiter gebracht hat und für eine wissenschaftliche Analyse ein wesentlich feinkörnigeres Instrumentarium vonnöten ist.

Was empirisch als Protowort zu werten ist, hängt natürlich auch damit zusammen, welche Wortformen man als bedeutungstragend respektive mit semantischem Gehalt klassifiziert. Auch wenn der semantische Aspekt in dieser Arbeit kein ausdrücklicher Untersuchungsgegenstand ist, so ist er für die Bestimmung früher phonologischer Wortformen und lexikalischer Repräsentationen eindeutig von ergänzender Wichtigkeit. Vor allem eine Umschreibung von Menyuk & Menn (1979: 62-63) zur Bedeutung von Protowörtern ist hier erhellend:

„They are, at this early stage, essentially vocal signals, and may be compared to adult words which have a very limited pragmatic range, like greetings and cries of *ouch*. The meanings of such items, for both adult and child are best characterized as 'what you say when you do X'. Proto-words do not remain restricted in this fashion, however, for more than a few months. They develop what is called 'symbolic autonomy' – the potential for being used for a variety of ends. Even the most impoverished of the adult language (e.g. *ouch*) has some degree of symbolic autonomy, for an adult can use it in reporting what he/she or another said as well when actually communicating present pain. [...] As a proto-word develops the full degree of symbolic autonomy possessed by a word of the adult language, it must be able to be used with a variety of intonations – raising, falling, and more complex contours as well. In this way, the semantic-pragmatic development is interwoven with the suprasegmental aspect of phonological development.“

Wie soll man sich eine solche Interaktion zwischen dem semantischen und phonologischen System modelltheoretisch vorstellen, dass sich die Bedeutung von „Was man sagt, wenn man X tut“ zu voller symbolischer Autonomie entwickelt? Und ist es vielleicht nur aufgrund der konzeptuellen Entwicklung, dass Protowörter von Kindern verwendet werden?

Vor allem die Arbeiten von Halliday (1975, 1979) waren für die Untersuchung des semantisch-pragmatischen Aspekts von Protowörtern wegweisend und sollen im nächsten Kapitel einführend dargestellt werden. Ausgehend von diesen gilt es auch für den semantisch-pragmatischen Erwerb, einen konzeptuellen Rahmen zu erarbeiten, der von seiner explanativen Reichweite hinreichend ist, um das Phänomen der Protowortproduktion adäquat zu erfassen.

3.2 Semantische und lexikalische Aspekte der Protowortproduktion

3.2.1 Erste Klassifikationen in der Spracherwerbsforschung

Wenn Menyuk und Menn (1979) hinsichtlich der Entwicklung der Protowortproduktion von einem Verlauf beginnend mit der Bedeutung „Was man sagt, wenn man X tut“ hin zu einer symbolischen Autonomie sprechen, könnte man versucht sein, Gebrauch und Bedeutungsentwicklung der Protowörter im Sinne Piagets auf die allgemeine kognitive Entwicklung zurückzuführen. So schreibt z.B. Szagun (1980: 202):

„Während der Einwortphase - und diese kann schon mit 9 Monaten beginnen - ist das Kind dabei, sich von einem sensomotorischen, auf Objekte und Ereignisse mit Handlungen reagierenden Wesen zu einem geistigen Wesen zu entwickeln. Aber die Erkenntnisse, die es mit den ersten Worten enkodiert, sind immer noch Erkenntnisse, die sich auf Handlungen beziehen [...].“

Diese Abbildbarkeit der allgemein kognitiven Entwicklung auf die Entwicklung sprachlicher Formen kann allerdings als widerlegt angesehen werden. Inzwischen gibt es solide Evidenz dafür, dass Objekte, Handlungen und Ereignisse von Kindern wesentlich früher wahrgenommen und repräsentiert werden, als es in der Theorie von Piaget angenommen wurde (für einen Überblick s. Hennon et al. 2000). Wie z.B. die Arbeiten von Baillargeon et al. (1985) und Baillargeon (1987) gezeigt haben, ist die Objektpermanenz, also das Wissen darum, dass Objekte unabhängig von der eigenen Wahrnehmung existieren, nicht wie bei Piaget erst ab dem neunten Lebensmonat im Entstehen begriffen (und mit 18-24 Monaten vollständig etabliert), sondern findet sich schon bei Kindern im Alter von dreieinhalb bis sechs Monaten. Baillargeon (ebd.: 662) betont, dass Objektpermanenz somit nicht auf die erst später stattfindende Koordination sensomotorischer Schemata zurückgeführt werden kann. Plausibel erscheint deshalb z.B. die Annahme von Spelke (1994), dass angeborene Beschränkungen wie Kohäsion, Kontinuität und Kontakt schon die frühe Objektwahrnehmung steuern. Allerdings ist dieses Wissen um das Objekt bei dreieinhalb Monate alten Kindern nur für wenige Sekunden präsent und wird erst mit steigendem Alter zeitlich länger behalten. Es ist also nicht das Vermögen selbst, was den Kindern fehlt, sondern eine Gedächtnislimitierung, ein klassischer Performanzfaktor, beschränkt die gegebene Kompetenz. Dies zeigen ebenfalls Untersuchungen von Diamond (1985) zu dem A-nicht-B-Irrtum aus Piagets Studien. Piaget schlussfolgerte aus dem Suchverhalten von 8-11 Monate alten Kindern, die bei mehrmaligem Sucherfolg an Stelle A weiter dort suchen, auch wenn ein Gegenstand vor ihren Augen an Stelle B versteckt wurde, dass die Kinder wesentlich von Handlungswissen und nicht von Wahrnehmungswissen gesteuert werden. Diamond zeigt nun, dass die Verzögerung vor dem Suchen entscheidend ist. Suchen acht Monate alte Kinder sofort, finden sie das Objekt auch an richtiger Stelle. Bei einer Verzögerung von zwei bis drei Sekunden suchen sie jedoch an der Stelle, wo sie früher Erfolg hatten. Für die Protowortbedeutung „Was man sagt, wenn man X tut“ ist dies insofern relevant, als dass sie eventuell weniger in einer nicht-symbolischen Bedeutungszuweisung begründet ist, sondern aufgrund der Gedächtnislimitierung die Bedeutungsrepräsentation sehr stark an die aktuelle Evokation durch einen spezifischen Kontext gebunden ist.

Den wesentlichen Unterschied zwischen der Erwachsenensprache und Protosprache, wie er letztere nennt, sieht Haliday (1979: 180) in Einklang mit obigen Ausführungen auch nicht in der grundlegenden Bedeutungszuweisung, sondern in der Einführung einer weiteren sprachlichen Informationsebene:

„[...] protolanguage as a system [...] consists of a semantics and a phonology (or other expressive means) with no lexicogrammar in between. Its elements are discrete signs, each one being a direct pairing of content with expression; there are no words, and no structures. The move into the adult language lies in introducing a new, purely abstract level of coding as intermediary between the two, so that meanings are coded as (lexicogrammatical) forms [...].“

Bedeutungstragende Akte (acts of meaning) werden nach Halliday bereits im Alter von neun Monaten ausgeführt. Er klassifiziert sie als symbolisch (ebd.: 172), da nicht mehr die Handlungen an sich das Ziel sind, sondern sie nun dazu dienen, eine spezifische Intention zu realisieren. Die ersten bedeutungstragenden Akte sind nach seinen Beobachtungen nicht auf sprachliche Äußerungen beschränkt, da sie auch gestisch transportiert werden können. Folgende Bedeutungen wurden z.B. von Hallidays Sohn Nigel im Alter von neun Monaten über den gestischen Modus ausgedrückt (ebd.: 173; eigene Übersetzung):

'Ich will das'	kurzes Anfassen des Objekts, dann loslassen
'Ich will das nicht'	leichtes und kurzes Anfassen des Objektes
'Mach etwas damit, so dass ich zuschauen kann'	festes Anfassen des Objekts für eine messbare Zeit

Zur gleichen Zeit allerdings produzierte er auch zwei Vokalisierung, die ebenfalls mit unterschiedlichen Bedeutungen versehen waren (ebd.: 173; eigene Übersetzung):

'Lass uns zusammen sein'	Vokal (halb geschlossen, vorne, gerundet) mit leicht fallendem Ton
'Das ist interessant'	Vokal (halb geschlossen, vorne, gerundet) mit mittelmäßig fallendem Ton

Schon ab diesem Zeitpunkt sieht Halliday in den gestischen und vokalischen Akten ein Set aus vier semantischen Funktionen realisiert:

- 1) Instrumental: Ziel, ein Objekt zu bekommen.
- 2) Regulatory: Ziel, eine Person zu einer Handlung aufzufordern.
- 3) Interactional: Ziel, eine Interaktion mit einer Person herzustellen.
- 4) Personal: Ziel, die eigene innere Teilhabe an einer Situation kund zu tun.

Schon ab dem Alter von 10 Monaten nutzte Nigel nur noch den sprachlichen Modus, um Bedeutung zu transportieren.

In Anbetracht der Ergebnisse von Halliday im Vergleich zu den Studien, welche die phonologische Form in den Mittelpunkt der Untersuchung stellen, kann man erneut die Frage aufwerfen, ab wann Protowörter angenommen und wie sie definiert werden sollen. Hallidays Definition ist, wie man unschwer erkennen kann, relativ agnostisch gegenüber der phonologischen Form, weshalb hier Formen als Protowörter klassifiziert werden, die eindeutig keine phonotaktischen und segmentalen Zielstrukturen beinhalten müssen. Wie zu Anfang des Kapitels schon einmal erwähnt, setzt sich für ihn Protosprache wie folgt zusammen (Halliday 1979: 180):

„[...] protolanguage as a system [...] consists of a semantics and a phonology (or other expressive means) with no lexicogrammar in between.“

Allerdings fällt in dieser Definition eine wichtige klassifikatorische Unterscheidung weg. Phonologie ist selbstredend nicht gleichzusetzen mit anderen expressiven Mitteln, sprich motorischen Routinen wie Gesten oder solchen der Artikulation. Sattelt man das Pferd also nur von der semantisch-pragmatischen Seite auf, wird nun alles uniform als Protowort klassifiziert, was als bedeutungstragender Akt eingeordnet werden kann. Naheliegender wäre es demnach, auch in der Protowortproduktion gemäß der Splitting-Hypothese (vgl. Kap. 1.5.2) selbst noch einmal diskrete Phasen zu unterscheiden, ein Vorgehen also, das bei Halliday schon angelegt ist, aber leider terminologisch nicht klar ausdifferenziert wird. In Anlehnung an Jackendoff (2002) kann man die früh produzierten Protowörter von Nigel oder auch, um ein anderes Beispiel zu nennen, das Grunzgeräusch [ʔəʔ] von Leopolds Tochter Hildegard, das im Alter von 0;8 bis 1;6 Jahren als demonstrative Interjektion verwendet wurde (Leopold 1947), als folgendes Informationskonglomerat wiedergeben:

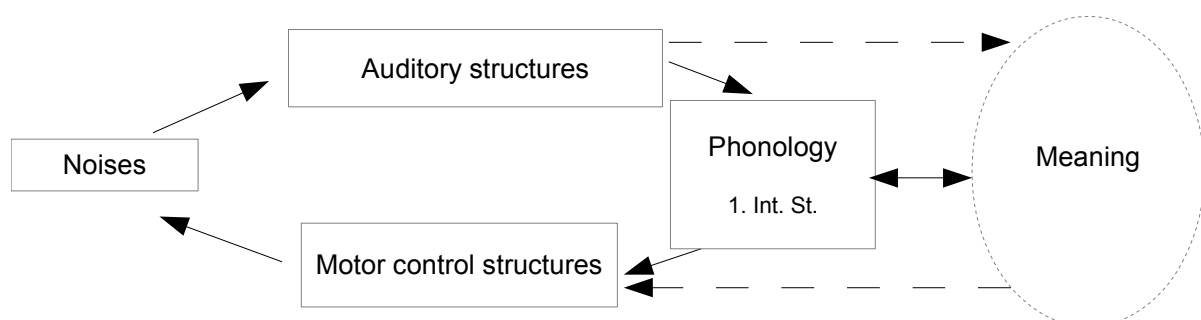


Abbildung 46: Verknüpfung von Subsystemen in der Produktion von frühen Protowörtern.

Sowohl bei Hildegards Grunzgeräusch als auch Nigels reinen Vokaläußerungen existiert ein Silbengipfel, so dass das Merkmal [\pm silbisch] in der Produktion demnach wirksam ist (1. Integrationsstufe (1. Int. St.), allerdings sind in den Äußerungen noch keine phonotaktischen Regularitäten gegeben, so dass die Oberklassenmerkmale (2. Integrationsstufe) in der Produktion noch nicht aktiv sind.

Es kann hier nicht genug betont werden, dass es sich bei dieser Darstellung nur um die Produktionspfade im Spracherwerb handelt, d.h., hier wird illustriert, welche Informationen in der Produktion bereits aktiv integriert und nutzbar gemacht wurden. Perzeptiv sind, wie bereits in vorigen Kapiteln dargelegt wurde (vgl. z.B. Kap. 2.2.1), die wesentlichen (ziel-)sprachlichen Strukturen bereits etabliert.

Nimmt man die Entwicklungsperspektive assoziierter und sich sukzessiv etablierender Subkomponenten (vgl. Jackendoff 2002: 264) ernst, so muss man m.E. auch die Bedeutungsebene für die Analyse früher Wortproduktionen ergänzend in Augenschein nehmen. Die Frage, die hier vor allem maßgeblich ist, wäre demnach, ob es frühe und systematisch verwendete Routen von Phonologie-Semantik-Paarungen gibt, die Vorläufer späterer zielsprachlicher Verbindungen sind, aber eventuell für die Zwecke einer detaillierten psycholinguistischen Beschreibung der Protowortproduktion noch nicht als fruchtbar erkannt wurden. Um dies allerdings bestimmen zu können, müssen zuvor grundlegende Fragen zum Bedeutungserwerb geklärt werden.

3.2.2 Das Problem der semantischen Analyse allgemein und im Spracherwerb

Hallidays Ansatz verschiedener bedeutungstragender Akte unterscheidet, wie wir oben gesehen haben, Wörter nach Maßgabe ihres funktionalen Gebrauchs. Ähnlich verfahren Vihman & McCune (1994), auch wenn sie ein anderes Set an Unterscheidungskriterien zugrundelegen (vgl. ebd.: 528). Die Begründung für ein solches Vorgehen ist im Grunde der Bewertung von frühen Wortformen aus phonologischer Perspektive analog: Genau so, wie Protowörter aufgrund ihrer fehlenden phonologischen Zieladäquatheit als motorische Routinen oder Muster und somit eben als nicht primär phonologisch klassifiziert werden, so werden sie auch aufgrund einer fehlenden zielsprachkonformen Referenz nicht als konzeptuell strukturiert, sondern vielmehr als unanalysierbare Akte vorgestellt. Velleman & Vihman (2002: 10) fassen die Grundzüge dieses Ansatzes, der im Erwerb auf allen

sprachlichen Ebenen den Ausgang von ungegliederten Templates bzw. Mustern postuliert, wie folgt zusammen:

„According to the middle-level entry account, then, knowledge of language is not built up brick by brick from the atoms of form or meaning - that is, the child does not first learn as individual units, roots and affixes, or semantic features such as “round shape” or “four-legged,” or phonetic segments or features, and then assemble them into words. Rather, the learner begins at the word (or phrase) level of each component (syntax, semantics, phonology). The assumption that the child begins at the middle level, which is immediately available in the speech the child hears, eliminates the need to posit the universal stock of linguistic units, segments, features, principles, rules, or word-learning constraints that some assume to be innately available [...].“

Wie auch in der Erforschung der phonologischen Entwicklung haben sich Studien des Wortbedeutungserwerbs im Rahmen der Generativen Grammatik vornehmlich auf die Syntax-Semantik-Schnittstelle konzentriert, sprich auf den Erwerb von Zielwörtern, die kategorial (Nomen, Verb usw.) klassifiziert werden können und auf semantische Relationen innerhalb von (prä-)syntaktischen Strukturen bzw. auf den möglichen Einstieg in die Syntax durch Wortbedeutungen. So wie Smith (2010: 53) für die Produktion von phonologischen Zielwörtern festhält „His real phonology was beginning“, konstatiert Kauschke (2000: 11) in Bezug auf den Worterwerb: „Mit etwa 12 Monaten tauchen die ersten „echten“ Wörter auf“. Als ein wesentlicher Unterschied wird bei Protowörtern ihr kontextgebundener Gebrauch angesehen:

„So äußert das Kind das Pseudowort **tschtschu** nur in einer ganz bestimmten Situation, in der es mit einem bestimmten Zug spielt und diesen bewegt. Das Wort wird nicht verwendet, wenn es um denselben Zug geht und dieser nicht bewegt wird.“ (Rothweiler, 2001: 45)

An diesem Beispiel erkennt man sehr schön, dass man auch in der Erforschung des Bedeutungserwerbs an einigen tradierten Unterscheidungen und Definitionen festhält, die feinkörnige strukturelle Entwicklungen nicht erfassen und dem generativen Potential des sich entwickelnden Sprachsystems in keinsten Weise gerecht werden. Natürlich ist es der Fall, dass Wortformen erst einmal nur für spezifische Objekte angewendet werden. Die entsprechenden Objektrepräsentationen müssen auf Basis von Erfahrung selbstverständlich erst erworben, aufgebaut und stabil repräsentiert werden. Eine wichtige strukturelle Unterscheidung wird aber in obigem Beispiel ausgedrückt, nämlich die zwischen einem sich bewegenden Ding und einem unbewegten Ding. Der Zug bildet hier vorerst das einzige Bezugsobjekt auf den diese wichtige Unterscheidung sprachlich bezogen wird. Nichtsdestotrotz zeigt sich hier eine zugrundeliegende Struktur, in welche später auch andere Objekte als mögliche Variablen

inseriert werden. Die wichtige Frage in der Analyse von frühen Wortbedeutungen scheint mir also zu sein, was als primäres Kriterium für solche definiert wird. Anders formuliert also: Bestimmt man als definatorisches Kriterium Wortgebrauch und Objektreferenz oder eine zugrundeliegende konzeptuelle Struktur. Dass letztere sich erst entwickelt, wenn eine annähernd zielsprachliche syntaktische und phonetische Etikettierung vorliegt, dürfte bei dem reichen konzeptuellen System von Säuglingen, wie die Forschung der letzten Jahre es offengelegt hat, sehr unwahrscheinlich sein. Im Rahmen sogenannter Kernwissenstheorien (vgl. besonders Carey 2009; Spelke & Kinzler 2007) geht man inzwischen auf Basis einer reichen Datenbasis davon aus, dass domänenspezifische Input-Analyse-Systeme angeboren sind und somit eine frühe Identifikation funktional relevanter Entitäten garantieren. In einem solchen Sinne leiten implizite Kernprinzipien im Laufe der Entwicklung die Elaboration objektbezogener kognitiver Prozesse (für eine ausführlichere Darstellung siehe Sodian 2014). Tabelle 19 listet domänenspezifische Kenntnisse, deren Auftreten im Kleinkindalter nachgewiesen werden konnten.

Kategorien	Globale Unterscheidung belebt-unbelebt	4-7 Monate
	Basic-level Unterscheidungen	9-12 Monate
	Sequential Touching	12-24 Monate
	Sortieren	Ab 18 Monaten
Mathematik	Repräsentation kleiner Mengen (bis 4 Elemente)	Ab 3-4 Monaten
	Intermodale Wahrnehmung von Mengen (Töne und Aktivitäten)	Ab 3-4 Monaten
	Repräsentation der Identität von mehr als einem Objekt	Ab 12 Monaten
Physik	Wissen über Solidität, Kontinuität und Trägheit	Ab 3-4 Monaten
	Differenzierung des Wissens über Unterstützungs- und Inhaltrelationen sowie Schwerkraft	1- Lebensjahr
Psychologie	Triadische Ausmerksamkeit	Ab 7-9 Monaten
	Emotional and social referencing	Ab 7-9 Monaten
	Frühes Intentionsverständnis	Ab 11-12 Monaten

Tabelle 19: Frühes domänenspezifisches Kernwissen nach Pauen & Vonderlin (2007: 22).

Es ist an dieser Stelle nicht möglich, auf die umfassende Datenlage und Literatur zum Wortbedeutungserwerb einzugehen. Prinzipiell kann man aber konstatieren, dass in der sogenannten Einwortphase (aber auch im späteren Erwerb) nicht-strukturelle Analysen, d.h., einfache Einteilungen in Wortkategorien allgemein üblich sind. Sei es die Verwendung der

klassischen Analysekriterien nach Bloom (1973) und Nelson (1973) oder der Kriterien in Kauschke (2000). Letztere unterscheidet z.B. eine Liste von 9 Wortkategorien (Nomen, Verben, Adjektive, personal-social words, relationale Wörter, Pronomen, Funktionswörter, Onomatopöien, Sonstige), die in weitere Subkategorien eingeteilt werden (ebd.: 109).

Elaborierte Analyseverfahren im Sinne einer kompositionellen Semantik sind m.W. innerhalb der Spracherwerbsforschung für die Beschreibung von frühen Wortbedeutungen nicht Usus, obwohl z.B. mit der Konzeptuellen Semantik von Jackendoff (2002) und auch Pinker (1989) ein ausgearbeitetes theoretisches Instrumentarium zur Verfügung steht. Allerdings ist in diesen Theorien der Fokus einseitig auf die Analyse von Verbbedeutungen gerichtet, so dass vornehmlich die späte Interaktion von semantisch-konzeptioneller Struktur mit syntaktischen Konstruktionen im Vordergrund steht.⁴⁷ Die Applikation solcher struktureller Analysen auf Erwerbsdaten des frühen Wortbedeutungserwerbs steht entsprechend noch aus und kann im Rahmen dieser Arbeit nicht ab initio geleistet werden.

Insgesamt bietet sich in Hinblick die Erforschung des frühen Bedeutungserwerbs leider ein ernüchterndes Bild. Löbach (2000: 5) erstellt eine wissenschaftliche Standortbestimmung, die so wohl auch noch heute gelten kann. Die wesentlichen Arbeiten zum Bedeutungserwerb erschienen in den 70er Jahren des 20. Jahrhundert und es muss auch heute immer wieder auf diese referiert werden:

„Symptomatisch für diese Situation ist schon die spärliche Behandlung des Semantikerwerbs in Standardwerken zur Psycholinguistik bzw. Kognitiven Linguistik. [...] Auch Meibauer/Rothweiler (1999) müssen noch auf diese Arbeiten verweisen. Sie konstatieren: „Wie die Bedeutung eines einzelnen Wortes entsteht...[ist] noch immer weitestgehend ungeklärt“ (Meibauer / Rothweiler 1999: 20). Ihre generelle Charakterisierung des Erwerbsprozesses ist entsprechend vage bis geheimnisvoll: „Der Bedeutungserwerb ist dynamisch, lang andauernd und verläuft verdeckt, und er basiert auf der Interaktion zweier sich entwickelnder Systeme, des kognitiven und des linguistischen Systems“ (Meibauer / Rothweiler 1999: 19).“

Was die Erforschung des Bedeutungserwerbs angeht, so dominierten innerhalb der Forschung in den vergangenen Jahren Ansätze, die sich auf gut quantifizierbare Erwerbsergebnisse, wie z.B. dem *fast mapping* konzentrierten. Letzteres ist im Grunde nur das Etikett für einen Prozess des schnellen passiven Wortlernens, in dem das Kind saliente Wortformmerkmale wie

⁴⁷ In der theoretischen Semantik scheint es inzwischen insgesamt gebräuchlich zu sein, aus komplexen Sätzen, die Bedeutung einzelner Wörter zu rekonstruieren: „As it has turned out in the history of the discipline, it is not possible to develop a method of doing so by simply manipulating semantic primitives like markers. Rather, one might instead pursue the opposite strategy, starting with a comparison between the meanings of entire sentences and then finding out more about the meanings of their parts and how they combine.“ (Zimmermann & Sternefeld 2013: 24) Bei einer solchen Konzentration auf elaborierte Sätze erwachsener Sprecher ist es kaum verwunderlich, dass es zum frühen Bedeutungserwerb kaum Arbeiten gibt, sondern hauptsächlich nur zu einem Erwerbsalter, in dem Semantik in der Produktion bereits relativ reibungslos über Syntax „supervenieren“ kann.

Silbenanzahl oder Anlaute nutzt, um eine erste Repräsentation mit unscharfer Wortbedeutung aufzubauen, die aber in einer längeren Phase ausgebaut werden muss (vgl. Rothweiler & Meibauer 1999: 20). Der Bedeutungserwerb wird durch Beschränkungen gesteuert, wie z.B. die *whole object assumption*, nach der Wörter sich auf ganze Objekte beziehen und nicht nur auf Teile oder Eigenschaften derselben (vgl. Markman 1989). Wie unschwer zu erkennen ist, wurde für die Analyse des frühen Worterwerbs durch diese Konzepte wenig gewonnen. Fast mapping beschreibt nur, wie schnell und unter welchen Bedingungen Wörter erworben werden, aber eben nicht, was genau erworben wird. Ähnlich verhält es sich mit den Beschränkungen. Der taxonomic constraint besagt z.B. folgendes:

„Labels refer to objects of the same kind rather than to objects that are thematically related.“
(Markman 1993: 155)

Wie aber Löbach (2000: 39) wiederum ausführt, ist das Aufstellen einer solchen Beschränkung für die eigentliche Untersuchung der Wortbedeutung irrelevant:

„M.E. ist dies nun aber offensichtlich keine Antwort auf die Frage, wie Wörter auf Objekte derselben Art angewendet werden; der constraint besagt nicht, wie eine Kategorie ausgebildet wird, oder welcher Mechanismus dafür verantwortlich ist, dass Entitäten als Mitglieder derselben Kategorie ausgewiesen werden können. Er beantwortet m.E. - kurz gesagt - keine der relevanten Fragen zum Bedeutungsbegriff oder zum Erwerbsprozess selbst.“

Offenbar befindet man sich also in Bezug auf die Wortbedeutungsanalyse in einem rechten Dilemma. Auf der einen Seite findet man die erstaunlichen Ergebnisse und Fortschritte der Säuglingsforschung vor, auf der anderen Seite ist das psycholinguistische Instrumentarium zur Beschreibung einfacher Wortbedeutungen innerhalb der letzten 40 Jahre eigentlich wie nicht vorwärts gekommen. Jackendoff betont ebenfalls die jüngsten Erkenntnisse in der Säuglingsforschung und kommt zu folgendem Schluss:

„In short, prelinguistic children and children in the very earliest stages of learning language apparently have concepts that structure both their understanding of events they observe and their attempts to assign meanings to words. Again, this gives us an independent fix on a theory of semantics: it should have at its foundations the kinds of concepts that prelinguistic children exhibit.“ (Jackendoff 2014)

Wie genau soll nun allerdings eine strukturelle Analyse früher Wörter aussehen, die berücksichtigt, dass Säuglinge und Kleinkinder Konzepte unterhalten, die ihre Beobachtungen und Wortbedeutungen strukturieren, aber aufgrund fehlender phonologischer Zielform und fehlender Objektreferenz nicht einmal ansatzweise entschieden werden kann, welche wortkategoriale Information den jeweiligen Produktionen unterliegt? Es gilt an dieser Stelle

noch eingehender die Möglichkeit einer semantischen Analyse früher Wortbedeutungen zu untersuchen. Wesentlich ist hierbei die Frage, ob eine semantische Analyse analog zur phonologischen mittels definierter Merkmale vorgenommen werden kann.

3.2.3 Fodor und die Komponentialsemantik

Frühe Untersuchungen zum Wortbedeutungserwerb versuchten im Rahmen der Komponentialsemantik die noch unkonkrete und in diesem Sinne „weite“ Bedeutung früher Produktionen durch den sukzessiven Ausbau universeller semantischer Primitiva wie [\pm belebt] oder [\pm mensch] zu erfassen (z.B. Clark 1973). Merkmalsbasierte Ansätze bzw. Ansätze der Komponentialsemantik standen und stehen allerdings häufig unter massiver Kritik. Starke Argumente gegen solche Ansätze finden sich vor allem in Fodor et al. (1975) und Fodor et al. (1980), die einen weitreichenden Einfluss auf die Erforschung semantischer Strukturen (und damit auch derselben im frühen Spracherwerb) hatten⁴⁸ und eng mit obiger Standortbestimmung zusammenhängen.⁴⁹ Eine kurze Zusammenfassung nach Piatelli-Palmarini (1994: 330-332) soll die Argumentation Fodors verdeutlichen. Die Fragestellung ist, wie Kinder neue Konzepte, z.B. [grau und quadratisch] aus den Einzelkonzepten [grau] und [quadratisch], lernen können. Es wird angenommen, ein Proband müsste anhand von Abbildungen grüner Kreise, grauer Quadrate und blauer Dreiecke usw. eine Zuordnung zu zwei Stapeln vornehmen, in der einer für die richtige Kategorie steht bzw. die richtigen Exemplare enthält (d.h. graue Quadrate). Diese Zuordnung steht in diesem Gedankenexperiment für das Lernen eines neuen Wortkonzepts *miv*. Die induktive Hypothese wäre demnach:

X ist ein Exemplar, wenn und nur wenn X ist Y.

Das Ziel des Lernprozesses wäre auf den Worterwerb bezogen, dass

Y = *miv* = „quadratisch und grau“.

Der Lernprozess durchläuft nun folgende drei Stufen:

⁴⁸ Piatelli-Palmarini (1994: 330) ging soweit, dass er Fodors Argumentation von 1980 im Nachhinein als „jeu de massacre“ bezeichnete.

⁴⁹ „Tatsache ist jedoch, dass die Semantikerwerbsforschung der letzten 20 Jahre im Schatten dieser Argumentationen steht und durch sie geprägt wird.“ (Löbach 2000: 6)

Stufe 1	[grau], [quadratisch]
Stufe 2	$Y \neq miv$
Stufe 3	$Y = miv =$ [grau und quadratisch]

Auf Stufe 1 besitzt der Lerner noch nicht das Konzept [grau und quadratisch], sondern nur Primitiva, während er auf Stufe 2 eine vorläufige Hypothese für das Konzept Y erstellt, die er testet. Auf Stufe 3 schließlich stellt er die richtige Hypothese auf, dass das Konzept $Y = miv =$ [grau und quadratisch] sein muss. Der wichtige Punkt ist nun, dass der Lerner auf Stufe 3 eine Hypothese aufstellt, die er nicht aus den Schritten 1 und 2 abgeleitet haben kann, sondern „dass er genau die Hypothese aufstellen muss, die es gerade zu erwerben gilt! Er kann also das Ziel des Erwerbsprozesses nur erreichen, wenn er über dieses Ziel schon verfügt“ (Löbach 2000: 110). Fodor selbst kommt nun zur entscheidenden Schlussfolgerung, dass Konzepte schon angeboren sein müssen, was einige zu der, m.E. merkwürdigen, Schlussfolgerung verleitet, dass alle Konzepte angeboren seien und in Form einer Konzeptliste a priori gegeben sei:

„[...] human nature gives us the concept [CLIMB] for free [...] This is the way we learn language. We simply learn the label that goes with the preexisting concept. So in other words, it is as if the child, prior to any experience, has a long list of concepts like [CLIMB], and then the child is looking at the world to figure out which sound goes with the concept.“
(Chomsky 1988: 190)

Es kann an dieser Stelle keine ausführliche Diskussion des Arguments erfolgen (für eine solche, s. Löbach 2000), es sollen aber zwei Punkte angeführt werden. Der erste ist primär kognitionswissenschaftlicher, der zweite mehr methodologischer Natur.

Fodor bedient sich eines logischen Argumentationsverfahrens, um die generelle Unmöglichkeit des Konzepterwerbs zu zeigen. Was er als Bestandteile in die logische Analyse einführt, kann nichts weiter sein, als mehr oder weniger wahrscheinliche atomare Modelle für in Echtzeit ablaufende Prozesse. Das atomare Merkmal [grau] ist unter letzterer Perspektive eine semantische Repräsentation, die für nichts anderes steht, als dass ein Kategorisierungsprozess stattfindet, der für bestimmte Sinnesdata funktional spezifiziert ist und aus diesen eine Einheit konstruiert.⁵⁰ Bierwisch (1970: 181-182) diese Anschauung in bestechend detaillierter Art und Weise zusammengefasst:

⁵⁰ Was in diesem Sinne nichts anderes ist, als die kategoriale Wahrnehmung innerhalb der Phonologie, so dass ein phonetisch messbares Lautspektrum als ein Konsonant- oder Vokal wahrgenommen wird. Diese kategoriale Identifikation und auch Speicherung kann man sich durch die Vermittlung (angeborener) phonologischer Merkmale vorstellen (vgl. Stevens 2002).

„It seems natural to assume that these [semantic] components represent categories or principles according to which real and fictitious, perceived and imagined situations and objects are structured and classified. The semantic features do not represent, however, external physical properties, but rather the psychological conditions according to which human beings process their physical and social environment. Thus they are not symbols for physical properties and relations outside the human organism, but rather for the internal mechanisms by means of which such phenomena are perceived and conceptualized. This then leads to the extremely far-reaching, though plausible, hypothesis that all semantic structures might finally be reduced to components representing the basic dispositions of the cognitive and perceptual structure of the human organism. According to this hypothesis, semantic features cannot be different from language to language, but are rather part of the general human capacity for language, forming a universal inventory used in particular ways by individual languages.“

Gerade aber das sprachliche Arbeitsgedächtnis ist funktional fähig, solche kategorial gebildeten Repräsentationen verschiedener Domänen zu einer neuen Repräsentation zu verschmelzen, wie z.B. Shusterman & Spelke (2005) zeigen. Im Sinne ihrer *Linguistic Combination Hypothesis* nutzten erst ältere Kinder Informationen aus verschiedenen kognitiven Modulen zur allgemeinen Orientierung.⁵¹ Im Besonderen waren es in dieser Untersuchung Inhalte aus dem Modul, das für die Verarbeitung geometrischer Informationen relevant ist, und dem Modul, das Objekteigenschaften verarbeitet (wie z.B. auch Farbe). Die Fähigkeit der Kinder Phrasen zu produzieren, wie *left/right of X*, korrelierte im Gegensatz zu anderen Aspekten der kognitiven Entwicklung (wie z.B. IQ, räumliches Arbeitsgedächtnis etc.) stark mit einer Orientierungsfähigkeit, die eine Repräsentation der Form *links von der roten Wand* voraussetzte (zuvor konnte entweder nur die Orientierung an geometrischer oder Objekteigenschaftsinformation festgestellt werden). Nach Shusterman & Spelke (2005: 106) folgt daraus,

„that the specific role of language is to allow the isolated contents of encapsulated representations to combine into unified representations.“

Carruthers greift diese Einsichten auf und integriert sie in seine eigene Konzeption von (massiver) Modularität, wobei er von sogenannten Aktivitätszyklen eines modular strukturierten Geistes ausgeht:

„Cycles of activity would thus become possible, as follows. In response to perceptual or linguistic input, the central modules generate a variety of domain-specific outputs. These are made available to the language faculty, which combines some of them into a sentence that is displayed in imagination, processed by the comprehension subsystem, and made available to the central modules once again. The latter process the resulting input, generating new domain-specific output, which is again made available to the production subsystem of the language faculty, which formulates some of it into a new sentence; and so on. While there is no reason to think that this could be the *whole* of human thinking, it does suggest a way in which—given

⁵¹ Kinder im Alter von 18-24 Monaten und Ratten orientierten sich hingegen nur an Merkmalen aus einem Modul.

sufficient cycles of domain-specific activity — new non-domain-specific ideas and beliefs might be generated, which could go well beyond anything manifest in the initial input.“ (Carruthers 2005: 83)

Die von Carruthers benannten Aktivitätszyklen machen somit einen Aspekt deutlich, der in der Modularitätskonzeption von Fodor nicht vorhanden ist, nämlich:

1. Den Zugriff des Sprachmoduls auf die domänenspezifische Inhalte anderer Module
2. Die damit zusammenhängende Möglichkeit der Kombination unterschiedlicher semantischer Gehalte zu neuen Repräsentationen
3. Die Möglichkeit der sukzessiven Einspeicherung und Erweiterung von Repräsentationen durch eine Input-Output-Schleife (der repräsentationale Output der Sprachproduktion kann wiederum der repräsentationale Input des eigenen Sprachverstehens sein)

Unter dieser Perspektive können wir in Bezug auf Fodors Argumentation also feststellen, dass er sowohl recht hat als auch falsch liegt. Piatelli-Palmarini (1994: 331) gibt zur Übersicht noch einmal Fodors Argument in Kürze wieder:

„Fodor's Argument: If this is the case, that is, if the language at stage 3 is really more powerful than the language at stage 1 and 2, then this transition cannot be the result of learning, it cannot come from induction.“

Es ist wohl deutlich geworden, dass Fodors Argument nur mittels einer postulierten Identität von LERNEN = INDUKTION schlüssig ist: Wenn eine neue Repräsentation gebildet wird, die nicht aus den vorherigen Zuständen erklärt werden kann, dann findet keine Induktion und somit kein Lernen statt. Dass kein Lernen stattfindet, ist allerdings eine Schlussfolgerung, die nur gilt, wenn man Lernen hier nur als induktive Verallgemeinerung versteht. Im Sinne von Carruthers Konzeption des mentalen Apparates findet aber sehr wohl Lernen im Sinne von Konzepterweiterung statt, dieses gerade deswegen, weil es das Sprachmodul selbst ermöglicht, auf domänenspezifische Inhalt zuzugreifen, diverse Repräsentationsformate zu kombinieren und diese neu gebildeten Repräsentationen erneut zu speichern usw. Lernen in Bezug auf Konzepterweiterungen kann somit als ein mentaler und genuin generativer Prozess verstanden werden und nichts zwingt unter dieser Perspektive, die Angeborenheit aller Konzepte anzunehmen.⁵² Diese Schlussfolgerung gilt dementsprechend nur, wenn man mit

⁵² Weiterhin entspricht diese Sichtweise auch dem Kriterium der downward adequacy (vgl. Kap. 1.2.2). Auch wenn Carruthers kognitive Modellierung gleichsam modular ist, so ist die Kombinatorik domänenspezifischer Inhalte in lexikalischen Elementen kompatibel mit der neuronalen Verarbeitung und Speicherung von lexikalischen Wörtern, die großflächig distribuiert und überlappend über verschiedene Hirnareale geschieht (vgl. Pulvermüller 2012).

Fodor Lernen als logische Prozedur der Induktion konzipiert und weiter eine Unterscheidung zwischen Lernen (=Induktion) und Angeborenheit postuliert.⁵³

Wenn wir somit annehmen, dass dem Spracherwerb eine kognitive Echtzeitverarbeitung unterliegt, so liegen wissenschaftshistorische Analogien in Bezug auf logisch-apriorische Argumentationen nahe. Einstein nannte es „eine der verderblichsten Taten der Philosophen, dass sie gewisse begriffliche Grundlagen der Naturwissenschaft aus dem der Kontrolle zugänglichen Gebiete des Empirisch-Zweckmäßigen in die unangreifbare Höhe des Denknötwendigen (Apriorischen) versetzt haben.“ (Einstein 1969: 6). Wenn man den gallileischen Stil (vgl. Kapitel 1.1) in der Untersuchung sprachlicher Repräsentationen durchhalten will, ist eine Deduktion im Sinne Fodors, auch wenn sie logisch stringent sein mag, nicht hinreichend für eine Verwerfung von empirisch überprüfaren Modellierungen.⁵⁴ In der Tat sprechen empirische Beobachtungen für eine Modellierung, die auf einer merkmalsbasierten Analyse fußt. Schon in den Arbeiten von Stachowiak (1979) und Poeck et al. (1974) konnten bei verschiedenen Aphasietypen Regularitäten in den semantischen Paraphrasen nachgewiesen werden, die sich mit Hilfe eines merkmalsbasierten Ansatzes beschreiben lassen:

Zielwort	Paraphasie
Bagger	Raupe, Bulldozer, Trecker, Kran, Kranwagen [PHYSIKALISCHES OBJEKT, NICHT BELEBT, KÜNSTLICH, FAHRZEUG, WERKZEUG]
Armbrust	Büchse, Harpune, Bogen, Flitzebogen [PHYSIKALISCHES OBJEKT, NICHT BELEBT, KÜNSTLICH, SCHIESSGERÄT]
Eimer	Schüssel, Topf, Glas [PHYSIKALISCHES OBJEKT, NICHT BELEBT, KÜNSTLICH, BEHÄLTER]

Tabelle 20: Semantische Paraphrasen und Merkmalsanalyse (nach Peuser 2000: 120).

⁵³ Dieser Dichotomie unterliegt die noch viel wesentlichere zwischen Empirismus und Rationalismus, die Fodor vertritt. Wie allerdings Löbach (2000: 117-118) in Rekurs auf Engfer (1996) betont, handelt es sich bei jeder Position „aber lediglich um den Entwurf einer idealtypischen empiristischen bzw. rationalistischen Position, d.h. eine Abstraktion und Idealisierung über tatsächliche Erkenntnistheorien, die in dieser „reinen“ Form nie formuliert wurden.“

⁵⁴ Kant hat bekanntlich in seiner Kritik der reinen Vernunft Raum und Zeit als apriorische Formen der sinnlichen Anschauung konzipiert. Der Weg, auf dem er zu diesem Verständnis kam, war die Einteilung von uns möglichen Urteilen nach ihrer logischen Zusammengehörigkeit und die Erkenntnis, dass synthetische Urteile a priori nur möglich seien, weil Raum und Zeit als Bedingungen der Möglichkeit von Wahrnehmung als die Form unserer Sinnlichkeit gegeben sind. Seine Deduktion kann man auch heute noch als vollkommen schlüssig anerkennen, wenn man die empirische Bestätigung der Allgemeinen Relativitätstheorie im Jahre 1919 negiert.

Die Verwendung von semantischen Merkmalen hat sich in der Neurolinguistik wie auch Neuropsychologie zum Teil etabliert (vgl. Moss et al. 2007) und dort zu fruchtbaren Erkenntnissen geführt. Z.B. konnten Caramazza & Shelton (1998) eine kategorienspezifische Dissoziation der semantischen Belebt-Unbelebt-Unterscheidung nachweisen, die unabhängig von der jeweiligen Modalität bestand. Eine Speicherung und ein entsprechender Abruf solcher semantischer Informationen in und aus merkmalspezifisierten Domänen werden in dem aktuellen Conceptual-Structure-Ansatz (z.B. Tyler & Moss 2001) angenommen.

Im Spracherwerb war es vor allem Clark (1973), die einen merkmalsbasierten Erwerb von Wortbedeutungen annahm, allerdings wurde dieser Ansatz nicht mehr weiterverfolgt (s.o.) und m.W. nicht auf sehr frühe Wortproduktionen angewandt. In Ingram (1971) findet sich einer der wenigen Ansätze, der mit der zugrundeliegenden Annahme der potentiellen Satzwertigkeit von frühen Äußerungen und dem sukzessiven Merkmalsausbau innerhalb eines transitiven Formats sowohl oben genannter konzeptueller Elaboriertheit als auch der „semantischen Weite“ sehr früher Äußerungen gerecht wird.

3.2.4 Ingrams Analysen: Erste Annäherungen an die strukturelle Elaboriertheit früher Wortbedeutungen

Die semantischen Merkmalsmengen zur Klassifikation eines Objektes werden nach Ingram innerhalb eines semantischen Gesamtkonzeptes gespeichert. Dieses ist in Anlehnung an Fillmore (1968) eine Satzrepräsentation, die sich wiederum aus den Repräsentationsebenen Modalität und Proposition zusammensetzt. Modalität bezieht sich hier auf Informationen wie demonstrativ oder interrogativ wie auch auf die Ausdrucksart (wie z.B. Zeigen) und modifiziert die gesamte Proposition. Letztere setzt sich bei intransitiven Repräsentationen aus der Prädikation eines Zustandes oder einer Funktion zu einem Objekt zusammen (s. Abb. 47), bei transitiven Repräsentationen zusätzlich aus Agens und Akt.

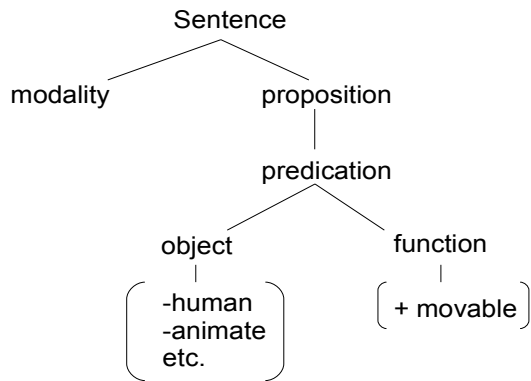


Abbildung 47: Beispielrepräsentation für *Stein* nach Ingram (1971: 890).

Einige Beispiele aus Ingrams Analyse früher Wörter aus Pollock (1878) und Leopold (1949) sollen die Analyse weiter verdeutlichen. Das Kind von Pollock benutzte im Alter von 1;1 die Form *wa-wa* für Wasser oder Trinken. Allerdings verwendete es diese Form auch im volitionalen Sinne (*Ich möchte Wasser trinken*), so dass beide Repräsentationen in Abbildung 48 denkbar wären:

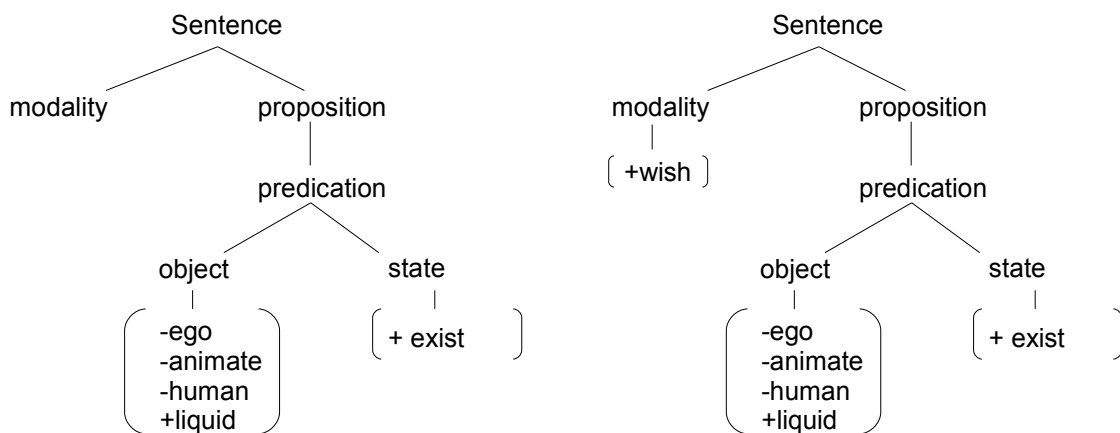


Abbildung 48: Repräsentationen von *wa-wa* nach Ingram (1971: 896).

Das Merkmal [+ego] definiert hier, ob sich das Wort auf das Kind selbst oder andere Objekte und Personen bezieht ([-ego]). Falls der Kontrast noch nicht etabliert ist, kann es auch unspezifisch [±ego] verwendet werden, wie z.B. vom selben Kind im Alter von 1;0 mit dem Wort *ba-ba* als generelles Demonstrativ (Ingram 1971: 895). Das Merkmal [±exist] definiert hingegen den Zustand des Da- oder Fortseins. So nutzte z.B. das Kind von Preyer (1882) die Wörter *atta* (= [-exist]) und *da* (= [+exist]), um den relevanten Kontrast auszudrücken.

Abbildung 49 zeigt die Repräsentation von *bebe*, ein Wort, mit dem Pollocks Kind im Alter von 1;3 auf sich selbst referierte und somit die Merkmale [+ego; +exist] gegeben sind.

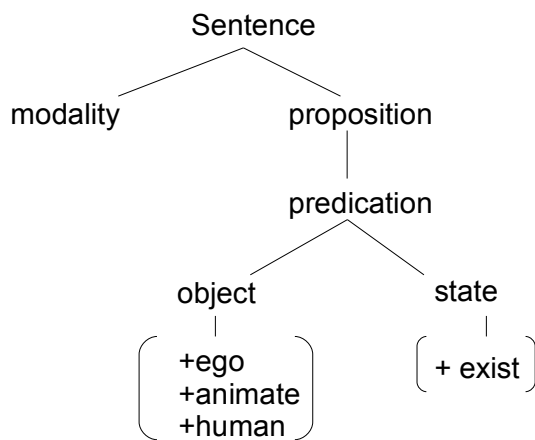


Abbildung 49: Repräsentation von *be-be* nach Ingram (1971: 897).

Volle Transitivität ist erst vorhanden, wenn die Äußerungen eine gesamte verbale Information enkodieren. So z.B. in Hildegards Äußerung *up* im Alter von 1;6, was soviel wie ein Wunsch oder eine Aufforderung der Form *Get up* war:

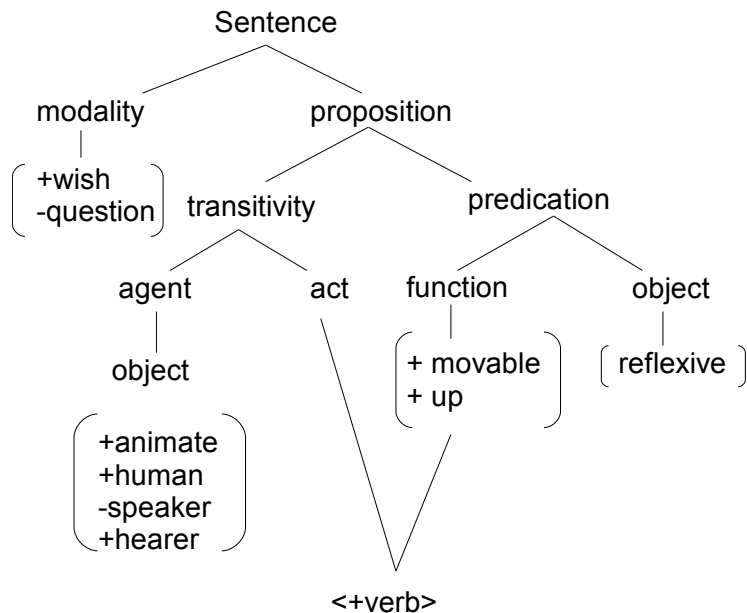


Abbildung 50: Repräsentation von *up* nach Ingram (1971: 906).

Die Merkmale [\pm hearer] und [\pm speaker] sind nach Ingram (ebd.: 893) keine semantischen Merkmale im eigentlichen Sinne, sondern werden als deiktische Merkmale gebraucht. Sie

dienen der kommunikativen Differenzierung. Während [\pm ego] nur ICH von NICHT-ICH unterscheidet und jüngere Kinder somit auch unabhängig von einem Konversationskontext mit sich, Tieren oder Puppen sprechen, führen die obigen Merkmale diese Unterscheidungsfähigkeit ein. Hildegard verwendete in der zweiten Hälfte des zweiten Lebensjahres zunehmend Imperative, was für Ingram eine solche Merkmalsrepräsentation plausibel macht (ebd.: 904).

Das hervorstechendste Merkmal dieser Bedeutungsanalyse für Wörter ist ihre holophrastische Grundstruktur. Es mag nicht für jeden einsichtig sein, dass frühe Wortbedeutungen vollständige Propositionen sein sollen. Sind Kinder in einem solch frühen Alter überhaupt fähig, ((in-)transitive) Relationen zu repräsentieren? Untersuchungen der letzten Jahre lassen den Schluss zu, dass schon Kinder im ersten Lebensjahr komplexe Repräsentationen von Handlungen im Sinne von Verursachung, Mitteln und intentionalen Zielen haben. So konnte z.B. in den Untersuchungen von Woodward (1998) gezeigt werden, dass schon sechs Monate alte Säuglinge signifikante Unterschiede einer Greifhandlung erkennen können. Nach einer Habituationsphase mit einer menschlichen Hand, die nach einem Teddy neben einem Ball griff, erfolgte eine Dishabituationsphase nur beim Wechsel des ergriffenen Objekts bei gleichbleibendem Bewegungsmuster. Griff eine mechanische Klaue nach dem Teddy oder vollzog die menschliche Hand eine andere Handlung (sie berührt das Objekt nur mit dem Handrücken), erfolgte keine Dishabituationsphase. D.h., die Säuglinge konnten die Relation zwischen Person und Zielobjekten im Sinne einer intentionalen Handlung repräsentieren, eventuell auch ein einfaches Konzept der Handlung *Greifen* (vgl. auch Henning et al. 2009). Es sei noch erwähnt, dass Säuglinge dieses Alters auch unvertraute menschliche Handlungen wie auch die Bewegungen einer mechanischen Klaue als auf ein Zielobjekt gerichtet wahrnehmen können. Hier muss allerdings ein wahrnehmbarer Handlungseffekt hinzutreten, wie z.B. das Verschieben des Objekts (vgl. Jovanovic et al. 2007). D.h., die Wahrnehmbarkeit von Effekten ist wesentlich für die kognitive Strukturierung einer Handlungssequenz und ist wohl auch ein Kernwissen, was nicht kleinteilig durch Erfahrungen erlernt werden muss:

„The fact that when the cues of goal-directedness are present goal attribution can be demonstrated as early as 6 months of age for unfamiliar human actions as well as for actions performed by novel objects lacking human features suggests that understanding goal-directedness does not develop in a piecemeal fashion as a function of learning about particular actions as they become familiar to the infant, but is due to a more general representational understanding of the concept of goal that indicates the very early presence of a teleological action interpretational system in infancy.“ (Király et al. 2003: 767)

Dass der Effekt im Versuch perzeptuell hervorgehoben werden muss, zeigt, dass die Objekte und ihre jeweiligen Merkmale für Säuglinge erst einmal wahrnehmbarer sind als deren Relationen untereinander. Mit dieser Präferenz geht einher, dass Säuglinge Relationen anfangs nur in spezifischen Kontexten kategorisieren können. So können bereits drei Monate alte Säuglinge die räumlichen Relationen *über* und *unter* unterscheiden, wenn diese Relationen mit familiarisierten, d.h. ihnen vertrauten Objekten angezeigt werden, während sechs Monate alte Säuglinge dies mit nicht-vertrauten Objekten vermögen (Quinn 2005). Eine solche Kategorisierungsentwicklung von spezifischen hin zu abstrakten Kategorisierungen findet bei Verwendung verschiedener Methoden, in Darstellung von statischen wie auch dynamischen Ereignissen wie auch im Gebrauch von formmäßig simplen und komplexen Objekten statt (vgl. Casasola 2008: 22). Auch wenn die räumlichen Konzepte von Säuglingen und Kleinkindern also erst einmal an spezifische, ihnen vertraute Objekte gebunden sind, so sind nichtsdestotrotz grundsätzlich dazu fähig, abstrakte Relationen zwischen Objekten zu repräsentieren.

Erscheint Ingrams Ansatz für eine Strukturanalyse mit der Postulation eines propositionalen Gehalts durchaus plausibel, so zeigen sich doch bei genauerer Betrachtung Schwierigkeiten mit der Merkmalsanalyse. Für die Merkmale $[\pm\text{ego}]$, $[\pm\text{exist}]$, $[\pm\text{hearer}]$ und $[\pm\text{speaker}]$ scheint es doch sehr zweifelhaft, ob sie zu einem universellen Set von Merkmalen der Wortbedeutung gehören könnten. Insgesamt wird die Merkmalsanalyse von Ingram natürlich von der wesentlichen Kritik von Fodor et al. (1980) getroffen, nämlich, dass es unmöglich ist, die Bedeutung von Wörtern nach Maßgabe anderer Wortbedeutungen zu definieren und so irgendwann zu einer semantisch nicht mehr reduzierbaren Basis zu gelangen. Wie allerdings Jackendoff (2002: 335-336) betont, ist das Argument richtig, aber nur wenn von einer spezifischen Variante der Merkmalsdefinition ausgegangen wird:

„However, his notion of definition is the standard dictionary sort: a phrase that elucidates a word meaning. So what he has actually shown is that word meanings cannot be built by combining other word meanings, using the principles that also combine words into phrases. But there are other options. [...] So here is the answer to Fodor's argument against definitions: although he has shown that lexical meanings cannot be decomposed definitionally, he neglects the possibility that there are non-definitional forms of decomposition of the sort found in phonology and physics. Since his alternative to non-definitional decomposition is genetic transmission of the whole meanings of *quark* and *fax*, our choice ought to be clear.“

Es lohnt sich m.E. also an dieser Stelle, die Analysen von Ingram nicht komplett zu verwerfen, sondern sie im Hinblick auf Protowörter unter einer erweiterten Perspektive

anderer Bedeutungstheorien und Erwerbsdaten zu betrachten. Eine solche Perspektive soll im anschließenden Kapitel erarbeitet werden.

3.2.5 Die konzeptuelle Struktur von Protowörtern und ihre Integration in die Parallele Architektur

Vihman (2014: 146) sieht die Vagheit von Protowortbedeutungen darin begründet, dass die Bedeutungen zu großen Teilen ein Ausdruck von Affekten seien.⁵⁵ Menn (2013: 172) weist auf die fehlende Symbolqualität, d.h. die beobachtbare Kontextgebundenheit in der Verwendung, hin. Weiter stellt sie heraus, dass Protowörter in der Produktion an andere gerichtet sein können, durchaus aber auch unabhängig von einer konversationellen Intention gebraucht werden (ebd: 171). Die Bedeutung ergibt sich hier also mehr aus dem „Was man sagt, wenn man X tut“, wobei die diskrete Kombination aus Wortproduktion und Situation des Tuns wiederkehrend und damit spezifizierbar ist.

Wenn man diese Angaben von Vihman und Menn zusammenfasst, kann man folgende Merkmale der Protowortbedeutung festhalten:

In einem gegebenen, wiederkehrenden Kontext einer Handlung werden Protowörter sowohl konversationell als auch rein personal gebraucht, wobei die Bedeutung vor allem innere (affektive) Zustände anzeigt, also primär subjektiver Natur sei.

Wie kann man eine solche Zusammenfassung nun in ein theoretisch explizites Modell transferieren? Es ist für eine adäquate Modellierung sinnvoll, noch einmal aufzulisten, was genau ein Kind an kognitiven Prozeduren durchlaufen muss, um die beschriebene Bedeutungszuweisung vorzunehmen:

1. Das Kind muss für eine Identifikation spezifischer wiederkehrender Situationen/Handlungen/Zustände Repräsentationen derselben unterhalten können, die sich der kontextuellen Bedingtheit entsprechend aus dem repräsentationalen Input verschiedener Modalitäten aufbauen müssen.
2. Die produzierte phonologische Form ist dergestalt mit dem minimalen konzeptuellen Gehalt assoziiert, dass eine Situation, eine Handlung oder ein Zustand identifiziert wurde.
3. Die Produktion der spezifischen Form-Gehalt-Korrespondenz wird begleitet von einer affektiven Akzentuierung bzw. zweiten Bedeutung, die entsprechend mit der spezifischen Form-Gehalt-Korrespondenz assoziiert ist. Die Produktion tritt in einem konversationellen als auch personalen Modus auf.

⁵⁵ „Most children also develop sound–meaning links of their own (‘protowords’), producing more or less stable sound patterns in conjunction with relatively broad or global meanings expressive of their personal needs and interests and other affects.“ (Vihman 2014: 146)

Zu 1.: Es wurde in Kapitel 3.2.4 schon darauf hingewiesen, dass Säuglinge und Kleinkinder fähig sind, Relationen zwischen Objekten und Akteuren in Ereignissen zu repräsentieren. Carter (1975) fasst frühe Wörter als sogenannte sensomotorische Morpheme, welche zusammen mit spezifischen Gesten produziert werden und im Erwerb die spätere Grundlage zielsprachlicher Wörter bilden. In Abbildung 51 ist ein solcher Erwerb dargestellt. Im Alter von ca. 12 Monaten produziert das Kind David diverse m-initiale Formen, die den Wunsch nach einem Objekt ausdrücken, das allerdings selbst noch lexikalisch un spezifiziert ist. Im Alter von ca. 1;3 Jahren benennt David die spezifischen Objekte, wobei er die m-initialen Formen nun phonologisch in zwei Varianten differenziert. Erste Variante bezeichnet den Wunsch nach der Wiederholung einer Handlung, zweite Variante steht für den Wunsch nach dem Besitz eines Objektes. Im Alter von 1;10.10 werden schließlich zieladäquat das Wort „more“ für die Wiederholung einer Handlung, „my, mine“ für die Besitzanzeige und schließlich auch der Name des spezifischen Objekt produziert, ohne noch in einer Relation zueinander stehen zu müssen.

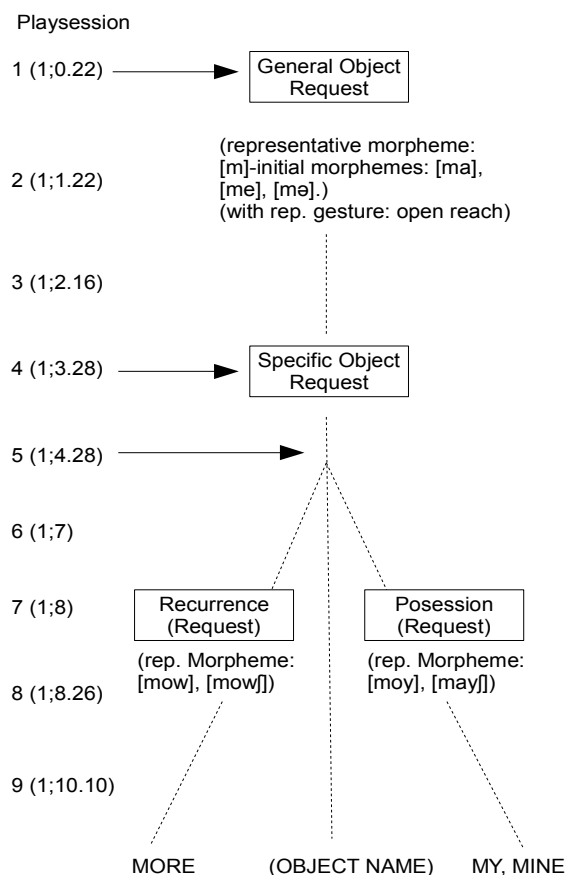


Abbildung 51: Veränderung von Wortbedeutungen nach Carter (1975).

Ohne die Terminologie von Carter übernehmen zu wollen, so zeigt das Beispiel doch Entscheidendes. In der Perzeption wird konzeptuelles Wissen vom Spezifischen zum Abstrakten aufgebaut und in der Produktion bilden die Objektrelationen bzw. die eigene Relation zu den Objekten den konzeptuellen Gehalt der ersten Wörter. Spezifisch kann somit als kontextuell gebundene Bedeutung, abstrakt oder allgemein als dekontextualisierte Bedeutung verstanden werden. Eine dekontextualisierte und in einem solchen Sinne symbolische Benennung spezifischer Objekte erfolgt erst, wenn die sogenannte Benenneinsicht stattgefunden hat, nach der alle Objekte einen Namen zugewiesen bekommen können (Goldfield & Reznick 1990). Ungeachtet der späteren Benenneinsicht und Expansion des Wortschatzes zeigen die Daten von Carter (vgl. Tabelle 21), dass auch schon vor Zielwortproduktionen Bedeutungs-differenzierungen (Bedeutungen sind in der Tabelle als *Schema* klassifiziert) über Laute vorgenommen werden, genauer hauptsächlich über die initialen Laute.

Schema	Gesture	Sound	Goal	No. of instances (1st 4 play sessions)
(1) Request Object	reach to object	[m]-initial	Get receiver's help to obtaining object	342
(2) Attention to Object	point, hold out	Alveolar-dental ([l] oder [d])-initial ([y]-initial – a few instances only)	Draw receiver's attention to object	334
(3) Attention to Self	sound of vocalization	Phonetic variants of <i>David, Mommy</i>	Draw receiver's attention to self	142
(4) Request Transfer	reach to person	[h]-initial (constricted & minimally aspirated)	Obtain object from, or give to, receiver	135
(5) Dislike	prolonged, falling intonation	Nasalized, especially [n]-initial	Get receiver's help in changing situation	82
(6) Disappearance	waving hands, slapping	[b]-initial	Get receiver's help in removing object	32
(7) Rejection	negative headshake	[ʔəʔə]	Same as for Dislike (above)	20
(8) Pleasure - surprise-Recognition	(smile)	flowing or breathy [h] sounds, especially <i>hi, ha, oh, ah</i>	Expresses pleasure	20

Tabelle 21: Übersicht über Davids produzierte sensomotorische Morpheme und ihrer kommunikativen Verwendung im Alter von 1;0 bis 1;4 (Foster 1990: 39).

Zu 2:

Barrett (1986) hebt noch stärker auf die Bindung des konzeptuellen Gehalts erster Wörter an definierte Situationen mit spezifischen Objekten ab. So beschreibt er, wie sein Sohn Adam das Wort *duck* verwendete:

„This word was initially produced by Adam only while he was engaged in the process of hitting one of his toy yellow ducks off the edge of the bath (which is where they were normally kept). He was never observed producing this word in any other situation at this initial stage: he never produced it while he was playing with his toy ducks in other situations, or while he was looking at or feeding real ducks. The behavior therefore tends to suggest once again that Adam had not yet learned that the word duck could be used to refer to either his toy ducks or real ducks. Instead, his behavior suggests that he had simply identified one particular event in the context of which it was appropriate for him to produce the word duck.“ (ebd.: 40)

Solch eine Verwurzelung in konkreten Situationen oder Ereignissen konnte Barrett für alle ersten Wörter ausmachen, unabhängig davon, ob sie phonologisch zielwortbezogen waren (wie *duck*, das zuerst als *du* [dat] realisiert wurde) oder nicht (wie z.B. die Wörter [hɪja:] von Kind Tina, das in Situationen verwendet wurde, in denen sie ein Objekt einer anderen Person gab, oder [kju:], das in Situationen verwendet wurde, in denen sie von einer anderen Person ein Objekt genommen hat⁵⁶ (vgl. Barrett 1981)). Insofern kommt Barrett zu der Schlussfolgerung:

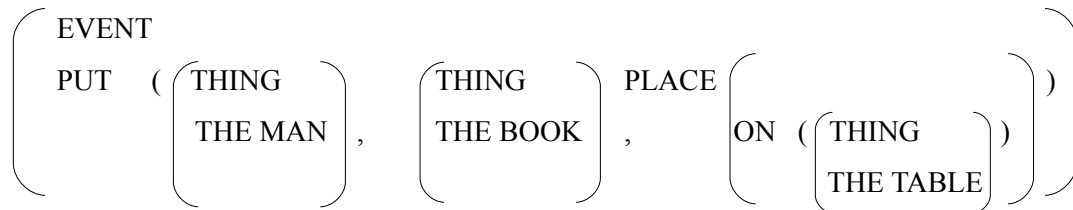
„[...] These words can often be interpreted merely as ritualized responses to the occurrence of particular events (in which the child may be involved either as participant or as a spectator), events which the child interprets, in a holistic manner, as providing appropriate contexts for the production of those words.“ (Barrett 1986: 42)

Auf Basis dieser Beobachtungen geht Barrett davon aus, dass frühe Wortbedeutungen aus Event-Repräsentationen bestehen, in denen die spezifischen Relationen, die wiederum spezifischer Teilnehmer enthalten, encodiert sind. Für das Wort *duck* entspricht dies in den ersten Verwendungen folgender Repräsentation, wobei mit den Zahlenindizes auf die drei unterschiedlichen Spielzeugenten von Adam referiert werden (vgl. ebd.: 43):

(Adam
hitting
toy yellow duck (duck₁, duck₂, duck₃))

⁵⁶ Interessanterweise fanden bei diesen nicht-zielwortbezogenen Wörtern keine weitere Generalisierungen statt, d.h., sie wurden invariant in der Zeit des frühen Worterwerbs verwendet (vgl. Barrett 1986: 60).

Ähnliche Formen der Event-Repräsentation werden nun von Jackendoff auch zur Analyse des konzeptuellen Gehalts von Verben verwendet. So z.B. für das Verb *put* in dem Satz „The man put the book on the table“ (Jackendoff 1983: 68):



Die konzeptuelle Struktur könnte noch ausführlicher dargestellt werden, dies würde aber über die Zielsetzung dieses Kapitels weit hinausführen (vgl. ebd.: 179-180). Grundsätzlich finden sich in jeder Struktur Relationen beschrieben, die durch eine allgemeine Funktions-Argument-Struktur der Form $[X^F ([Y\dots], [Z\dots])]$ dargestellt werden können. Eine Funktion F bildet hier eine konzeptuelle Konstituente des Typs Y und des Typs Z in eine Konstituente des Typs X ab (vgl. Jackendoff 2002: 11). Für das Verb *put* bzw. dessen konzeptuelle Struktur wäre dies die Funktion CAUSE, die drei konzeptuelle Konstituenten als Ergänzung benötigt.⁵⁷

Analog zu der Mehrdimensionalität phonologischer Ebenen unterscheidet Jackendoff (1990) in der konzeptuellen Komponente zwischen einer Thematischen Ebene (engl.: thematic tier) und einer Handlungsebene (engl.: action tier). Während auf der Thematischen Ebene eine komplexe räumliche Struktur eines Events (bzw. Satzes) auf Basis von Bewegung und Ort, wie oben angedeutet, gebildet wird, stellt die Handlungsebene eine gröbere Repräsentation der AKTEUR-PATIENS-Relationen (eines Satzes), also quasi eine intentionale Struktur, dar. Die Rollen AKTEUR (derjenige, der eine Handlung unternimmt) und PATIENS (derjenige, der von einer Handlung affiziert wird) sind zu einem gewissen Grad unabhängig von den üblichen thematischen Rollen, die syntaktischen Konstituenten zugewiesen werden, wie folgende Beispiele verdeutlichen (vgl. Jackendoff 1990: 126-127):

⁵⁷ Das relativ umfangreiche Set der Funktionen nach Jackendoff (2002: 364):

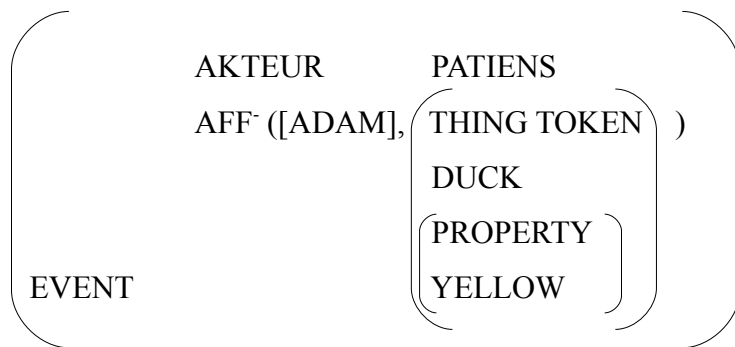
- a. BE: $\langle (X,Y), \text{State} \rangle$, X and Y an ordered pair, where the types of X and Y depend on semantic field
- b. STAY: $\langle (X,Y), \text{Event} \rangle$
- c. GO: $\langle (\text{Object}, \text{Path}), \text{Event} \rangle$
- d. EXT, ORIENT: $\langle (\text{Object}, \text{Path}), \text{State} \rangle$
- e. TO, FROM: $\langle x, \text{Path} \rangle$, where the type of X depends on semantic field
- f. INCH: $\langle \text{State}, \text{Event} \rangle$
- g. PERF: $\langle \text{Event}, \text{State} \rangle$
- h. CAUSE, HELP, LET (three-argument version): $\langle (\text{Object}/\text{Event}, \text{Object}, \text{Event}), \text{Event} \rangle$
- i. CAUSE, LET (two-argument version): $\langle (\text{Object}/\text{Event}, \text{Event}), \text{Event} \rangle$

1.	Sue	hit	Fred.	
	Thema		Ziel	Thematische Ebene
	Akteur		Patiens	Handlungsebene
2.	Pete	threw	the ball.	
	Ursprung		Thema	Thematische Ebene
	Akteur		Patiens	Handlungsebene
3.	Bill	entered	the room.	
	Thema		Ziel	Thematische Ebene
	Akteur			Handlungsebene
4.	Bill	received	a letter.	
	Ziel		Thema	Thematische Ebene
				Handlungsebene

In Jackendoff (2007b: 200-201) wird die einzige Funktion auf der Handlungsebene, *AFF* („affect“) in der Form *X AFF Y* angegeben, was ungefähr die Bedeutung *X affiziert Y* encodiert. Drei mögliche Konfigurationen werden hier angenommen, nämlich das Vorhandensein von AKTEUR und PATIENS oder in zwei Fällen jeweils nur das Vorhandensein einer Rolle. Zudem wird eine Spezifikation *AFF⁺* und *AFF⁻* angenommen, die den Unterschied zwischen negativer (PATIENS) und positiver (BENEFIZIENT) Affizierung markiert.

Nach Wagner & Lakusta (2009) gibt es hingegen Evidenz dafür, dass die Handlungsebene als grobe Repräsentation intentionaler Relationen von Kleinkindern durchaus repräsentiert wird. So wurden in einer Studie von Baldwin et al. (2001) Kleinkindern im Alter von 10-11 Monaten Szenen zur Habituation gezeigt, in der eine zielgerichtete Handlung erfolgt, wie z.B. eine Frau, die ein Handtuch auf dem Boden sieht, sich danach bückt, es erfasst und es dann auf einen Tresen legt. Anschließend wurden die Szenen noch einmal gezeigt, wobei „natürliche“ und „unnatürliche“ Unterbrechungen bzw. Pausen eingebaut waren. Die Pause wurde z.B. eingefügt, wenn die Frau das Handtuch auf den Tresen gelegt hat (natürliche Unterbrechung), das Ziel der Handlung also erreicht war, oder vor dem Ablegen auf dem Tresen (unnatürliche Unterbrechung), das Ziel der Handlung bzw. ihr natürlicher Endpunkt also nicht erreicht war. Die Kinder schauten länger auf die Szenen mit unnatürlichen Unterbrechungen, so dass angenommen werden kann, dass sie zielgerichtete Handlungen mit Ausführungsintentionen und intendierten Endzustände repräsentieren können (für weitere Evidenz vgl. Woodward (1999); Woodward & Somerville (2000)). Wagner & Lakusta (2009) betonen auch, dass die bisherigen Evidenzen für Repräsentationen auf der thematischen Ebene bei Kleinkindern diesen Alters nicht im gleichen Sinne eindeutig sind. Eine minimale Annahme für frühe Wortbedeutungen wäre entsprechend, dass Kinder Repräsentationen auf der Handlungsebene unterhalten können.

Die von Barrett (1986) erstellte Event-Repräsentation des Wortes *duck* könnte gemäß Jackendoff (2007b) also in folgender Form reanalysiert werden:



Diese konzeptuelle Struktur ist über Schnittstellenregeln mit dem phonologischen Wort *duck* assoziiert sowie mit modalitätsspezifischen Repräsentationen (haptische, visuelle usw.) die eine Handlung *schlagen* charakterisieren. In dieser Struktur ist mit dem Merkmal THING TOKEN die Spezifität respektive Individualität der Entität *Ente* markiert. Jackendoff (1983: 78) definiert diese Unterscheidung derart: „We will refer the representation of the thing being categorized as a [TOKEN] concept and that of a category as a [TYPE] concept.“

Tokens sind insofern Entitäten, die konkret wahrgenommen werden, während Types niemals wahrgenommen werden können. Das Spezifische im frühen Spracherwerb ist nun, dass ein [TYPE]-Konzept der individuellen Objekte noch nicht vorhanden ist und erst über die individuellen Tokens gebildet werden muss (was auch die Ergebnisse von Quinn (2005) anzeigen).⁵⁸ Dass Adam nur die spezifischen Spielzeugenten mit *duck* benennen konnte (s.o.), zeigt hier das Fehlen eines generalisierten [Type]-Konzeptes *Ente*, unter das auch andere wahrgenommene Exemplare mit entenspezifischen Merkmalen subsumiert werden könnten.

Bis zu diesem Punkt haben wir also eine (approximale) Strukturanalyse früher Wortbedeutungen im Rahmen eines aktuellen und etablierten Ansatzes erstellt, die mit den Daten experimenteller Forschung und Beobachtungsstudien kompatibel ist. Offen ist allerdings immer noch eine Einordnung der Beobachtung, dass mit Protowörtern eine Kundgabe emotionaler Zustände erfolgt und sie nicht an einen Adressaten gerichtet sein müssen.

⁵⁸ Wie genau die Bildung von TYPE-Konzepten verläuft, muss an dieser Stelle offen bleiben. Jackendoff (1983) bezieht sich in diesem Zusammenhang auf das Lernen durch ostensive Definition, d.h., eine Liste einzelner exemplarischer Tokens bildet letztlich das Type-Konzept.

Zu 3:

Auch für diese letzte Einordnung kann auf die Parallele Architektur rekurriert werden. Es wurde schon dargestellt, dass die primäre Bedeutung früher Wörter eine konzeptuelle Struktur auf der Handlungsebene ist, die intentionale Relationen mittels der Funktion *AFF* abbildet. Folglich ist eine Reduktion der Bedeutungsstruktur auf eine rein emotionale Kundgabe in Anbetracht oben dargestellter Ergebnisse nicht zulässig. Allerdings besteht die Problemstellung weiter, die Systematizität affektiver Akzentuierungen beim Gebrauch früher Wörter erfassen, die eine Form der Bedeutungsnuance darstellt. In Jackendoff (2007a) wird eine Unterscheidung zwischen *content features* und *valuation features* gemacht. *Content features* sind mentale Strukturen, die als funktionales Korrelat des Bewusstseins gelten können, während andere Strukturen unbewusst bleiben. Im Fall sprachlicher Repräsentationen bilden phonologische Strukturen die *content features*⁵⁹, während sie z.B. bei der Wahrnehmung von Musik durch die musikalische Oberfläche bereitgestellt werden.⁶⁰

Syntaktische und konzeptuelle Strukturen sind bei Produktion wie Perzeption nicht bewusst und sind dem Bewusstsein auch nicht zugänglich. Wie Jackendoff (2007b: 82) in Bezug auf das Tip-of-the-tongue-Phänomen betont:

„In other words, meaning without phonology leads to an absence of the qualia that gives experience a form – or perhaps there is a quale of absence or formlessness.“

Es ist also sowohl das funktionale Korrelat des Bewusstseins dem Gewahrsein zugänglich als auch über diese *content features* ein mit ihnen korrelierter, subjektiver Erlebnisgehalt (Qualia). Die sogenannten *valuation features* encodieren, wie eine jeweilige Input- oder Output-Repräsentation einer jeweiligen Modalität (sprachlich, visuell, haptisch, olfaktorisch usw.) erlebt bzw. intern bewertet wird. Binäre Merkmale und ihre Kombinationen definieren ein „Gefühl von“, etwas ist so und so oder eben nicht. Jackendoff nennt ein Set solcher binärer Merkmale (vgl. Jackendoff 2002: 312-313; Jackendoff 2007b: 87-98):

[± external]: Das Gefühl davon, dass etwas als außerhalb erfahren wird oder nicht (wird etwas als Perzept oder als Vorstellung erfahren).

⁵⁹ In Jackendoff (2007b: 80-84) werden verschiedene Argumente dafür angeführt, warum phonologische Strukturen die funktionalen Korrelate von Bewusstsein darstellen, d.h. am ehesten für bewusste Erfahrung zugänglich sind. Ein Argument ist z.B., dass es jedem intuitiv leicht möglich ist, bestimmte Strukturmerkmale der phonologischen Komponente, wie Silbenanzahl oder Betonungsmuster anzugeben, während dies für die syntaktischen und konzeptuellen Strukturen nur anhand einer linguistischen Analyse mediiert durch die phonologische Struktur möglich ist.

⁶⁰ „[...] the musical surface is an "external representation:" it is a two-dimensional structure (of pitch and time) from which information can be derived.“ (Robinson 2009)

[± self-initiated]: Das Gefühl davon, dass etwas selbst hervorgebracht wurde oder nicht (unterscheidet z.B. zwischen intendierten Vorstellungen und Halluzinationen).

[± familiar]: Das Gefühl davon, dass etwas vertraut ist oder nicht (dies bezieht sich nicht nur auf Perzepte, sondern auch auf den Unterschied zwischen Erinnerungen und Imaginationen).

[± affective]: Das Gefühl davon, dass etwas als positiv oder negativ erlebt wird.

[± meaningful]: Das Gefühl davon, dass etwas kohärent und sinnvoll ist oder nicht.

[± committed]: Das Gefühl davon, dass etwas der eigenen Überzeugung entspricht oder nicht.

Die hier aufgelisteten Merkmale bilden in Kombinationen komplexe Einstellungen. Da die konzeptuelle Struktur im frühen Spracherwerb wesentlich auf der Handlungsebene verortet wird, sollen hier nur die sogenannten actional attitudes aufgelistet werden (ebd.: 96):

[+ committed: valence +]: intending

[+ committed: valence -]: avoiding

[- committed]: considering

[- committed; + affective: valence +]: desiring

[- committed; + affective: valence -]: dreading

Das Merkmal [± valence] steht hier für eine kognitiv-kategoriale Repräsentation einer positiven oder negativen Tönung. [+affective: valence+] und [+affective: valence-] stehen demnach für eine positiv-wünschende bzw. negativ-vermeidende Einstellung. Mit dem alleinigen Merkmal [-affective] hingegen, ist die Repräsentation einer rein sachlichen Nuancierung gegeben, wie z.B. in „Der Stuhl steht neben dem Tisch“.

Während also die content features in einer komplexen hierarchischen Struktur angeordnet werden, sind die valuation features binäre Merkmale, die an die Konstituenten der primären konzeptuellen Struktur gebunden sind. In Abbildung 52 ist die hier relevante sprachliche Verarbeitung dargestellt. In der Sprachperzeption würde vermittelt über phonologische und syntaktische Strukturen eine konzeptuelle Struktur 1 erstellt. Parallel zu der Verarbeitung des jeweiligen Inputs ist diese über die phonologische Struktur dem Bewusstsein bzw. der Introspektion zugänglich (Language introspection processor). Zudem findet ein Monitoring statt, wie der Input intern bewertet werden soll (affect monitor) und diese Bewertung wird ebenfalls in die Introspektion (Language introspection processor) eingespeist. Aus diesen Informationen wird eine konzeptuelle Struktur 2 erstellt, die im Langzeitgedächtnis gespeichert wird, somit abgerufen und mitgeteilt werden kann. Auch in der Produktion findet über die genannten Komponenten eine Bewertung des entsprechenden Outputs statt. Diese sekundäre konzeptuelle Struktur ist allerdings eine relativ vage Repräsentation. Diese beiden konzeptuellen Strukturen können, wie Jackendoff (1987) betont, miteinander verschmelzen:

„The two conceptual structures are thereby in registration with each other. When a report of introspection is produced, these two structures in fact become conflated, so that the forms detected are categorized by their meanings. For example, „I see a dog“ reports the detection of a Privileged visual form, but the description of this form is derived from conceptual structure₁, the category in terms of which the form is comprehended.“ (Jackendoff 1987: 316)

Entsprechend können diverse Informationen aus verschiedenen Modalitäten gemeinsam der Introspektion zugänglich sein, und damit stellen solche sekundären konzeptuellen Strukturen mehr die interne Bewertung einer erlebten Szene oder Ereignisses dar als die eines Einzeleindrucks. Das Wiedererkennen eines Ereignisses würde demgemäß sowohl eine Repräsentation im Sinne der primären konzeptuellen Strukturen (konzeptuelle Struktur 1) als auch der sekundären (konzeptuelle Struktur 2) evozieren, was in der Folge auch sprachlich mitgeteilt werden kann. Der entsprechende Output encodiert nach dieser Logik auch die eigene Teilhabe als Akteur oder Beobachter eines Ereignisses mittels der primären und sekundären konzeptuellen Struktur.

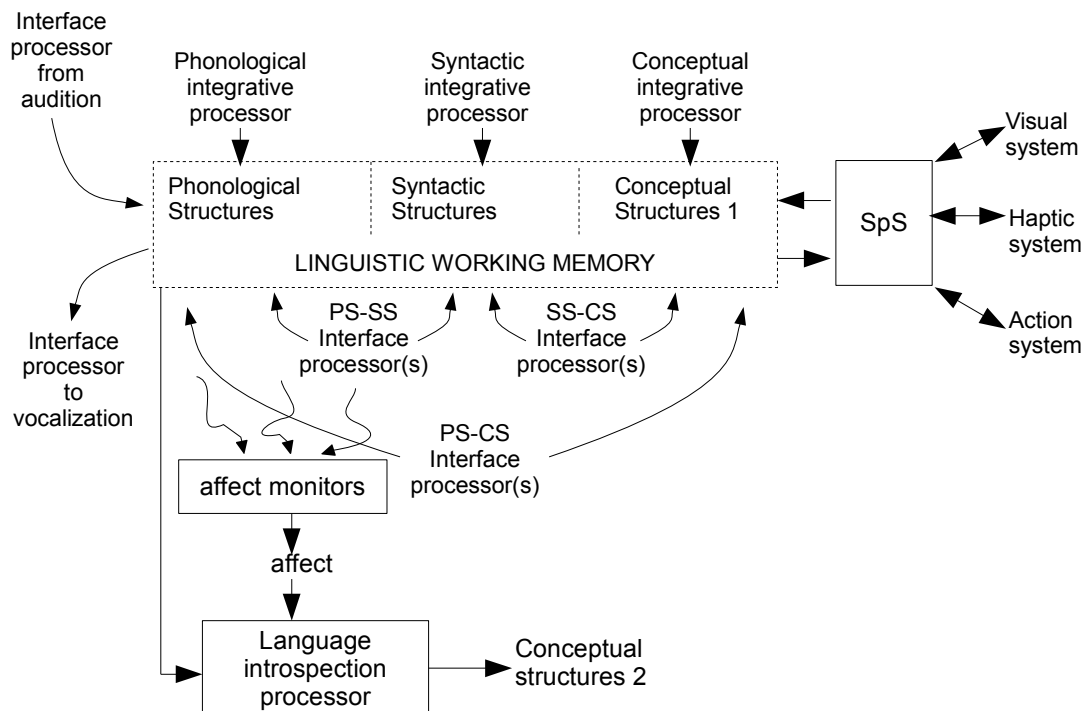


Abbildung 52: Die Parallele Architektur erweitert um die räumliche Struktur (SpS), Modalitäten und interne Registrierungskomponenten.

Was haben wir durch diese Beschreibungen nun gewonnen? Vihman (2014: 158) spricht in Bezug auf Protowörter davon, dass ihre Bedeutungen „are both global, or relatively undifferentiated, and essentially subjective, reflecting the lack of distance between child,

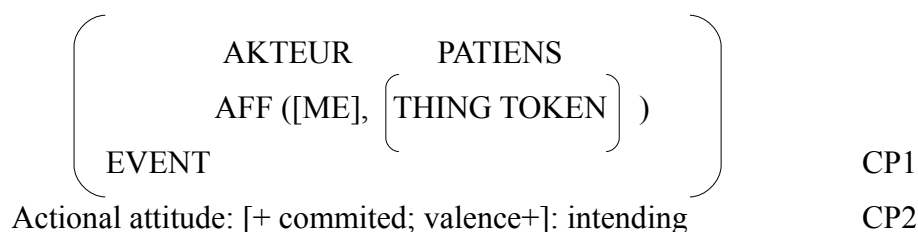
expressive vehicle and expressed content.“ Unter Zugrundelegung der Parallelen Architektur könnte man auch einfach sagen, dass bei kleinen Kindern primäre und sekundäre konzeptuelle Strukturen simultan produziert werden. Es liegt nahe, dass am Anfang der kognitiven Entwicklung das Gewahrsein von einem Wahrgenommenen (sekundäre konzeptuelle Strukturen) noch nicht getrennt von dem etwas, was wahrgenommen wird (primäre konzeptuelle Strukturen), verarbeitet werden kann. Dies deutet auch Jackendoff (1987: 317) an:

„As observed above, these structures tend to be conflated in ordinary circumstances. However, they can be kept distinct, in which case the „detected form“ functions conceptually as a different individual than its „interpretation.“ It can therefore be reported linguistically, producing another Privileged phonological structure that appears in awareness. Then, through the introspection processor, this new phonological form itself is encoded as a distinct individual, the „detection in awareness of a form reporting the detection of awareness,“ in short, „awareness of being aware of something.“ Thus, the possibility of „higher order“ awareness is a straightforward consequence of the present account.“

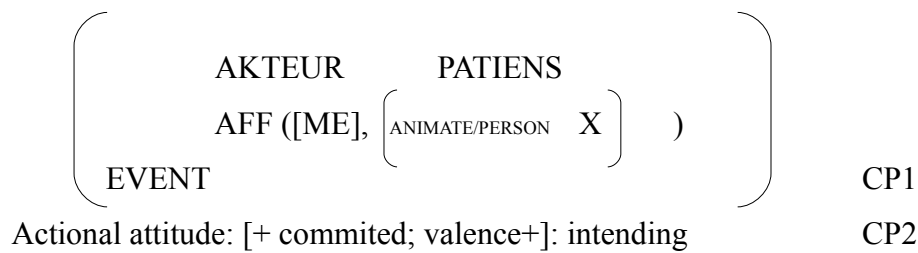
Die Verschmelzung von primären und sekundären konzeptuellen Strukturen in Produktion wie Perzeption kann nach dieser Modellierung interner Prozesse im frühen Spracherwerb als der Normalfall gedeutet werden. Dennoch ist es möglich, anhand obiger Qualifikationen eine Analyse von möglichen Wortbedeutungen vorzunehmen, eine Analyse, die m.E. vielversprechender ist, als globale Charakterisierungen oder einseitige Rekurse auf postulierte pragmatische Akte.

Zieht man noch einmal Hallidays vier grundlegende acts of meaning (vgl. Kap. 3.2.1) heran, so lassen sie sich unter Zugrundelegung obiger Darstellungen der Handlungsebene und der Actional attitudes wie folgt übersetzen:

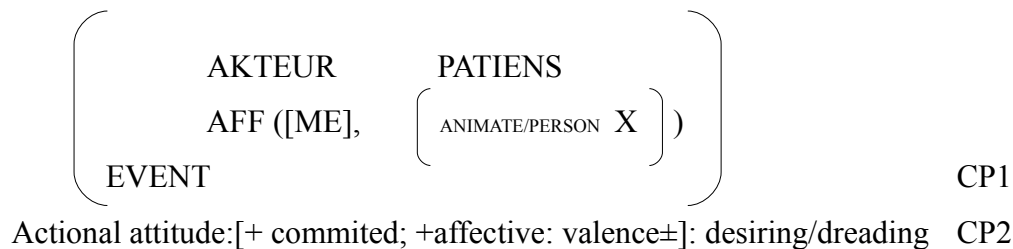
Instrumental: Ziel, ein Objekt zu bekommen oder „Was ich (Akteur) tue, ist ein Objekt (Patiens) zu intendieren (intending)“.



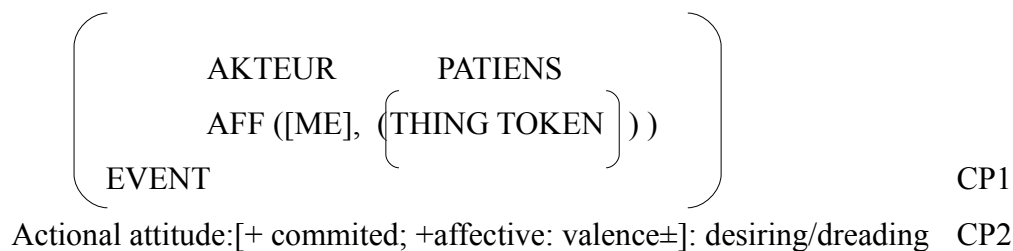
Regulatory: Ziel, eine Person zu einer Handlung aufzufordern, zu manipulieren oder „Was ich (Akteur) tue, ist einer Person (Patiens) meine Intention anzuzeigen (intending)“.



Interactional: Ziel, eine Interaktion mit einer Person herzustellen, Austausch von Aufmerksamkeit oder „Was ich (Akteur) tue, ist einer Person (Patiens) mein Bedürfnis (desiring/dreading) anzuzeigen“.



Personal: Ziel, eigene innere Teilhabe an einer Situation kund zu tun oder „Was ich (Akteur) tue, ist meine eigene Teilhabe (desiring/dreading) an der Situation ((mit einem eventuell gegebenen Objekt (Patiens)) auszudrücken“.



Im Rahmen solch einer Analyse wird innerhalb eines Merkmalsystems, in dem bedeutungsunterscheidende Kontraste gebildet werden, transparent und beschreibbar, warum die acts of meaning spezifisch gebraucht werden und wieso sie sich später durch die Integration weiterer Merkmale und zusätzlicher Ebenen eventuell weiterentwickeln. Die Darstellung der Bedeutungskontraste früher Wortbedeutungen in Tabelle 22 machen deutlich, dass sekundäre konzeptuelle Strukturen wesentlich an der Kontrastbildung beteiligt sind. Erst

wenn die Systemgrenzen zwischen den zwei konzeptuellen Strukturtypen stabil etabliert sind, ist eine rein sprachliche Analyse der primären konzeptuellen Strukturen möglich.

Actional attitude	[+ committed; valence+]: intending	[+ committed; +affective: valence±]: desiring/fearing
Action tier (AFF)		
Event(AFF(ME,THING))	INSTRUMENTAL	PERSONAL
Event(AFF(ME,ANIMATE/PERSON))	REGULATORY	INTERACTIONAL

Tabelle 22: Merkmalsanalyse der acts of meaning.

Ein wesentlicher Unterschied zu den Annahmen von Halliday muss hier betont werden. In allen hier angeführten Bedeutungsanalysen ist eine Repräsentation des Kindes (ME) als Akteur der EVENT-Repräsentation immanent, d.h., das eigene aktive Handeln als Akteur ist selbst ein wesentliches Element der Gesamtrepräsentation. Man kann nach dem Standardtest für die Akteur-Rolle von Jackendoff (2007b: 198) in der Umschreibung „What X did was ...“ auch die Akteur-Rolle des Kindes in den Repräsentationen umschreiben (s.o.):

- Instrumental: Was ich tat, war ein Objekt zu intendieren.
- Regulativ: Was ich tat, war einer Person meine Intention anzuzeigen.
- Interaktional: Was ich tat, war einer Person mein Bedürfnis anzuzeigen.
- Personal: Was ich tat, war meine eigene Teilhabe an der Situation auszudrücken.

Anhand dieser losen Umschreibungen wird klar, dass wir mit der Klassifizierung actional attitude sehr vorsichtig sein müssen. Das Kind bzw. seine Rolle als Akteur ist, wie beschrieben, in die Gesamtrepräsentation eines Ereignisses eingebettet, insofern ist die actional attitude implizit gegeben. Bei Halliday überwiegt allerdings der funktionale Charakter eines Aktes, d.h., das Kind will von jemandem, dass er etwas für es tut oder etwas wahrnimmt. Halliday nimmt also eine explizit intentionale Einstellung des Kindes an. Auch eine solche explizitere Repräsentation könnte man nach Jackendoff (2007b: 258) allgemein formalisieren:

[Situation, -Eventive [Animate/Person X] ATTITUDE [Situation, +Action Y]]

Etwas grob übersetzt, würde eine solche Formalisierung bedeuten: Das Kind [Animate/Person X] befindet sich in einem Zustand [Situation, -Eventive...] einer Einstellung (ATTITUDE), in dem es von jemandem (Y) eine Handlung [Situation, +Action...] intendiert.

Es sollte hierbei bedacht werden, dass eine solch explizite intentionale Haltung der Charakterisierung von Protowörtern in der Form „Was man sagt, wenn man X tut“ nicht

entspricht, sondern eher „Ich will/entscheide/intendiere [...], dass X geschieht“. Es müsste insofern in dem Fall der Protowortproduktion eine Repräsentation des SELBST des Kindes, seiner Einstellung wie auch des intendierten Ereignisses gegeben sein. Eine solche Repräsentation geht in ihrer Komplexität weit über alle bisher dargestellten Charakterisierungen des Wortgebrauchs und allgemein kognitiver Leistungen in dieser frühen Wortbildungsphase hinaus. Aus diesem Grund sollen die oben dargestellten Repräsentationen, in denen eine konzeptuelle Konstituente des Kindes als Akteur implizit in eine ganze Event-Repräsentation eingebettet ist, im weiteren Verlauf der Arbeit als zugrundeliegend betrachtet werden. Die Klassifikationen INTERAKTIONAL, REGULATIV, PERSONAL und INSTRUMENTAL, die im Untersuchungsteil dieser Arbeit (Kapitel 4) verwendet werden, sollen entsprechend nur noch als eine Etikettierung der oben ausführlich ausgearbeiteten Event-Repräsentationen gelten.

Für die Bestimmung der wesentlichen Eigenschaften ist es nun relevant, inwiefern obige konzeptuelle Repräsentationen genuin sprachlicher oder allgemein-kommunikativer Natur sind. Auch nach Halliday ist die Protowortproduktion nicht ausschließlich sprachlich. Vielmehr können nach seiner Interpretation auch Gesten oder andere kommunikative Marker die Funktion der Bedeutungsvermittlung übernehmen (vgl. Kap. 3.2.1). Insofern bildet die phonologische Verarbeitung nach dieser Konzeption kein ausdrückliches Kennzeichen der protosprachlichen Produktion, sie sitzt vielmehr einer schon zuvor stattfindenden Kommunikation auf und setzt diese ohne qualitative Neuerung kontinuierlich fort. Eine solche Genese zielsprachlicher Wortproduktionen aus einer allgemeinen symbolischen Kommunikation wird auch von neueren Ansätzen vertreten. Besonders Zeigegesten nehmen hier für die Erklärung des frühen Spracherwerbs eine Schlüsselrolle ein. Für eine genaue Bestimmung dessen, welche Rolle die phonologische Komponente in der Protowortproduktion einnimmt, ist eine genauere Beleuchtung dieses Erklärungsansatzes zwingend erforderlich.

3.2.6 Zeigegesten und ihr Zusammenhang mit dem frühen Worterwerb

In den Arbeiten von Tomasello wird der Verwendung von Zeigegesten eine basale Bedeutung für den Erwerb erster Zielwörter zugewiesen. Damit Zielwörter als konventionelle Symbole erworben werden können, ist der Gebrauch von Zeigegesten in spezifischen sozialen Interaktionen notwendig. Vor allem sogenannte Szenen der gemeinsamen Aufmerksamkeit,

häufig wiederkehrende Situationen oder Tätigkeiten, in denen eine Lenkung der Aufmerksamkeit mithilfe von Zeigegesten (und später Wörter) zuerst durch den Erwachsenen, später auch durch das Kind stattfindet. Damit das Kind einer solchen Aufmerksamkeitssteuerung durch symbolische Kommunikation zugänglich bzw. dazu fähig ist, muss es allerdings zuvor verstehen, dass sein Gegenüber ein Akteur mit Intentionen wie es selbst ist. Dieses Verständnis beginnt nach Tomasello ca. im Alter von neun Monaten. Was Tomasello hierbei hervorhebt, ist, dass der Eintritt und die Häufigkeit der Szenen der gemeinsamen Aufmerksamkeit eng mit Sprachverstehen und Sprachproduktion korrelieren, wie zum Beispiel in Carpenter, Nagell & Tomasello (1998) gezeigt wird. Häufigere Szenen der gemeinsamen Aufmerksamkeit bedingen somit auch ein größeres rezeptives und expressives Vokabular. Gemäß der Logik von Tomasellos Theorie werden früh produzierte Zeigegesten nicht symbolisch verwendet. Die spätere symbolische Verwendung von Gesten geht der Produktion von sprachlichen Symbolen allerdings klar voraus:

„[...] our claim is that in human ontogeny today infants from 12 to 14 months of age, before language acquisition has begun in earnest, already participate in the species-unique activity of human cooperative communication. This fact provides an existence proof that human-style cooperative communication does not depend on language, and rather suggests that language depends on it. Pointing may thus represent a key transition, both phylogenetically and ontogenetically, from nonlinguistic to linguistic forms of human communication.“ (Tomasello, Carpenter & Liszkowski 2007: 720)

Dieser Ansatz steht der hier vertretenen Vorstellung einer phonologie-basierten Selbststeuerung innerhalb eines kognitiven, modularen Systems maximal entgegen. Wie oben dargestellt, bildet hier die phonologische Struktur das interne Trägermedium, mit dessen Hilfe ein rudimentäres Selbstbewusstsein und mit ihm auch eine Repräsentation des intentionalen Selbst aufgebaut werden kann. Genuin sprachliche Bedeutung, die nichts mit sozial motivierter Intention zu tun hat, kann allerdings schon vorher generiert werden. Insofern kann man auch mithilfe dieser Theorie potentiell die Frage beantworten, „warum die frühe Sprache zusammen mit nicht sprachlichen sozio-kognitiven und sozial-interaktiven Fertigkeiten auftaucht.“ (Tomasello 2006: 146).

Was nach Tomasello allerdings nicht auftauchen dürfte, sind eben Protowörter, da sie mit der fehlenden Distanz zwischen Kind, Ausdrucksmittel und Ausdrucksinhalt keinen intersubjektiven Gehalt encodieren. Zudem dürften aus dem selben Grund keine Wörter mit

der Spezifikation PERSONAL produziert werden.⁶¹ Die empirische Überprüfung wird in dieser Arbeit entsprechend geleistet.

Es seien an dieser Stelle einige abschließende und allgemein kritische Anmerkungen zu Tomasellos Theorie angebracht. In dieser werden fast gänzlich die Ergebnisse zum frühen Spracherwerb, welche die eminenten Inputanalysefähigkeiten von Säuglingen empirisch bestätigen, ignoriert, obwohl gerade diese zeigen, dass der Spracherwerb schon mit prosodischem Input im Uterus beginnt und dies direkt nach der Geburt mittels Saugratenmessung nachgewiesen werden kann (vgl. Hennon et al. 2000). Da hier noch keine einzige soziale Interaktion stattgefunden hat, scheinen sie nach Tomasellos Anschauung gänzlich wegfallen zu können. Dies gilt insgesamt für die prosodische und phonologische Komponente, die anscheinend nicht als sprachlich qualifiziert werden muss. Diese rhetorische Figur, frühe Sprache simpel mit pragmatisch-kommunikativ begründeten Zielworterwerb zu identifizieren, finden sich m.W. in allen einschlägigen Veröffentlichungen von Tomasello. Eine Textstelle, wo es zuvor um den statistisch relevanten Zusammenhang zwischen gemeinsamer Aufmerksamkeit und sprachlichen Fähigkeiten geht, sei hier ausführlich exemplarisch angeführt:

„Dieser Befund ist deshalb von Bedeutung, weil er zeigt, daß die bekannte Altersentsprechung zwischen Fertigkeiten gemeinsamer Aufmerksamkeit und der Sprache – beide treten in den Monaten um den ersten Geburtstags herum auf, wobei die nichtsprachlichen Fertigkeiten gemeinsamer Aufmerksamkeit etwas früher erscheinen – kein Zufall ist. Die Schwierigkeit, die dieser Befund für Theorien des frühen Spracherwerbs aufwirft, die ihren Fokus nicht auf den sozialen Dimensionen des Prozesses haben, ist unmittelbar und schwerwiegend. Für Theorien, die sich in erster Linie auf die kognitiven Dimensionen des Wortlernens konzentrieren oder auf die beteiligten Prozesse des Assoziationslernens, stellt sich die Frage, warum der Spracherwerb genau zu dieser Zeit beginnt. Warum setzt er unmittelbar nach dem Erscheinen von Fertigkeiten gemeinsamer Aufmerksamkeit ein? [...] Meines Wissens ist die einzige Theorie des frühen Wortlernens und des Spracherwerbs, die diese Befunde erklären kann, die sozial-pragmatische Theorie, die von Bruner, Nelson und Tomasello vertreten wird.“ (ebd.: 145-146)

Zusammengefasst ist das Argument hier: Weil „die Sprache“ gleichzusetzen ist mit Zielworterwerb und dieser erst mit oder nach Szenen der gemeinsamen Aufmerksamkeit auftritt, können andere Theorien den eigentlichen Beginn des Spracherwerbs nicht erklären. Schopenhauer (1995) bestimmt ein solches Vorgehen in seiner Eristischen Dialektik als Kunstgriff 2⁶². Indem Kommunikation mit Sprache gleichgesetzt wird, wird durch diese

⁶¹ Entsprechend werden solche Optionen m.W. von Tomasello auch nicht erwähnt, was wiederum einer Identifikation von frühen Wörtern mit konventionellen Zielwörtern entspricht, wie wir sie auch schon bei anderen Ansätzen beobachten konnten.

⁶² „Die Homonymie benutzen, um die aufgestellte Behauptung auch auf das auszudehnen, was außer dem gleichen Wort wenig oder nichts mit der in Rede stehenden Sache gemein hat, dies dann lukulent widerlegen,

Homonymie eine *mutatio controversiae* (Veränderung der Streitfrage) auf den Weg gebracht, so dass genuin psycholinguistische Theorien plötzlich nichts Relevantes mehr zum Thema Sprache beitragen können. Diese Argumentation ist allerdings so krude, dass auch Schopenhauer schreibt: „Jedoch das offenbare Sophisma der Homonymie wird nicht im Ernst täuschen.“ In der hier vorliegenden eigenen Untersuchung soll gezeigt werden, dass genuin sprachliche Produktionen in der Form von Protowörtern unabhängig von Zeigegesten und Szenen der gemeinsamen Aufmerksamkeit auftreten. Auf diesem Wege soll auch das offenbare Sophisma von Tomasellos Argumentationen verdeutlicht werden. Eine weiteres von Tomasello ignoriertes Phänomen sind Spezifische Sprachentwicklungsstörungen. Diese sind gerade dadurch gekennzeichnet, dass isoliert sprachliche Beeinträchtigungen in Sprachproduktion und -verstehen auftreten, die sich nicht auf sensorische, organische, neurologische oder sozial-emotionale Defizite (inklusive extreme Milieuumstände) zurückführen lassen (vgl. z.B. Grimm 2000). Genau deshalb bilden solche sprachlichen Beeinträchtigungen auch ein Kardinalargument gegen Theorien, die, wie diejenige Tomasellos, soziale Interaktionen unter Aufgabe angeborener Kenntnissysteme als maßgeblich definieren. Ein Beispiel für eine Spezifische Sprachentwicklungsstörung ist der Fall P. (Dümig & Leuninger 2013), der hier aufgrund seines besonderen Störungsmusters mit einer feinkörnigen phonologischen Beeinträchtigung dargestellt werden soll.

3.3 Protowörter und morphosyntaktische Strukturen: Der Fall P.

Bei den bisher dargestellten Studien handelte es sich vornehmlich um Beobachtungen an sehr jungen Kindern, um festzustellen, in welchem genauen Lebensalter mit welchen sprachsystematischen Eigenschaften Protowörter erstmalig auftreten. Mit der Darstellung des Falls P. ist insofern ein *Novum* gegeben, als dass es m.W. die bisher einzige Studie ist, in der eine so früh einsetzende und so lang persistierende phonologische Plateubildung auf der Ebene der Protowortproduktion erfasst und untersucht wird. Die zusammenfassende Darstellung der Untersuchung und der Ergebnisse bezieht sich im Wesentlichen auf Dümig (2007), Dümig (2009), Dümig & Frank (2008) und Dümig & Leuninger (2013).

Schon seit dem Alter von ca. 3;0 Jahren war Kind P. in logopädischer Behandlung. Längere Krankenhausaufenthalte im Alter von 4;7 Jahren (fünf Wochen Intensiv-Sprachtherapie, dies war auch der Beginn der Studienaufnahmen) und 5;2 Jahren (vier Wochen Intensiv-

und so sich das Ansehn geben, als habe man die Behauptung widerlegt.“ (Schopenhauer 1995).

Sprachtherapie) machten es möglich, eine detaillierte Diagnose der Spracherwerbsstörung zu erstellen. Die nachfolgende Behandlung erfolgte danach einmal wöchentlich durch einen Sprachtherapeuten des Krankenhauses.

In spezifischen Testungen konnte gezeigt werden, dass P.s Spracherwerbsstörung nicht durch sprachexterne Faktoren bedingt sein konnte. Gemäß der Diagnose Spezifische Sprachentwicklungsstörung lagen keine neurologischen, sensorischen oder emotionalen Schädigungen sowie auch keine geistige Behinderung vor. Die nonverbale Testintelligenz lag zudem im Normbereich. Im sprachlichen Bereich konnte eine starke Dissoziation zwischen Sprachverstehen und Sprachproduktion festgestellt werden. Ersteres war völlig altersadäquat entwickelt, was sich in zwei Testungen mit den Reynell-Sprachentwicklungsskalen (Sarimski 1985) zeigt. Hier erreichte P. im Alter von 4;7 Jahren im Bereich Sprachverständnis ein Test-Entwicklungsalter von 4;10 Jahren und mit 5;2 Jahren eines von 5;5 Jahren. Ein anderer Sprachverstehenstest (TROG-D; Fox 2006) mit dem Hauptgewicht auf dem Verstehen syntaktischer Strukturen untermauerte diese Diagnose eines guten rezeptiven Sprachvermögens. Standardisierte Tests zur Sprachproduktion konnten zuerst aufgrund der sehr frühen Plateaubildung auf dem Niveau der Protowortproduktion nicht durchgeführt werden, da eine annähernde Zielwortproduktion die Voraussetzung zu deren Anwendung ist. Tabelle 23 zeigt P.s Wortschatz im Alter von 5;5 Jahren, der aus ca. 60 Wörtern bestand mit ungefähr 50% Ziel- und 50% Protowörtern.

Protoformen	Bedeutung der Protoformen	Zielformen ⁶³
am	<i>los</i>	o(ben)
amam	<i>fahren</i>	du (zu)
oapuff	<i>hin- und runterfallen</i>	Bäbi (Baby, klein)
Schlaferäusch	<i>Bett, schlafen</i>	ich
mmh	<i>mögen</i>	ao (auch)
Weingeräusch	<i>weinen</i>	auf
puff	<i>schießen, springen, Känguruh</i>	einch (eins)
hamham	<i>Lebensmittel, essen</i>	Uno (das Spiel)
m-hamham	<i>McDonald's</i>	Bä(r)
ham-ham-Ei	<i>Überraschungsei</i>	U-Ba(hn)
hamham-o + Gebärde für braun	<i>Doughnat</i>	Ei
iaobär	<i>Katzenbär</i>	Aichauto (Eisauto)
dudu	<i>Schuhe, evtl. Fuß, Bein</i>	Main (der Fluss)
i	<i>nass, glitschig, Regen</i>	Hai
heia	<i>Koalabrüder (aus Kinderbuch), fliegen</i>	(h)eiß
heiaheia	<i>Flugzeug</i>	eich/eis (Winter, Schnee, Eis, weiß)
wawa	<i>Hund</i>	Hammer
iao	<i>Katze</i>	Tomante (Tomate)
pu	<i>stinken</i>	Anananch (Ananas)
buabua	<i>Affe</i>	Baun (Baum)
buabuapiaopuff	<i>SOS-Affendorf</i>	Biene
auf-O	<i>Buchstabe C oder Hörnchen</i>	Ha(h)n
jai	<i>Lied, Musik</i>	Er(d)beer(e)
bababa	<i>Ente</i>	Bau (Ball)
hoho	<i>Weihnachtsmann</i>	
bibobi	<i>bitte</i>	
aua	<i>rot, heiß</i>	
Popo	<i>hinten, sitzen, Sitzmöbel, frech</i>	(Oma, Opa, Mama, Papa)
Anzahl: 28		Anzahl: 24 (28)

Tabelle 23: P.s Wortschatz im Alter von 5;5 (Dümig & Leuninger: 86-87).

3.3.1 P.s phonologische Entwicklung

Besonders auffallend in P.s Spontansprachproduktion war, dass er keine Koda-Plosive produzierte. In Abbildung 53 und 54 ist die prozentuale Verteilung von Demisilben innerhalb

⁶³ In Klammern stehen Ergänzungen oder die entsprechenden zieladäquaten Formen.

eines Datensets von 1884 Silben dargestellt. Initial wurden keine Liquide produziert, Plosiv-Vokal-Demisilben bilden eindeutig die frequenteste Form (Abb. 53).

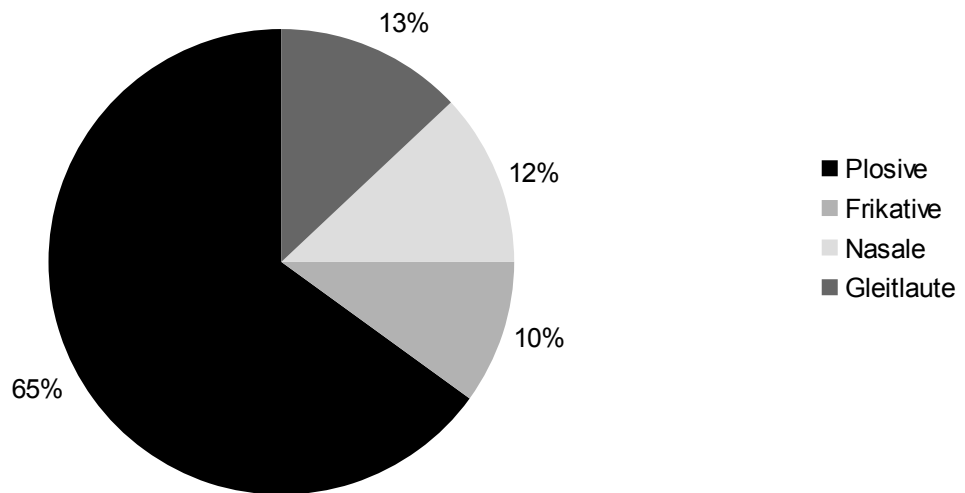


Abbildung 53: Frequenzangaben initialer Demisilben in P.s Produktion.

Die Verteilung von finalen Demisilben stellt sich wesentlich anders dar. Mit 82% bilden reine Vokal-Demisilben die klar häufigste Variante, während Plosive und Liquida überhaupt nicht vorkommen. Während diese Ergebnisse für den Erwerb von Liquida zeigen, dass sie als Lautklasse noch nicht erworben wurden, scheinen Plosive, die ja initial durchaus auftreten und deshalb erworben wurden, einer spezifischen Silbenpositionsbeschränkung zu unterliegen (Abb. 54).

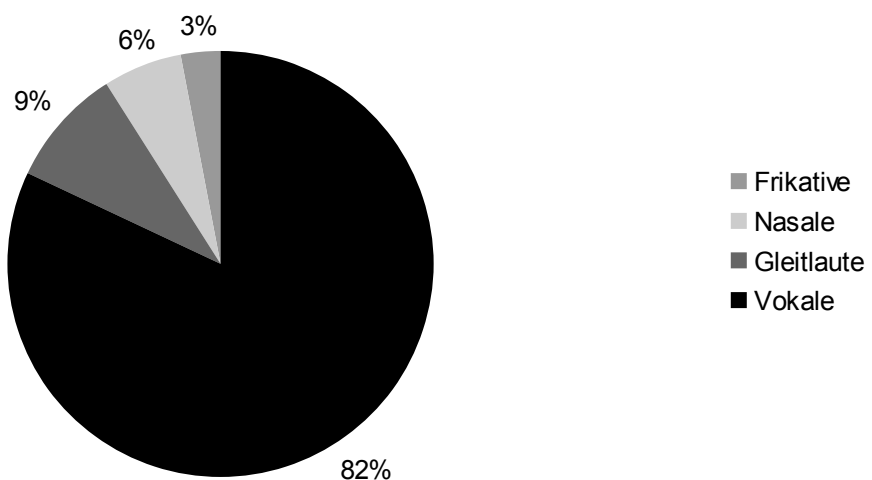


Abbildung 54: Frequenzangaben finaler Demisilben in P.s Produktion.

Diese Silbenpositionsbeschränkung wird in Dümig & Leuninger (2013: 91) in Bezug auf die Markiertheitstheorie von Clements (1990) erfasst:

„Plosiv-Vokal-Verbindungen sind initial nicht-komplex bzw. unmarkiert, hier gibt es keine Probleme für den Sprachprozessor. Vokal-Plosiv-Verbindungen sind final allerdings die komplexeste, das heißt markierteste Form. Hier zeigen sich eindeutig Sprachverarbeitungsschwierigkeiten.“

Diese Beschränkung der Produktion der komplexesten bi-positionalen Demisilbe der Form Vokal-Plosiv hatte, wie wir sehen werden, gravierende Folgen für den Ausbau sprachlicher Strukturen über Modulgrenzen hinweg. In Dümig (2007) wurde gezeigt, dass sich im Zuge der ambulanten Therapie im Alter von 5;3 bis 5;5 Jahren in der Spontansprache drei Mikroschritte in der phonologischen Entwicklung identifiziert ließen. Die Entwicklung verlief über die fehlende Produktion von Koda-Plosiven (Phase I) über die Ersetzung silbenfinaler Plosive durch sonorantere Laute (Nasale) (Phase II) hin zu einer zieladäquaten Produktion silbenfinaler Koda-Plosive (Phase III) (vgl. Tabelle 24).

Lautklasse	Phase I (Plateauphase)		Phase II (Übergangsphase)		Phase III (Zielformphase)	
	Fehlende Produktion von Koda-Plosiven		Substitution von Koda-Plosiven durch Nasale		Produktion von Koda-Plosiven	
	Produktion	Ziel	Produktion	Ziel	Produktion	Ziel
Nasal	[ʔam]	an	[majn]	Main	[bajn]	Bein
	[ham]	essen	[bawn]	Baum	[nojn]	Neun
Frikative	[ʔeç]	ich	[hawf]	Haus	[maos]	Maus
	[ʔo:x]	hoch	[ʔajs]	Eis	[paf]	Fass
Plosive	[bo]	Bob	[bon]	Boot	[bo:t]	Boot
	[pao]	papp	[ban]	Bad	[ba:t]	Bad
			[don]	tot	[hu:t]	Hut
			[bawn]	baut	[be:t]	Beet
			[han]	hab'		

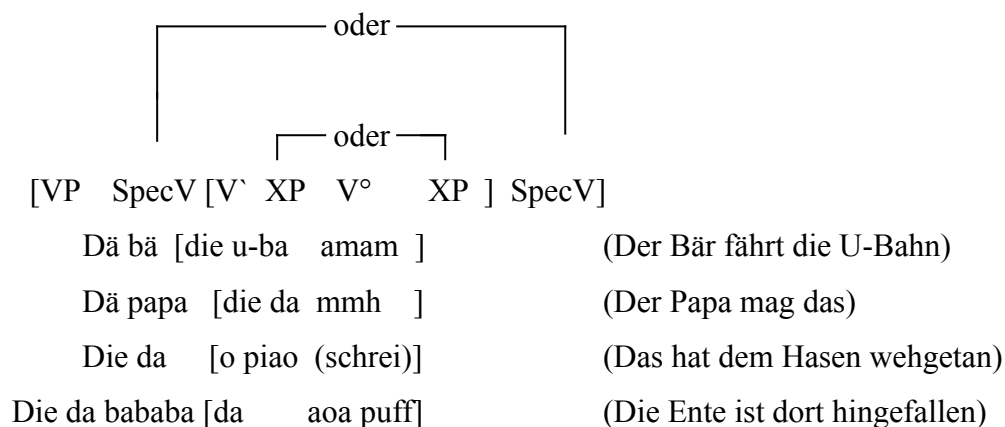
Tabelle 24: Phasen des Demisilbenerwerbs (nach Dümig & Leuninger 2013: 93).

Problematisch gestaltet sich eine Einordnung der beobachteten Produktionen in die Erwerbsmodelle von Fikkert (1994) und Demuth (1995). P. konnte bis zur Phase III keine

Plosive in der Silbenkoda oder extra-reimischen Position produzieren, so dass schon Stufe 1 des Fikkert-Modells nicht adäquat gemeistert scheint. Andererseits produzierte er schon lange Formen, die sich strukturell den Stufen 2,3 und 4 zuordnen ließen. Auch wenn man die Silbenentwicklung von P. nach Demuth in die Phase der Minimalen Wörter einordnet, fehlt eine Passung mit den Daten. P. reduplizierte auch CVC-Silben (wie z.B. [hamham]) und produzierte zudem schon extra-reimische Konsonanten, ohne dass er die Vokallängenunterscheidung schon systematisch repräsentieren konnte. Der oben beschriebene Zugang, P.s Produktionen im Rahmen der Markiertheitstheorie von Clements zu erklären, ist in Anbetracht der Schwierigkeiten mit den anderen Modellen also der probate. Welche Auswirkungen das phonologische Plateau und seine Überwindung auf die morphosyntaktische Entwicklung hatte, soll im folgenden Kapitel betrachtet werden.

3.3.2 P.s morphosyntaktische Entwicklung

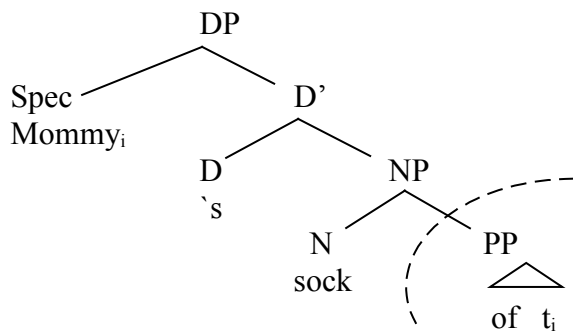
Trotz seines reduzierten Wortschatzes nutzte P. syntaktische Strukturen, welche denjenigen sprachunauffälliger Kinder in einem früheren Lebensalter völlig gleichen. So lag bei P. eine Plateaubildung auf Meilenstein 2 mit dem infiniten Verb in Verbendstellung (nach Tracy 1995) vor, wobei die grundlegende Parametrisierung für diese Zwischengrammatik vollkommen zieladäquat, d.h., in Übereinstimmung mit der Zielgrammatik war:



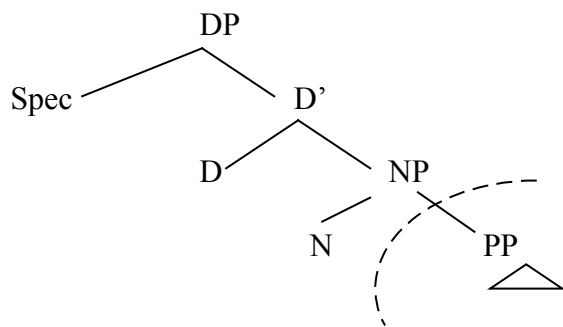
Theta-Rollen-Zuweisung und Subkategorisierung waren völlig adäquat. Interessanterweise hatten selbst onomatopoetische Formen oder Lautmalereien wie *mmh* eine Argumentstruktur und lizensierten als Kopf V° eine komplette VP:

[SpecV' *Dä papa* [V' [NP *die da mmh* V°]]] (der Papa mag das)

P. produzierte spezifische Konstruktionen, um den Genetiv zu markieren. Diese Konstruktionen können im Rahmen von Penner und Weissenborns Universal Determiner Requirement Hypothesis (UDR) (1996) analysiert werden. Da im Berner Schweizer Deutsch im Alter von ca. 1;07 Formen wie *Mami sis* (Mami ihre (sis neutrum)) vorkommen, im Deutschen aber N + N-Strukturen wie *Nomi Bett*, wird von Penner und Weissenborn die UDR aufgestellt, welche eine Phrasenstruktur wie in (1) für alle Entwicklungsstufen postuliert:



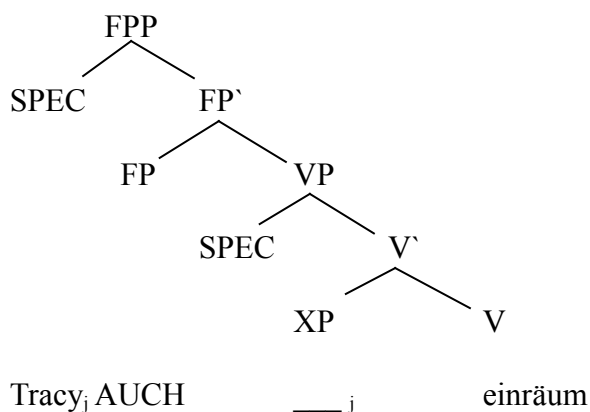
Mit Hilfe dieses phrasenstrukturellen Formats kann man besitzanzeigende Äußerungen von P. analysieren, der in der Phase I (Plateauphase) keine Präpositionen oder Possessivpronomina bilden konnte.



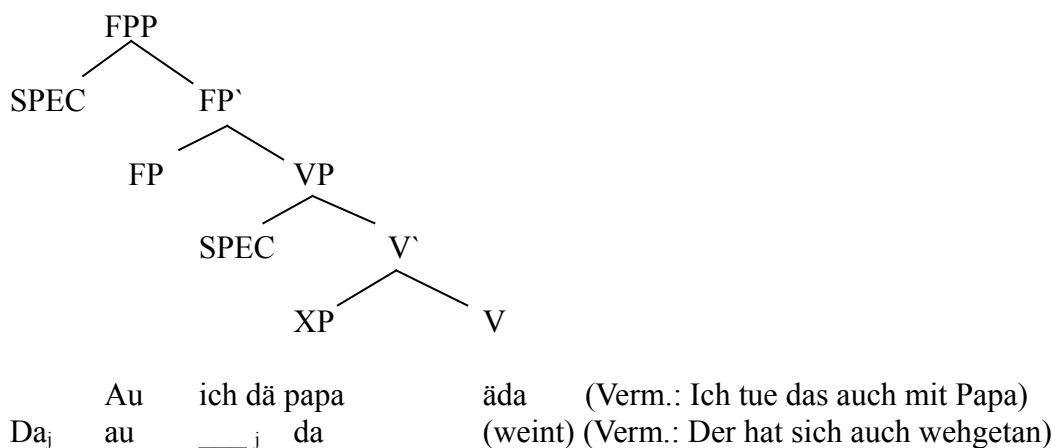
die mama _i	die	mami	(von)___t _i	(die Mutter von der Mama)
ich _i	die	wawa	(von)___t _i	(mein Hund)
ich _i	de	papa	(von)___t _i	(mein Papa)
ich _i	die	mami	(von)___t _i	(meine Mami)

P. produzierte zudem systematisch Konstruktionen mit der Fokuspartikel AUCH (phonologisch reduziert als [ʔau]). Nach Penner, Tracy & Weissenborn (1999) „übernimmt

die Fokuspartikel AUCH eine maßgebliche Rolle für den weiteren Strukturaufbau, indem sie als eigenständiger Kopf eine vollständige funktionale Strukturschicht oberhalb der (Wurzel-)VP eröffnet und damit diese VP zu ihrem Komplement macht.“ (ebd.: 236). Die so entstehenden Fokuspartikelphrasen (FPP) beinhalten mit der Spezifikatorposition (SPEC) einen zusätzlichen Landeplatz für Argumente aus der VP. So kann man eine Art „AUCHzweit“-Effekt beobachten, weil wie bei späteren V2-Strukturen des Deutschen nur eine Position zur Verfügung gestellt wird, in welche Argumente bewegt werden können:

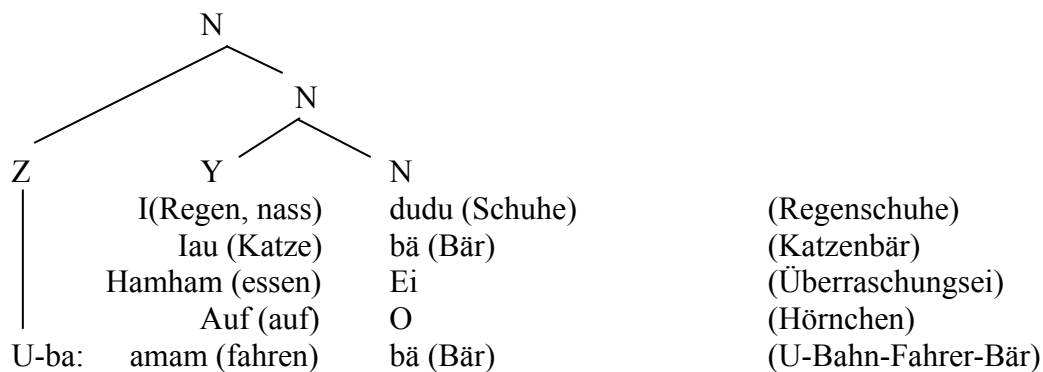


Wie Tracy feststellt (2000: 30), ist beim Auftreten von *auch* zu beobachten, dass Verben entweder völlig fehlen (z.B. „alle haben Augen...Sabine auch Augen“), nicht finit sind (z.B. „Julia auch Buch angucken“) oder auch als finite Verben am Satzende verbleiben (z.B. „Eichhörnchen auch noch mehr steht“). In P.s Produktionen lässt sich genau dieses syntaktische Verhalten von *auch* beobachten. Tritt die Fokuspartikel auf, dann fehlen Verben entweder völlig (z.B. „ao ich die mami“) oder verbleiben am Satzende (z.B. „au ich dä papa äda“). Analog lassen sich in P.s Produktionen einige Äußerungen, mit dem dargestellten AUCH-Format analysieren:



	Au	ich die mami	(Verm.: Ich habe auch eine Mama)
Ich;	au	___j dä da	(Verm.: Ich habe auch das)

Auch Komposita werden von P.s Sprachsystem hinsichtlich des Kopfrichtungsparameters zielkonform generiert. Im Deutschen ist der Kopfrichtungsparameter für Kopf rechts spezifiziert. Endozentrische Determinativkomposita sind im Deutschen dadurch charakterisiert, dass das Erstglied das Zweitglied des Kompositums spezifiziert. P. verwendet Protowörter in solchen Komposita, um die Bedeutung existierender Zielformen wiederzugeben. Er folgt dabei den morphologischen Regeln des Standarddeutschen und setzt die phonologischen Formen nach Maßgabe der thematischen Rollen der strukturellen Repräsentation ein. Interessanterweise greift er auch auf exozentrische Strukturen zurück, um auf etwas zu referieren, wie z.B. „auf o“ (offenes O) für „Hörnchen“.



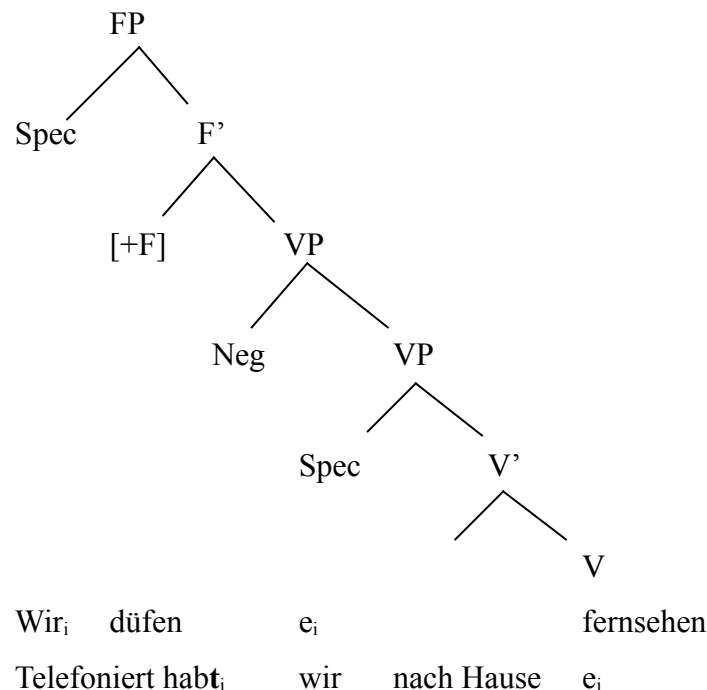
Zusammenfassend kann man festhalten, dass P. auf morphologischer und syntaktischer Ebene keine Abweichung von den Strukturen der Zielgrammatik produzierte und auch Protowörter in solchen Strukturformaten adäquat verwendet wurden. Darüberhinaus konnte P. mit Hilfe seiner strukturellen Möglichkeiten mehr Information anzeigen, als ihm durch seinen phonologischen Stand möglich war.

Nach der etablierten Option Koda-Plosive zu produzieren, erzielte P. im Alter von 5;8 Jahren in einem Wortschatztest (AWST-R; Kiese-Himmel 2005) einen Prozentrang von 77% (t-Wert: 57). Ein t-Wert von 40-60 ist ein durchschnittliches Ergebnis für das Lebensalter von 5;5 Jahren (Höchstalter für diesen Test). Da P.s Ergebnis aber im oberen Bereich angesiedelt ist, kann man für sein Lebensalter von einem durchschnittlichen Wortschatz ausgehen.

Im Alter von 5;9 wurde der morphosyntaktische Sprachstand mittels ESGRAF untersucht. Hier konnten entscheidende Veränderungen auf morphosyntaktischer Ebene beobachtet werden.

P. produzierte nun die overt morphologischen Affixe -t (wie z.B. in „Telefoniert **habt** wir nach Hause“ (Wenn wir telefoniert haben, dann kommen wir nach Hause) und die Infinitivform -n (wie z.B. in „Ich erst ma **suchen**“ (Ich suche erst mal)), wobei das Affix -e weiter systematisch ausgelassen wurde (wie in Ich schieb**X** (Ich schiebe)).

P. befand sich nun nach Clahsen auf Stufe II des Flexionsmorphemerwerbs. Man konnte an den Produktionen von P. erkennen, dass das Flexiv -t wie bei sprachunauffälligen Kindern ein anderes syntaktisches Verhalten aufweist, als -n und Ø. Die Flexive -n und Ø wurden unter V basis-generiert, während das Flexiv -t als Kopf der kategorial un spezifizierten FP basis-generiert wurde. Zur Verdeutlichung noch einmal die von Clahsen vorgeschlagene Struktur der FP mit Beispielen aus P.s Produktionen:



Auch konnte man in P.s Äußerungen einen ersten zielgerechten Gebrauch von indefiniten Artikeln ausmachen, wobei aber der definite Artikel noch weitestgehend ausgelassen oder mit inkorrekt markierter Kasusmarkierung produziert wurde. Dass die indefiniten Artikel korrekt produziert wurden, ist eventuell auf ihren kategorialen Status im Vergleich zu definiten Artikeln zurückzuführen. So fasst z.B. Vater (1984) den indefiniten Artikel als Quantor auf, das heißt, er ist hinsichtlich der Definitheit von Nominalphrasen unmarkiert. Ohne auf diese spezifische Diskussion genauer einzugehen, kann man aber festhalten, dass in P.s

Produktionen mit dem indefiniten Artikel der Einstieg in das Konzept der Quantifikation nominaler Referenten zu beobachten war.⁶⁴

Zudem sieht man an den Daten, dass noch nicht systematisch dreigliedrige Nominalphrasen produziert werden konnten. Entweder fand man die Kombination indefiniter Artikel + Nomen (wie z.B. „ein Schnüssel“ (ein Schlüssel)) oder Adjektiv + Nomen (wie z.B. in „rotes Bonbon“ (rotes Bonbon)) vor. Dies lässt sich in eine geordnete Erwerbsreihenfolge einordnen, wie sie Bittner (1998) beobachtet hat. Es findet in dieser eine sukzessive Zunahme von referenzspezifizierenden Merkmalen statt (Bittner 1999: 56):

ein Haus → *rotes Haus* → *das Haus* → *das rote Haus*

Von den fünf deutschen Pluralallomorphen (overt: -(e)n, -s, -e, -er; covert: -Ø) produzierte P. nun alle Formen bis auf den -(e)n Plural. Das Pluralflexiv -s ist dasjenige, das als einzige Pluralmarkierung übergeneralisiert wurde, was sich mit Beobachtungen aus dem ungestörten Spracherwerb deckt. Hier wählen sprachunauffällige Kinder zwar häufig auch den -e- und -(e)n-Plural als Default-Wert, der -s Plural wird aber am häufigsten übergeneralisiert⁶⁵ (vgl. Bartke 1998).

P. produzierte nur noch Zielwörter, so dass nun der Wert für die Mittlere Satzlänge ermittelt werden konnte (engl. Mean Length of Utterance = MLU). Er betrug 3.11.

Wie in Dümig & Leuninger (2013) betont wurde, ist dieser Sprung von einem Spracherwerb in Zeitlupe zu einem Spracherwerb im Zeitraffer nur adäquat mit dem Erwerb der komplexen Demisilbe Vokal-Plosiv in der Produktion zu erfassen, der einen phonologischen Bootstrapping-Effekt auf die flexionsmorphologische und syntaktische Ebene hatte:

„Dass morphologische Flexionsformen realisiert werden konnten, die letztlich auch elaboriertere Satzkonstruktionen verursachten, hatte anscheinend als phonologische Voraussetzung den Erwerb aller Lautklassen in finalen, zweipositionalen Demisilben. Solange die Beschränkung qua Demisilbenkomplexität dominierte, wurden vor allem nicht-komplexe -V-Demisilben gebildet, so dass keine Flexionsformen angefügt werden konnten.“ (ebd.: 94)

⁶⁴ Bittner (1999) verbindet diesen Einstieg mit dem Erwerb der Merkmalopposition [+einzahlig] vs. [-einzahlig].

⁶⁵ Das -s Flexiv ist vor allem aufgrund seiner hohen Kontextfrequenz als „Notplural“ geeignet, d.h., es kann z.B. in Kontexten wie in den Pluralformen von Namen (die Meier-s), Kurzformen (Lok-s), Akronymen (LKW-s), Entlehnungen (Kiosk-s) und Kunstwörtern mit ungewöhnlicher Struktur (pnähf-s) verwendet werden. Der Plural mit der höchsten Oberflächenfrequenz ist aber der -(e)n-Plural (vgl. Niedeggen-Bartke 1999).

Ein Aspekt der von Dümig & Leuninger (ebd.) nicht erwähnt wird, ist der Zusammenhang zwischen dem Ausbau der Silbenpositionen und der Bauplanüberfrachtung durch Affrikata, wie sie in Kapitel 1.4.1 beschrieben wurde. Da im Deutschen Affrikaten eine Sequenz aus Plosiv+Frikativ sind, sehen wir an P.s phonologischem Plateau die feinkörnige Interaktion von Demisilbenkomplexität, Affrikatenbildung und Silbenpositionsausbau. Da er aufgrund der Demisilbenkomplexität keine Plosive in der Koda bilden konnte, war die Produktion von Affrikaten entsprechend auch nicht möglich und mit ihnen entsprechend keine Bauplanüberfrachtung, die weitere Silbenpositionen freischalten könnte. Die Unmöglichkeit, die Demisilbe Vokal-Plosiv zu produzieren, führte also direkt zur fehlenden Applikationsmöglichkeit der Suffigierungsregel mit dem Flexionsmorphem -t und indirekt zu weiteren fehlenden Positionen für Flexionsmorphemen in der Produktion. Mit der Möglichkeit die komplexe Vokal-Plosiv-Demisilbe zu produzieren, erfolgte entsprechend eine Aufhebung der Plateaubildung. Auch die Produktion von Protowörtern hing offensichtlich eng mit der beschränkten Demisilbenproduktion zusammen. Dieser Aspekt soll im Folgenden genauer untersucht und eingeordnet werden.

3.3.3 Der Fall P. unter modelltheoretischer Perspektive

Protowörter sind per definitionem Formen, die in ihrer phonologischen Form keinen Bezug zu Zielformen der Erwachsenensprache haben. Sie können also im Erweiterten Garrett-Modell analog zu den Produktionen von Jargon-Aphasikern als Substitute für Formen aus dem Zielwortlexikon betrachtet werden. Während allerdings bei Jargon-Aphasikern der Zugriff aufgrund eines blockierten Output-Lexikonzugangs nicht vonstatten geht, ist im frühen Spracherwerb ein nicht hinreichend gefülltes Output-Lexikon vorhanden. Die Ersatzformen werden aus zwei unterschiedlichen Gründen inseriert, erfüllen nichtsdestotrotz aber in beiden Fällen die Funktion, ein minimales Set phonologischer Elemente in der Sprachplanung bereit zu stellen. Wie bei den Produktionen von Jargon-Aphasikern bestehen auch die Produktionen von P. vorwiegend aus nicht-komplexen Demisilben. Auch wenn die die morphosyntaktischen Formen bis zum Alter von 5;8 Jahren nicht altersadäquat sind, so scheint doch im Angesicht der kreativen Verwendung dieser Formen der Zugang zu Informationen der funktionalen Ebene vorhanden zu sein. In Dümig & Leuninger (2013) wird darauf hingewiesen, dass die relevanten Flexionsinformationen für satzstrukturelle Repräsentationen wahrscheinlich durch die phonologischen Silbenkoda-Beschränkungen nicht eingefügt und insofern die

morphosyntaktische Plateaubildung als ein Beiprodukt der phonologischen betrachtet werden kann. Ähnlich wie die frühen Protowörter von Nigel (vgl. Kapitel 3.2.1) lassen sich auch die Protowörter von P. nun im Rahmen der Parallelen Architektur charakterisieren. Sie bestehen aus nicht zieladäquaten phonologischen Repräsentationen, die mit morphosyntaktischer und konzeptueller Struktur assoziiert sind und insofern einen stabilen lexikalischen Eintrag bilden. In Anlehnung an Jackendoff (2002) lässt sich dieses Informationskonglomerat in der Sprachproduktion als fortgeschrittene Aktivierung von Subkomponenten innerhalb der Parallelen Architektur darstellen:

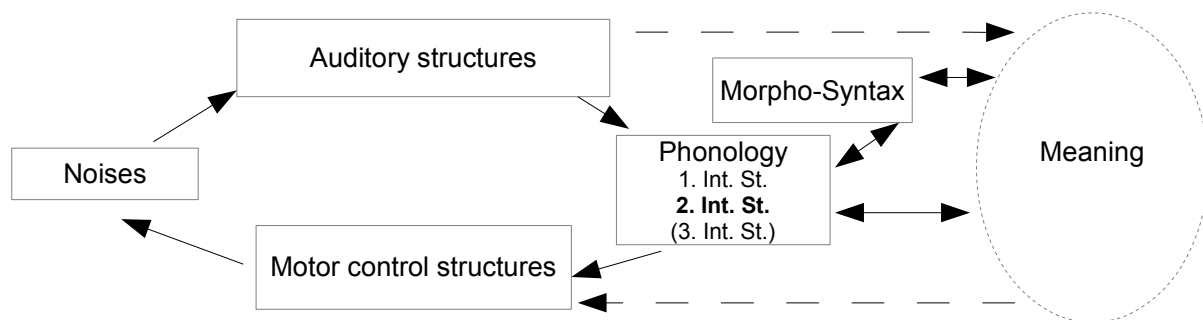


Abbildung 55: Verknüpfung von Subsystemen in der späten Protowortproduktion von P.

Da P. keine Wörter produzierte, die die phonotaktischen Regeln des Deutschen verletzen und auch schon anteilig adäquate Zielwörter produzierte, sind alle drei Integrationsstufen (im Phonologie-Modul der Abb. 55 mit Int. St. abgekürzt) zu verzeichnen. Allerdings verhindert sein selektives Defizit der Kodaplosivproduktion (auf der 2. Integrationsstufe, der Produktion von Oberklassenmerkmalen (durch Fettdruck gekennzeichnet)) die weitere Entfaltung des Gesamtsystems (so auch die Produktion von Artikulationsart und -ortsmerkmalen auf der 3. Integrationsstufe (durch Klammerung gekennzeichnet)).

Der Fall P. macht deutlich, dass es ungenügend ist, Protowörter einerseits nur unter der Perspektive einer umschriebenen Spracherwerbsphase und, was eng mit dieser Perspektive zusammenhängt, andererseits nur als Kombination eines artikulatorischen Musters mit einem Bedeutungsgehalt zu betrachten. Vielmehr kann unter der Protowortproduktion auch ein Zustand des phonologischen Sprachproduktionssystems gefasst werden, in dem Protowörter als genuin phonologische Formen zusätzlich zu einem konzeptuellen Gehalt auch eine morphosyntaktische Spezifikation beinhalten. Die Berechnung der Demisilbenkomplexität scheint in diesem Zustand des phonologischen Sprachproduktionssystems ein entscheidender Faktor zu sein. Insofern zeigen sich hier Analogien zur Sprachproduktion von Jargon-Aphasikern, die in Kapitel 1.5.2 modelltheoretisch erfasst wurde. Im nächsten Kapitel soll

geprüft werden, ob sich die beobachteten Analogien für die modelltheoretische Beschreibung der Protowortproduktion fruchtbar machen lassen.

3.4 Protowortproduktion: Eine vorläufige modelltheoretische Synthese

3.4.1 Protowortproduktion im erweiterten Garrett-Modell

Nach Sichtung der vorhandenen Literatur zum Protoworterwerb ergibt sich sowohl in der phonologischen als auch der semantischen Klassifikation ein recht uneinheitliches Bild. Es ist zur Klärung des Phänomens von immenser Wichtigkeit in einer modelltheoretischen Beschreibung die partiellen Ergebnissen der Protowortforschung zu synthetisieren und in einer solchen Abstrahierung ein umfassenderes Verständnis sowohl von der Klassifikation, was ein Protowort sprachsystematisch sei, als auch von der Funktion, die es im sich aufbauenden Sprachsystem einnimmt, zu gewinnen. Hierzu sollen zuvor noch einmal die phonologischen und lexikalisch-semantischen Charakteristika benannt werden, die in der bisherigen Forschung hervorstechen und somit als Grundlage einer Arbeitshypothese gelten können:

1. Die (CV-)Silbe scheint die maßgebliche Einheit der Protowortproduktion zu sein.
2. Reduplikation ist ein aktiver Prozess in der Protowortproduktion.
3. Protowörter treten je nach Definition in verschiedenen Varianten auf, wobei die Verarbeitung innerhalb der phonologischen Komponente hauptsächlich nur auf der 1. oder der 2. Integrationsstufe geschieht. Folgende Kombinationen mit anderen Subkomponenten konnten beobachtet werden:
Artikulation/Geste + Phonologie + Bedeutung
Artikulation/Geste + Phonologie + Morphosyntax + Bedeutung
4. Der Erwerb von Protowörtern korreliert negativ mit dem Erwerb von Zielwörtern. Ob dies durch einen unterschiedlichen lexikalischen Status oder einer fortschreitende Phonologisierung bedingt ist, bleibt unklar.
5. Auch wenn Halliday Protowörter als voll symbolisch klassifiziert, so fehlt ihnen hierzu aber das entscheidende Kriterium, nämlich Dekontextualisierung (vgl. Bickerton 2009). Der konzeptuelle Gehalt ist vielmehr, gleichfalls nach Halliday, intentional, d.h. diskret und funktional spezifizierbar, dennoch aber auf einen konkreten Kontext mit gegebenen Personen und Objekten bezogen bzw. durch diesen evoziert.

Wenn man mit Jackendoff (2002) lexikalische Einträge als Konglomerate sprachlicher Informationen versteht, dann definiert die Unterscheidung im konzeptuellen Gehalt bereits, dass ein stratifiziertes Lexikon im Sinne differenter Abrufpfade zu Beginn des Spracherwerbs der Normalfall ist. Die Informationskorrelate, welche Protowörter konstituieren, sind nicht nur durch einen spezifischen semantischen Gehalt, sondern natürlich auch durch eine

spezifische morpho-phonologische Form spezifiziert. Unter konsequenter Berücksichtigung der Precedence-Based Phonology (vgl. Kap. 1.7.1) muss konstatiert werden, dass Totale Reduplikation als unmarkierter Fall, d.h. als Defaultoption der UG, die primäre Option in der Wortproduktion darstellen sollte. Die Nähe von totaler Reduplikation zur Präfigierung und Suffigierung machen es wahrscheinlich, dass diese Optionen durchaus im Sinne von Yangs Variational Model (vgl. Kap. 2.2.4.1) parallel zur Verfügung stehen und eine vollständige Etablierung der morphologischen Markierung qua Affigierung in der Produktion erst durch den Erwerb der Kodaposition als Landeplatz für das phonologische Material möglich ist. Die Produktion liegt also in der zeitlichen Entwicklung hinter perzeptiv bereits belegten Parameteroptionen zurück. Im Produktionsprozess wird dementsprechend die realisierbare, aber in der Kompetenz schon überholte Option ausgewählt. Sobald allerdings in der Produktion die strukturellen Voraussetzungen gegeben sind, werden die in der Perzeption schon verfügbaren Formen realisiert. Auf diese Interaktion von morphologischer Markierung und Silbenstruktur wurde in Kapitel 3.3.2 hingewiesen. Welche phonologischen Formen werden nun aber in der Protowortproduktion realisiert? In Kapitel 1.5.2 wurde dargestellt, dass Jargon-Aphasiker bei fehlendem Zugang zu den phonologischen Formen des Output-Lexikons Zufallssilben produzieren, deren Komplexität nach Maßgabe des Dispersionsprinzip gering ist. Die kognitive Komponente, die für diese Generierung verantwortlich ist, bezeichnet Buckingham (1990) als Random Generator, der allerdings bei funktionierendem Zugang zum Form-Lexikon im Erwachsenenalter nicht in Gebrauch ist; er ist sozusagen im Standby-Modus. Die Annahme, die an dieser Stelle nun gemacht werden soll, ist die, dass der Random Generator auch aktiv ist, wenn keine oder nur wenige Zielformen im Output-Lexikon gespeichert sind. Wir finden also innerhalb des Sprachproduktionssystems von Jargon-Aphasikern und Kindern im frühen Spracherwerb eine analoge Situation vor, nämlich die Nicht-Verfügbarkeit von phonologischen Einträgen im Output-Lexikon. Entsprechend den Untersuchungen an Jargon-Aphasikern wäre demnach die Voraussage für Protowortproduktionen, dass sie aus Demisilben mit geringer Komplexität bestehen. Anhand von Abbildung 56 soll das hier propagierte Modell der frühen Sprachproduktion noch einmal exemplifiziert werden. Wie in Kapitel 1.2.2 ausgeführt, existiert m.E. kein neuroanatomisches Argument gegen die Annahme, dass auch im frühen Spracherwerb eine modulare Komposition des Sprachproduktionssystems gegeben ist. Basis soll dementsprechend hier das durch Shattuck-Hufnagel erweiterte Garrett-Modell bilden, wobei auch dieses noch weiter komplementiert werden muss. Auf der Bedeutungsebene ist hier eine konzeptuelle Struktur

auf Basis von Jackendoff (2007a; vgl. Kap. 3.2.5) angegeben. Auf morphologischer Ebene ist diese semantische Information mit keinem spezifischen segmentalen Gehalt assoziiert, wobei als Default die abstrakte Regel für Totale Reduplikation appliziert wird. Der segmentale Gehalt wird deshalb vom Random Generator in den Buffer inseriert, in dem eine linearisierte und reduplizierte Phonemkette zwischengespeichert wird. Weiter wird diese Phonemkette standardmäßig in das mögliche silbenstrukturelle Format integriert. Da dieses im frühen Spracherwerb eine CV-Silbe ist, besteht das Endprodukt der Silbifizierung aus einer reduplizierten CV-Silbe.

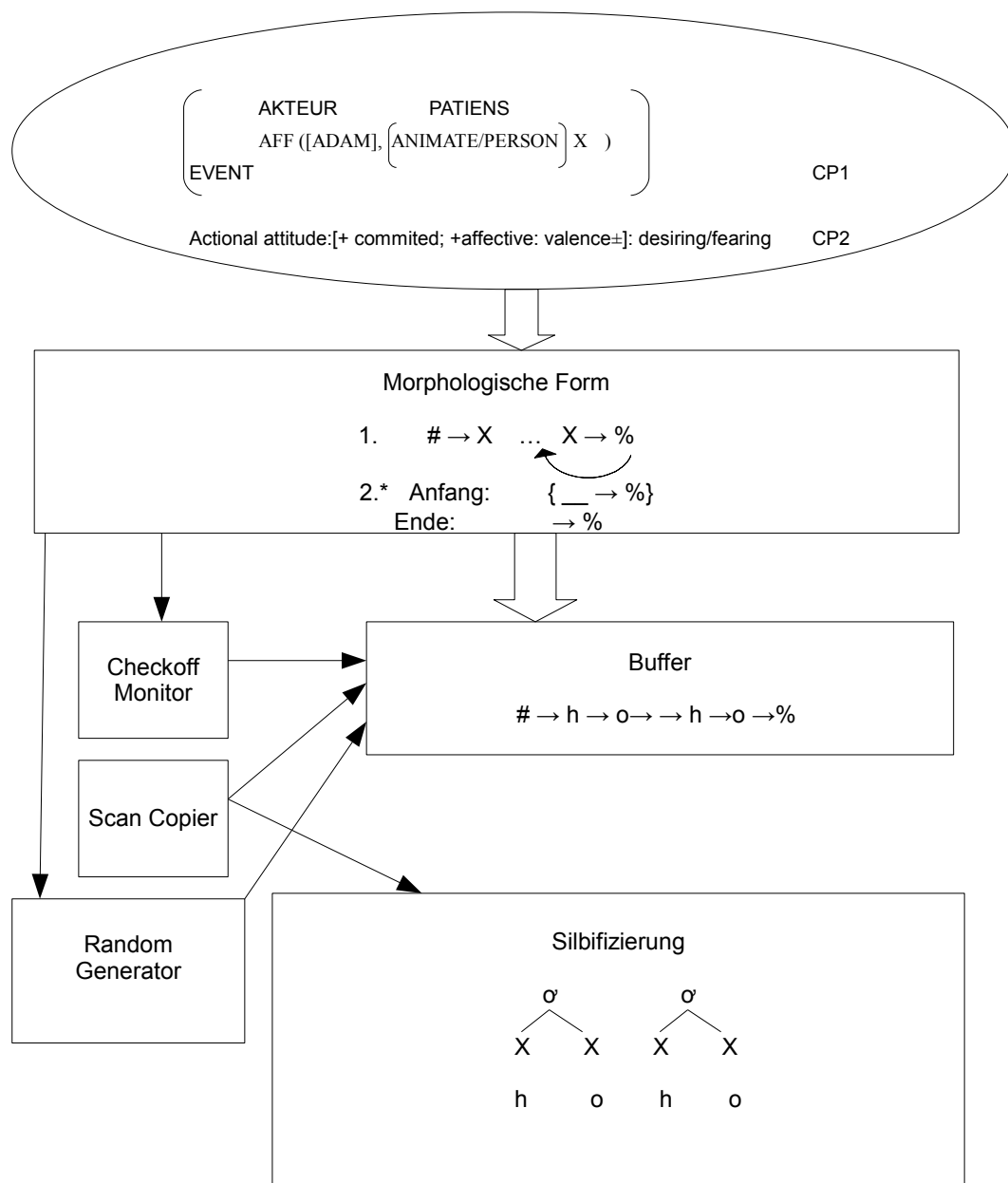


Abbildung 56: Protowortproduktion im erweiterten Garrett-Modell.

In Kapitel 1.5.2 wurde anhand von Versprecherdaten ein Modell der sukzessiven Integration segmentaler Informationen in die jeweiligen Silbenpositionen motiviert. D.h., man muss für den Silbifizierungsprozess sowohl eine syntagmatische als auch eine paradigmatische Ebene annehmen. In Kapitel 3.3.2 wurde bereits auf die Rolle der Komplexität von Demisilben für den syntagmatischen Aufbau hingewiesen. Paradigmatisch bedeutet ein Mehr an Struktur auch ein mehr an Komplexität, was dementsprechend eine erhöhte Integrationsleistung erfordert. Ein Strukturaufbau im Sinne von Rice & Avery (1995) in der Produktion lässt sich mit Hale & Reiss (2008) dementsprechend als eine im Spracherwerb stattfindende, durch Performanzfaktoren limitierte, Integration von bereits vorhandener Information umdeuten. Dies bedeutet, dass die drei Verarbeitungsstufen paradigmatischer Information (vgl. Kap. 1.5.2) eine sukzessive Etablierung erfahren bzw. immer mehr Information zugeführt werden kann. Dies ergibt für Produktionen, die nur auf der 1. Integrationsstufe verarbeitet werden, dass sie noch keinen phonotaktischen Regularitäten der Zielsprache gehorchen, da in der Produktion noch keine Konsonant-Vokal-Distinktion gegeben ist. So verweist Piske (2001: 214) z.B. auf die Form [gⁿɨ], die von Kind LEO produziert wurde und die er als Beleg gegen frühe CV-Silben anführt. Was allerdings auch in dieser Form gegeben ist, ist eine Abfolge von [-silbisch] und [+silbisch], eine Spezifikation für eine vokalische Merkmalsmenge ist hier allerdings noch nicht gegeben. Diese erfolgt erst, wenn auf der 2. Integrationsstufe die Oberklassenmerkmale etabliert wurden. Relevant ist vor allem, dass bis zur Etablierung der 3. Integrationsstufe eine overte Ähnlichkeit zu Zielwörtern nicht gegeben sein kann. Deutlich sollte aber auch werden, dass schon ab der Etablierung der 1. Stufe von einer phonologischen Verarbeitung in der Produktion die Rede sein muss, da auch hier phonologische Kontraste verarbeitungsrelevant sind. Es lässt sich an dieser Stelle natürlich fragen, ob man unter der Perspektive der o.g. sukzessiven Integration von zielsprachlicher Information weiterhin einen Random Generator annehmen muss. Diese Frage soll in Kapitel 4.4 behandelt werden.

Die hier dargestellte modulare Konzeption der Sprachproduktion macht an sich keine Aussagen darüber, ab welchem Erwerbszeitpunkt welche Informationen miteinander assoziiert sind. Im Gegenteil macht sie die Vorhersage, dass domänenspezifische Kenntnisse zeitlich unabhängig voneinander zur Verfügung stehen. Man kann also sagen, dass im ungestörten Spracherwerb eine quasi regelhafte Kontingenz von separaten Strukturweiterungen zu beobachten ist, wobei unterschiedliche Reifungsgeschwindigkeiten der submodularen Komponenten das implizieren, was als intra- und interindividuelle

Variation verstanden werden kann. In grober Abfolge lassen sich diese Integrationsstufen der sprachlichen Submodule allerdings mithilfe der Parallelen Architektur skizzieren.

3.4.2 Die Protowortproduktion als Indikator informationeller Integrationsstufen

Nach Maßgabe der Parallelen Architektur (vgl. Kap. 1.6) lässt sich die Entwicklung der Protowortproduktion noch etwas übersichtlicher illustrieren, als es im vorherigen Kapitel geschehen ist. Es muss betont werden, dass in den Arbeiten von Jackendoff der Spracherwerb nicht ausdrücklich im Fokus steht, aber vor allem in Besprechungen zur evolutionären Entwicklung des Sprachsystems Parallelen zu diesem gezogen werden. Zudem konzipiert er Produktion und Perzeption aller sprachlicher Strukturen ausdrücklich als arbeitsgedächtnisbasiert. Legt man die Modellierung der Parallelen Architektur der Untersuchung früher Wortproduktionen zugrunde, so lassen sich zum einen die divergenten Definitionen und Ergebnisse früherer Studien klären und ordnen, zum anderen erhält man eine Klassifikation von Protowörtern nach Maßgabe der Integration von perzeptiv schon vorhandenen Informationen. In diesem Sinne soll hier ebenfalls auf die Strong Identity Hypothesis von Hale & Reiss (2008) rekurriert werden und der Erwerb der Sprachproduktion als ein Prozess verstanden werden, in dem vollständig gespeicherte Informationen aufgrund von Performanzfaktoren nur sukzessiv integriert werden können. Hieraus ergibt sich ein Muster von Integrationsstufen, die sich nicht immer klar voneinander abgrenzen lassen, sondern graduelle Übergänge, insbesondere hinsichtlich der Informationen innerhalb der einzelnen Submodule, angenommen werden müssen. Insofern bilden die Stufen natürlich nur prototypische Abstraktionen zum Zwecke einer allgemeinen Klassifikation.

Insgesamt sollen hier drei Stufen der Integration angenommen werden. In den Analysen von Halliday (1975; 1979) und Piske (2001) werden Formen als Protowörter klassifiziert, die zum Teil die Phonotaktik der Zielsprache verletzen, zum Teil Phone beinhalten, die in der Zielsprache als Phoneme nicht vorkommen. Zum Beispiel produziert das Kind LEO [ŋgⁿgⁿŋ] (zeigen) im Alter von 1;03,17 (ebd.). Es wurde auch schon darauf hingewiesen, dass nach Ferguson (1978: 280) Protowörter u.a. aus einem einfachen oder wiederholten Vokal, silbischen Nasalen oder Frikativen bestehen können. In Kapitel 3.2.1 wurden solche Erscheinungsvarianten von Protowörtern auf die Integration des Merkmals [±silbisch] als 1. Integrationsstufe zurückgeführt, während alle weitere Information der Folgestufen

unsystematisch assoziiert wird. Genau deshalb kann es zu Verletzungen der Phonotaktik kommen, bei kontinuierlicher Produktion eines Silbengipfels.

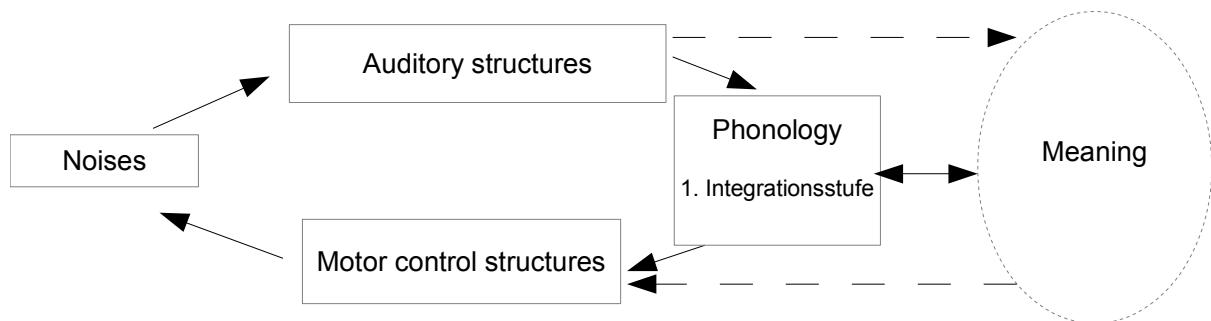


Abbildung 57: Etablierte Produktionspfade von Protowörtern der 1. Integrationsstufe.

Was wir also auf einer solch frühen Stufe nicht vorfinden, ist die Integration von zielsprachlichen Phonemen in zielsprachliche Silben. Solche Formen sollen hier als Protowörter der Integrationsstufe 1 (PWIS-1) bezeichnet werden. Diese können natürlich auch in späteren Spracherwerbsstufen noch erhalten bleiben (sowie auch Erwachsene diese Route noch nutzen können), allerdings muss man sie klar von späteren Protowortformen unterscheiden.

Grundsätzlich zu unterscheiden von Protowörtern der Integrationsstufe 1 sind Protowörter, die zusätzlich Information der Integrationsstufe 2 realisieren (PWIS-2). Sie unterscheiden sich genau dadurch, dass sie aus zielsprachlichen Phonemen mit einer zielsprachlichen Phonotaktik aufgebaut sind, ohne natürlich in ihrem segmentalem Gehalt einem Zielwort entsprechen zu müssen. Ein Beispiel wäre hierfür ['n̩n̩n̩] ((Auto-)fahren), das LEO im Alter von 2;00,25 produzierte (Piske 2001). Hier erkennt man klar eine reduplizierte CV-Struktur mit trochäischer Betonung und einem zielsprachlichen Nasal in der jeweiligen Onset-Position der Silben. Innerhalb der phonologischen Komponente hat also eine Etablierung der Oberklassenmerkmale stattgefunden, an den grundlegenden Produktionspfaden hat sich allerdings nichts geändert (vgl. Abb. 58).

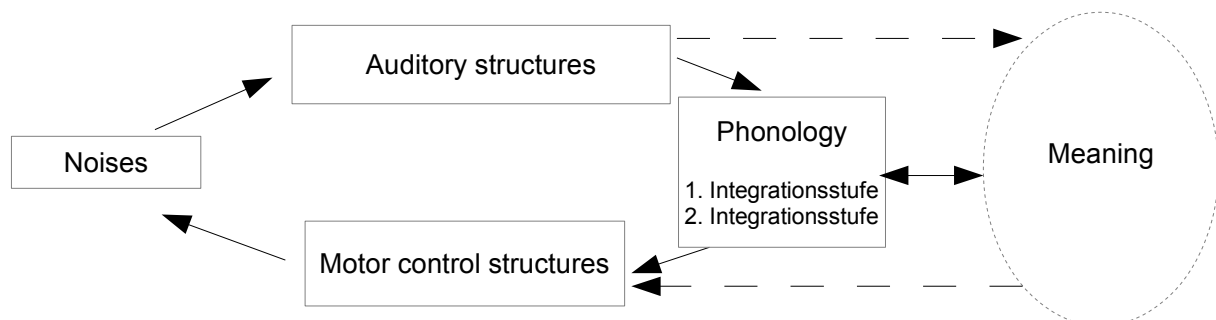


Abbildung 58: Etablierte Produktionspfade von Protowörtern der 2. Integrationsstufe.

Mit dem Fall P. liegt ein dokumentierter Fall einer verlängerten Protowortproduktion mit integrierter morphosyntaktischer Information vor. Die letzte Klasse sind entsprechend Protowörter, mit deren phonologischer Information auf Integrationsstufe 2 morphosyntaktische Informationen⁶⁶ assoziiert sind (PWIS-3) und die phonologische Information der Integrationsstufe 3 noch nicht oder noch nicht vollständig etabliert sind (vgl. Abb. 59). M.W. existiert kein dokumentierter Fall, in dem eine Assoziation von phonologischer Information der Integrationsstufe 1 mit morphosyntaktischer Information gegeben wäre. Dies wäre theoretisch durchaus denkbar, aufgrund der fehlenden Evidenz sollen hier aber nur die drei durch Beobachtung bestätigten Protoworttypen angeführt werden.

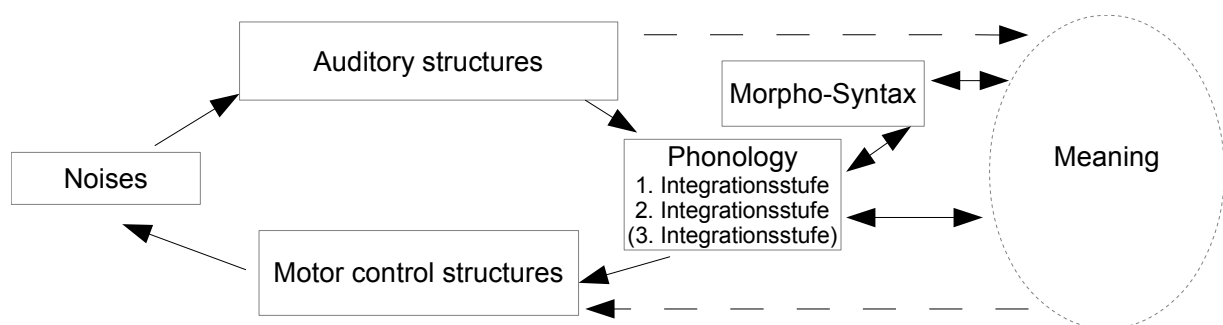


Abbildung 59: Etablierte Produktionspfade von Protowörtern mit morphosyntaktischer Information.

Mithilfe der oben ausgeführten Klassifikationsschemata ist es uns nun also möglich, die divergenten Daten systematisch zu ordnen und eine detaillierte Bestimmung dessen zu leisten, was bisher unter Protowortproduktion subsumiert wurde. Tabellarisch zusammengefasst lässt sich der chronologische Entwicklungsverlauf der Protowortproduktion jetzt wie folgt darstellen:

	Motorische Struktur	Phonologische Struktur	Morphosyntaktische Struktur	Bedeutung
PWIS-1	+	±silbisch	-	+
PWIS-2	+	(±silbisch) ± sonorant ± approximant ± vocoid	-	+
PWIS-3	+	±silbisch ± sonorant ± approximant ± vocoid	+	+

Tabelle 25: Klassifikationsschema der Protowortproduktion.

⁶⁶ Jackendoff (2002) unterscheidet noch einmal zwischen Phrasaler Syntax, Morphosyntax und einer GF(Grammatical Function)-Struktur. Zur einfacheren Übersicht sind diese Teilsysteme hier unter dem Etikett Morpho-Syntax zusammengefasst.

Selbstverständlich bleibt eine solche Darstellung vorläufig deskriptiv. Auch der Zielworterwerb in der Produktion vollzieht sich gemäß der beschriebenen Entwicklungslinie. Allerdings ist in der Protowortproduktion das gleichbleibende Moment die fehlende overte Ähnlichkeit mit Zielwörtern. Wie oben schon erläutert wurde, ist dies in einer fehlenden Integration phonologischer Information begründet. Die zugänglichen phonologischen Repräsentationen bleiben nach dieser Annahme auf phonotaktisch irreguläre Formen (PWIS-1) oder Demisilben mit geringer Komplexität (PWIS-2 und PWIS-3) beschränkt.

In Tabelle 26 findet sich eine Zuordnung von dokumentierten Protowörtern nach Maßgabe des hier vorgeschlagenen Klassifikationssystems. Die vorgenommene Zuordnung kann an dieser Stelle nur annäherungsweise durchgeführt werden, da es sich einerseits um Angaben aus Tagebuchstudien handelt (Elsen 1991; Leopold 1954), andererseits die Äußerungen der Kinder eng transkribiert wurden, da der Fokus auf der Untersuchung artikulatorischer Muster gelegt wurde (Piske 2001; Vihman & Miller 1988). Viele Wörter von Elsen (1991) wurden in die Liste nicht übernommen, da sie als onomatopoetische Wörter keine Protowörter darstellen, wie sie hier in der Arbeit verstanden werden sollen. Anders gesagt: Onomatopoetika sollen in der vorliegenden Arbeit als Zielwörter eingeordnet werden.

Protowort	Bedeutung	Alter	Quelle + Kind	Klassifikation
[ʔəʔ]	Demonstrativ, Interjektion	0;8-1;6	Leopold (1954) Hildegard	PWIS-1
[kχ]	schlecht schmecken	1;7, 1;10	Leopold (1954) Hildegard	PWIS-1
[næŋnæŋæ:] [næ'ŋa] [ŋ:aŋ:]	los/schneller	1;00,00	Piske (2001) LEO	PWIS-1
[vaʋ̃], [wæɔ'wə], [ʔəwa:v:ä]	drehen	1;00,14	Piske (2001) LEO	PWIS-1
[g'nŋ:], [ŋg'nŋŋ], ['g'nŋg'nŋ]	zeigen	1;03,17	Piske (2001) LEO	PWIS-1
[hæ'G'nŋ], [G'nŋ], [ʔɛ' G'nŋ]	ich	1;11,26	Piske (2001) LEO	PWIS-1
[G'nŋ], [ʔa.g'nŋ]	du	2;01,01	Piske (2001) LEO	PWIS-1
[pwi]	Reaktion auf unbekannten Besucher, salientes Geräusch, Mutter kommt zurück	0;9,17 bis 0;11,11	Vihman & Miller (1988) Deborah	PWIS-1
['ɛɪ], [ba.bai]	Wasser	0;11,24	Elsen (1991) Annalena	PWIS-2
[ʔaʔa]	Aufmerksamkeit, Rufen zu Hunden	1;0 – 1;2	Leopold (1954) Hildegard	PWIS-2
[bu::]	Donnern	1;11	Leopold (1954) Hildegard	PWIS-2
[dɪdɪdɪ]	beschimpfen, trösten	0;9-0;10	Leopold (1954) Hildegard	PWIS-2
[nenene]	beschimpfen	0;10	Leopold (1954) Hildegard	PWIS-2
[b̥əb̥ə]	Keks	1;04,00	Piske (2001) LEO	PWIS-2
[b̥'ɪɪ], [b̥ɪɪ],[b̥ɪɪb̥ɪɪ]	Räder	2;02,06	Piske (2001) LEO	PWIS-2
[b̥əb̥əb̥ə], [b̥əb̥ə], ['b̥əb̥ə]	Schnuller	1;02,21	Piske (2001) LEO	PWIS-2
[wɪwɪ]	waschen	2;01,22	Piske (2001) LEO	PWIS-2
([nm]), ['n̥an̥a], ['n̥ən̥ən̥ə]	(Auto)fahren	2;00,25	Piske (2001) LEO	PWIS-2
[h̥a ^h ̥a], [h̥ä ^h ̥ä],[h̥ə ^h ̥ə]	Zustimmung	1;06,01	Piske (2001) LEO	PWIS-2
[h̥o ^h ̥]	Reaktion auf Hundebellen, Telefonklingeln, aufgehängter	0;9-1;0	Vihman & Miller (1988) Jonah	PWIS-2

	Pinguin schlägt mit Flügeln, Untersucher			
[dada]	Reaktion auf Hundenamen, Hundebellen, am Fenster mit Blick auf Hundeplatz, Mutter baut Turm mit Bausteinen, krabbeln zu grünem Licht von der Videoanlage	keine Angaben	Vihman & Miller (1988) Jonah	PWIS-2
[oʊ]	Aufmerksamkeit auf neues Ereignis, ab 11 Monate mit Zeigen	0;9-1;0	Vihman & Miller (1988) Emily	PWIS-2
[ʔʌgæ]	mit Zeigegeste; erst Bilder, dann alle Objekte für die Namen unbekannt (z.B. Waal, Schildkröte)	0;11-1;1	Vihman & Miller (1988) Timmy	PWIS-2
[hə] + Nicken	Antwort auf Frage „Willst du, dass ich es mache?“, anschließend generalisiert zum Ausdruck für Wunsch für Handlung	Ab 1;2 (für zwei Wochen)	Vihman & Miller (1988) Timmy	PWIS-2

Tabelle 26: Klassifikation von dokumentierten Protowörtern nach Maßgabe der Parallelen Architektur.

In welcher qualitativen und quantitativen Relation die hier angeführten Protowörter zu den Zielwörtern der jeweiligen Kinder stehen, kann für keine der vorliegenden Studien genau bestimmt werden. Ob die in diesem Kapitel angeführten theoretischen Erwägungen und Modellvorstellungen stichhaltig sind, soll deshalb im Folgenden anhand eigener empirischer Daten untersucht werden.

4 Die empirische Untersuchung der Protowortproduktion

4.1 Methode

4.1.1 Zielsetzungen und Ausgangshypothesen

Die nicht bestimmbare Auftretenswahrscheinlichkeit von Protowörtern macht ein besonderes empirisches Design zu ihrer adäquaten Erfassung notwendig. Durch die Befunde der in

Kapitel 3.1 vorgestellten Untersuchungen zu Protowörtern kann man folgende methodenrelevante Aspekte zu Protowortproduktionen festhalten:

1. Nicht jedes Kind zeigt eine Protowortphase.
2. Treten Protowörter im ungestörten Spracherwerb auf, so ist der Produktionszeitraum von relativ kurzer Dauer.
3. Protowörter zeigen phonologisch keinen Bezug zu Zielwörtern.
4. Protowörter werden von unterschiedlichen Kindern jeweils unterschiedlich produziert, wir finden also inter- wie intraindividuelle Variation.

Aspekt 1 wirft ein Problem auf, das in der kognitiven Neuropsychologie im Hinblick auf die Einteilung von Patientengruppen in Syndromgruppen diskutiert wurde (vgl. Caramazza 1986; Caramazza & Badecker 1989; Ellis 1987; Ellis & Young 1991; Shallice 1979). In dieser Diskussion ging es um eine Form der Patientenkategorisierung in Gruppenstudien, in der häufig auftretende Symptome bestimmten Syndromgruppen, wie z.B. der Broca- oder Wernicke-Aphasie, zugeordnet werden. Die grundsätzliche Frage, die der gesamten Diskussion zugrunde lag, war und ist, inwiefern es mit Hilfe einer homogenen Kategorisierung und mit ihr den homogenisierten Leistungen von ganzen Patientengruppen möglich sei, relevante Informationen über kognitive Verarbeitungsprozesse zu erhalten. In den 80'er Jahren des letzten Jahrhunderts wurde diese Frage in der kognitiven Neuropsychologie grundsätzlich verneint und eine Menge beachtlicher und wissenschaftlich maßgebender Einzelfallstudien durchgeführt. Die traditionelle Neuropsychologie schien überholt, Ellis (1987: 405) sprach von den „bad old days of comparing twenty 'aphasics' with twenty 'controls'“. Die „bad old days“ scheinen heute in der Aphasieologie wie auch in der Spracherwerbsforschung wiedergekehrt. Aus diesem Grunde ist es angebracht, an die ertragreiche methodische Diskussion, die vor allem in Caramazza (1986) stattgefunden hat, anzuschließen und diese ausführlicher darzustellen.

4.1.2 Methodische Begründung des Untersuchungsverfahrens: Einzelfallstudien

Ausgangspunkt für die Kritik von Caramazza (ebd.) an Gruppenstudiendesigns ist die schlichte Feststellung, dass die Wertung von Daten als relevante Evidenz in kritischem Maße von dem Erklärungsmodell abhängt, das einer Untersuchung zugrunde gelegt wird. Für eine solche Bestimmung der Relevanz muss das jeweilige Modell, für das bestimmte Beobachtungen kritisch sein soll, die Beziehungen zwischen den Daten und postulierten

Verarbeitungskomponenten in hinreichender Genauigkeit darstellen können. Die komplexe Beziehung von Evidenz und Theoriebildung fasst Caramazza in folgender Formalisierung:

„A set of observations E_i obtained under initial condition $c_1, c_2, c_3 \dots c_n$, constitutes relevant evidence in support of hypothesis H and related auxiliary hypotheses $a_1, a_2 \dots a_n$, in model M provided that E_i can be derived computationally from H in M and the specified initial conditions.“

Das Beispiel, das Caramazza zur Verdeutlichung gibt (ebd.: 48; vgl. auch Abb. 60), ist ein Fall, in dem ein Set aus Beobachtungen zur Reaktionszeitmessung gegeben ist (E_i), die aus anfänglichen Bedingungen (c_1, \dots, c_n) eines experimentell manipulierten Faktors (morphologisch einfache und morphologisch komplexe Wörter) hervorgingen. Die Hypothese (H) könnte sich auf das graphematische Input-Lexikon beziehen (z.B., dass morphologische Dekomposition für den Zugang desselben notwendig ist), während Zusatzannahmen bestehen (a_1, \dots, a_n), durch welche eine Relation zwischen H und den Reaktionszeitmessungen hergestellt wird (z.B., dass Reaktionszeiten für morphologisch komplexere Wörter länger als für morphologisch einfache Wörter sind, wenn alle anderen Faktoren kontrolliert wurden). M ist eine ausgearbeitete funktionale Architektur, von der das graphematische Input-Lexikon eine Komponente darstellt.⁶⁷ Eine wesentliche Annahme für M ist, dass für sie eine universelle Gültigkeit beansprucht wird, d.h., M gilt für alle ungestörten menschlichen mentale Systeme/Gehirne.

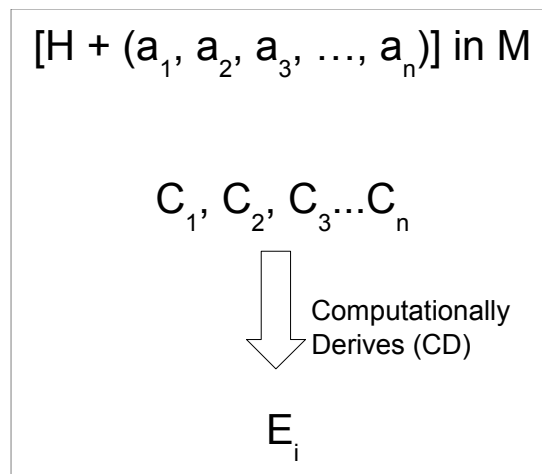


Abbildung 60: Schema des Zusammenhangs von Theorie und Evidenz.

⁶⁷ H und a_1, \dots, a_n werden im weiteren Verlauf der Darstellung zur besseren Übersicht unter M subsumiert.

Die Frage, die sich nun zu Untersuchungen bei gestörten menschlichen mentalen Systemen/Gehirnen stellt, ist natürlich, welche Schlussfolgerungen aus den vorliegenden Störungen gezogen werden können. Nimmt man eine nicht-transparente Beziehung zwischen ungestörten und gestörten kognitiven Systemen an, so muss eine Störung des menschlichen mentalen Systems/Gehirns zu einer Neubildung kognitiver Operationen führen (M'), die zu M in keiner ableitbaren Relation stehen. Gegen diesen Standpunkt macht Caramazza die Transparenzhypothese stark, in der M , also das kognitive System als Ganzes, sowohl in gesunden als auch in hirngeschädigten Personen fundamental gleich aufgebaut ist. Was bei letzteren lediglich auftritt, ist eine lokale Modifikation des kognitiven Systems (L_i), so dass die Performanz eines Patienten über L_i direkt zu M in Beziehung gesetzt werden kann und folgende, modifizierte Formalisierung der Beziehung von Evidenz und Theoriebildung gegeben werden kann (ebd.: 51):

„A set of observations E'_i that is, the cognitive performance of a brain-damaged patient, is evidence for M just in case it is possible to computationally derive E'_i from M , L_i and initial condition c_1 through c_n , where L_i is a complex hypothesis about the locus of damage to a functional architecture and associated assumptions concerning the effects of the hypothesized damage on the cognitive system.“

Insofern kann auch L_i über M , E'_i und c_1, \dots, c_n nur a posteriori abgeleitet werden und in dem Maße wie eine Spezifikation von L_i bei gegebenem E'_i erfolgreich gelingt, wird M bestätigt (vgl. Abb. 61).

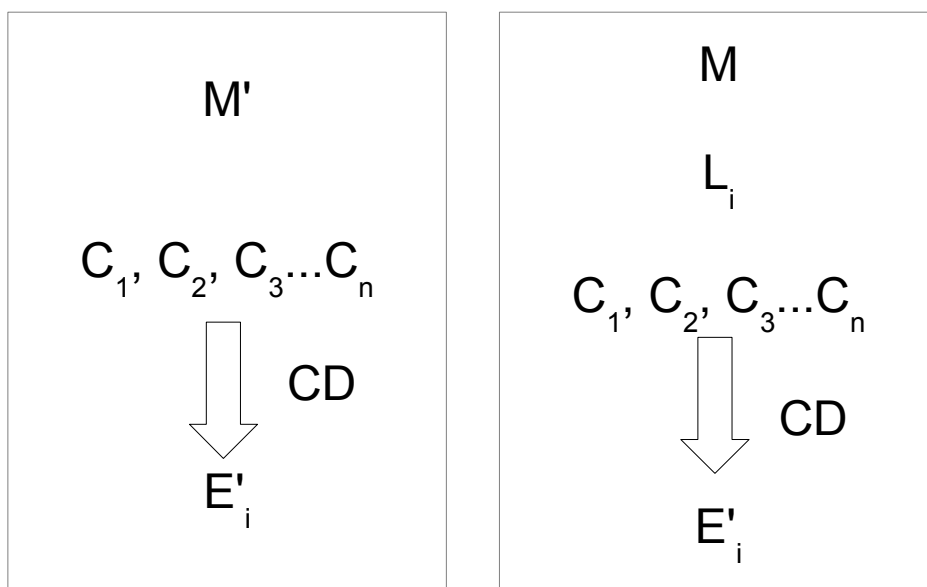


Abbildung 61: Schema des Zusammenhangs von Theorie und Evidenz mit lokaler Modifikation L_i des kognitiven Systems (rechts).

Ein solche Ableitung kann aber nur aus Einzelfallstudien gelingen. Da in Gruppenstudien mit ungestörten Probanden M als universell gegeben ist und die Bedingungen c_1, \dots, c_n konstant gehalten werden, können individuelle Leistungsprofile $E_1 \dots E_n$ als homogen angesehen und statistisch problemlos zu E'_i gemittelt werden. Der unterscheidende Faktor bei Patienten mit Hirnschädigung ist natürlich L_i . Da dieses durch ein gegebenes E'_i erst im Nachhinein erschlossen werden kann, bzw. E'_i durch L_i bedingt ist, kann man keine für Gruppenstudien notwendige Homogenität postulieren. Dies wäre nur möglich, wenn man für alle Patienten die Identität von L postulieren würde ($L_1 = L_2 = \dots = L_n$), was durch die notwendige a posteriori Ermittlung von L_i durch E'_i theoretisch absurd ist. Dementsprechend sind nach Caramazza nur (multiple) Einzelfallstudien für die Untersuchung hirngeschädigter Personen methodisch valide.

Nach Caramazza (1984) werden demnach Einzelfallstudien im optimalen Falle im Rahmen des kognitiv-linguistischen Paradigmas durchgeführt, das durch zwei wesentliche methodologische Prämissen charakterisiert ist: Modularität/Selektivität und Transparenz.

Die Modularitätskonzeption, die Caramazza verwendet, sieht das Sprachsystem aus mehreren Komponenten zusammengesetzt (vgl. Kap. 1.2.1). Komplexe kognitive Funktionen, wie z.B. das Satzverständnis, bestehen wiederum aus verschiedenen Verarbeitungseinheiten. Allgemein kann so eine komplexe kognitive Funktion X durch das Zusammenspiel der Komponenten a , b , c und d erklärt werden. Fällt nun z.B. die Komponente d aufgrund einer Hirnschädigung selektiv aus, so besagt die Transparenzannahme, dass aus den Leistungen des Patienten Rückschlüsse auf das (als uniform postulierte) ungestörte kognitive System gezogen werden können bzw. unter Zugrundelegung des ungestörten kognitiven Systems eine lokale Modifikation im Sinne einer gestörten Teilkomponente identifiziert werden kann (ebd.: 10). Als Nullhypothese des kognitiv-linguistischen Paradigmas schließlich gilt, dass sprachliche Auffälligkeiten so lange als ungestört angesehen werden, wenn nicht eine Alternativhypothese durch Evidenz bestätigt und so das Gegenteil belegt wird.

Inwiefern kann man die methodischen Argumente von Caramazza auf die Untersuchung des frühen Spracherwerbs übertragen? Wie in Kapitel 1.2.2 dargestellt, gibt es keine hinreichenden Gründe dafür, eine modulare Komposition des kognitiven Systems nicht auch für den frühen Spracherwerb anzunehmen. Alle Faktoren eines Untersuchungssettings (M , E'_i und c_1, \dots, c_n) können als analog angesehen werden. So ist z.B. der Zufallsgenerator (vgl. Kap.

1.5.2) als Teil von M, wenn er nicht bei Aphasikern als eine Ad-hoc-Installation postuliert wird, auch für das sich ausbauende Sprachproduktionssystem anzunehmen.⁶⁸

Der entscheidende Faktor, der unter dieser analogen Sichtweise für Einzelfallstudien im frühen Spracherwerb sprechen würde, wäre entsprechend obiger Ausführungen L_i . Wie im Zuge dieser Arbeit mehrfach ausgeführt, ist es möglich, L_i für den Spracherwerb analog zu konzipieren. Die lokalen Modifikationen des kognitiven Systems bilden dementsprechend hier nicht Schädigungen, sondern differente Reifungszustände der Submodule. Da die Variabilität solcher intra-modularen Reifungszustände im Spracherwerb den Normalfall darstellt, Homogenität also nicht postuliert werden kann, sind Gruppenstudien im frühen Spracherwerb ebenso abzulehnen wie in der Aphasieforschung. Dass Protowörter als genuin individuelle Produktionsformen selbstverständlich durch standardisierte Testungen wie Elternfragebögen nicht erfasst werden können, stellt hierfür wohl ein deutliches Beispiel dar. Eine Komponente wie der Zufallsgenerator würde dementsprechend in Gruppenstudien sowohl in einer Überkategorie Wernicke-Aphasiker als auch in der Überkategorie „Normkind“ schlicht statistisch ausgemittelt werden. In diesem Sinne soll die Definition von Caramazza für die Rolle der Evidenz aphasischer Daten im Hinblick auf das kognitive System für den frühen Spracherwerb in folgender Form analog übernommen werden:

Eine Menge von Beobachtungen E'_i , d.h. die kognitive Performanz eines Kindes im Spracherwerb, bildet Evidenz für M im Falle dessen, dass es möglich ist, E'_i von M, L_i und anfänglichen Bedingungen c_1 zu c_n verarbeitungsbasiert abzuleiten, wobei L_i eine komplexe Hypothese über den Reifungszustand einer funktionalen Architektur und mit dieser verbundene Annahmen über die Effekte dieses hypothetisierten Reifungszustandes darstellt.

Anzumerken ist hierbei, dass durchaus Parallelen zur Regressionshypothese von Jakobson (1941) gegeben sind. Mit Caramazza (1994: 121) kann man bei dieser zwei unabhängige Annahmen unterscheiden. Zum einen die Formbeschränkung, die besagt, dass es Übereinstimmungen gibt zwischen den zugänglichen sprachlichen Formen nach Hirnschädigung und solchen Formen, die auf verschiedenen Stufen des Spracherwerbs auftreten. Zum anderen die Folgebeschränkung, nach der die zeitlich später erworbenen Formen nach einer Schädigung auch anfälliger sind für Störungen. Dass beide Beschränkungen in ihrer starken Variante nicht haltbar sind, wurde vielfach durch Evidenz belegt (vgl. z.B. Caramazza & Zurif 1978). Dies würde allerdings auch einen zeitlich uniformen Verlauf über die gesamte funktionale Architektur des Sprachsystems hinweg

⁶⁸ Auch Caramazza (1994: 126) verweist auf die Uniformität von M in Bezug auf den Spracherwerb.

postulieren, was bei der gegebenen Variabilität von Spracherwerbsverläufen des intra- und intermodularen Ausbaus nicht angenommen werden kann. Selbst wenn die funktionale Architektur uniform ist, so ist es ihr Ausbau mitnichten. Eine Bestätigung der Folgebeschränkung wäre demnach m.E. nur möglich, wenn der individuelle Erwerbsverlauf eines Patienten mit seinen sprachlichen Formen nach einer späteren Hirnschädigung korreliert werden könnte. Es ist wohl unnötig zu betonen, dass eine solche Langzeitstudie nicht durchführbar ist. Ein Zurückschalten auf frühere Programmstufen, d.h. auf repräsentationale Optionen, die im Spracherwerb zur Verfügung standen, ist dennoch eine denkbare, wenn auch nicht notwendige Variante.⁶⁹ In diesem Sinne wäre die Hypothese, dass der Zufallsgenerator und mit ihm die Produktion nicht-komplexer Demisilbenformen sowohl bei einer Untergruppe von Aphasikern als auch bei einer Untergruppe von Kindern im frühen Spracherwerb aktiv ist, eine schwache Form der Regressionshypothese, die hier auch vertreten werden soll.

⁶⁹ Fälle des Fremdsprachenakzentsyndroms (Blumstein et al. 1987, Blumstein & Kurowski 2006) deuten m.E. sogar darauf hin, dass selbst phonologische Optionen, die im Spracherwerb zur Verfügung standen und nicht realisiert wurden, nach einer Hirnschädigung immer noch zur Verfügung stehen und aktiviert werden können. Bisher ist diese Möglichkeit m.W. allerdings noch nicht untersucht worden (für einen Überblick, s. Moen 2000).

4.1.3 Methodische Begründung des Untersuchungsverfahrens: Spontansprachuntersuchungen

Protowörter treten im ungestörten Spracherwerb in einem relativ kurzen Zeitfenster auf. Dies macht zum einen Einzelfallstudien notwendig, da in einer, durch Gruppenstudien gewonnenen, standardisierten Testung die individuelle phonologische Form in einem individuellen Kontext nicht erfasst werden kann. Zum anderen ergibt sich durch ihn zwangsläufig die Frage nach der adäquaten Methode zur Erfassung des (Proto-)Lexikons. Kauschke (2000) hat unter diesem Aspekt die Vor- und Nachteile von Fragebögen bzw. Checklisten und Spontansprachuntersuchungen zusammengefasst (vgl. Tab. 26).

Fragebogen / Checkliste		Spontansprachuntersuchung	
Vorteile	Nachteile	Vorteile	Nachteile
Direkt aus dem natürlichen Umfeld des Kindes entnommen (Bezugspersonen)	Keine Ermittlung, wie häufig bestimmte Wörter und Wortarten verwendet werden	Ermittlung, wie häufig bestimmte Wörter und Wortarten verwendet werden, ist möglich	Unökonomisches Verfahren aufgrund der Beobachtungszeit und Auswertung
Ökonomisches Verfahren – große Fallzahlen möglich	Kontext des Gebrauchs bleibt offen	Kontext des Gebrauchs ist bestimmbar	Es wird nur ein selektiver Ausschnitt lexikalischen Wissens erfasst
	Keine Rückschlüsse auf phonologische Entwicklung möglich	Phonologische Entwicklung kann anhand von Transkripten erschlossen werden	Es können keine Aussagen über rezeptive Sprachleistungen gemacht werden
	Wortartenverteilung richtet sich an der vorgegebenen Verteilung der Checkliste aus	Tatsächliche Verwendung von Wörtern wird erfasst	
	Fehlerquote der Eltern kann nicht ermittelt werden	Verschiedene sprachliche Parameter können analysiert werden	
	Wortsammlung ist vorgegebenen, Bezugspersonen richten sich meist nach Vorgabe	Input kann zusätzlich analysiert werden (Sprache der Bezugspersonen)	

Tabelle 27: Gegenüberstellung der Vor- und Nachteile von Checkliste-/Fragebogen-Methode und Spontansprachuntersuchung nach Kauschke (2000: 77-78).

Anhand der Zusammenfassung wird klar deutlich, dass für eine Untersuchung der frühen phonologischen Entwicklung eine Spontansprachstudie die angemessene methodische Variante ist. Da diese es ermöglicht, auch andere sprachliche Parameter in die Analyse einzubeziehen sowie den Kontext einer Äußerung genau zu bestimmen, lassen sich zudem

Aussagen darüber machen, mit welchem konzeptuellen Gehalt Protowörter produziert werden und inwiefern andere Teilkomponenten des Sprachproduktionssystems bereits aktiv sind.

Eine wichtige Ergänzung zu Spontansprachuntersuchungen können Tagebuchstudien darstellen. Die Einschränkungen durch eine vorgegebene Normliste sind bei dieser Form der Datensammlung nicht gegeben. Die Stärke dieser Erhebungsmethode besteht vor allem in der Detailschärfe durch die kontinuierliche und langfristige Dokumentation, so dass auch kurzzeitig auftretende Muster erfasst werden (vgl. Kauschke 2000; Pomnitz & Siegmüller 2013). Der Nachteil, dass sehr wenige Kinder erfasst werden können (Kauschke 2000: 75), wird unter Berücksichtigung obiger Argumentation und dem Ziel, Protowörter zu identifizieren, hier nicht als solcher gewertet. Detailschärfe wird selbstverständlich nur durch Detailkenntnis garantiert, d.h., spezifische Phänomene, wie hier die Protowortproduktion, werden den linguistisch ungebildeten Eltern mit hoher Wahrscheinlichkeit entgehen. Auch wenn Elternangaben für die Untersuchung des zielsprachlichen Lexikonerwerbs reliable und valide Daten liefern (vgl. z.B. Bates et al. 1988), ist anzunehmen, dass im Schnitt einer konventionellen Definition von Wörtern gefolgt wird. In diesem Sinne werden in den Untersuchungen auch die Instruktionen für die Eltern unter Zugrundelegung eines konventionellen Wortbegriffs konzipiert. Dass derart wichtige Informationen über individuelle Formen verloren gehen, wiegt m.E. schwerer, als die Bedenken von Pomnitz & Siegmüller (ebd.: 53), dass durch die vorausgesetzte kontinuierliche Mitarbeit der Bezugspersonen ein Überhang von Eltern mit höherem sozio-ökonomischen Status entstehen könne. Diese Bedenken sind für eine phonologische Untersuchung des frühen Spracherwerbs insofern nicht relevant, als dass in einer Studie von Dodd et al. (2003) mit 684 Englisch sprechenden Kindern im Alter von 3;0 bis 6;11 Jahren gezeigt wurde, dass der sozio-ökonomische Status keinen Einfluss auf die phonologische Entwicklung hat (und demnach auch im früheren Erwerbssalter nicht vorhanden ist; vgl. auch Smit et al. 1990). Was das Geschlecht der untersuchten Kinder angeht, so fanden Dodd et al. (2003) hier bis zum Alter von 5;6 Jahren keine Einflüsse, so dass dieser Faktor für eine phonologische Analyse früher Sprachproduktionen ebenfalls zu vernachlässigen ist.

Zusammenfassend lässt sich festhalten, dass für eine phonologische Analyse früher Wortproduktionen die Methode einer Spontansprachuntersuchung die adäquate Methode darstellt. Im Sinne der Ausführungen in Kapitel 4.1.2 ist eine Einzelfallstudie hinreichend und zielführend, um Rückschlüsse auf die zugrundeliegende kognitive Architektur zu ziehen.

4.2 Durchführung

4.2.1 Datenerhebung

Die Daten des sprachunauffälligen Kindes J. (männlich) wurden ab einem frühen Entwicklungsalter (ab 1;4 Jahre) zu Beginn seiner ersten Wortproduktionen aufgenommen. Die Frage des Aufnahmekontextes ist für Spontansprachstudien von keiner geringen Bedeutung und für die Protowortproduktion gilt noch stärker, dass zum einen eine vertraute Umgebung notwendig ist (da frühe konzeptuelle Strukturen an spezifische Ereignisse in der Umgebung gebunden sein können, vgl. auch Kapitel 3.2.5), zum anderen aber auch eine Umgebung gegeben sein muss, die genügend sprachliche Äußerungen evoziert. Ein enges Setting mit nur zwei Erwachsenen (Bezugsperson und Untersucher), in dem die Anwesenheit und zielsprachliche Interaktion mit dem Untersucher zu bedeutsam wird, ist für die Untersuchung der Protowortproduktion in diesem frühen Entwicklungsalter abzulehnen. Aus diesem Grunde wurden die Aufnahmen von J. in der Krabbelstube vorgenommen, die er bereits seit ca. vier Monaten täglich besuchte. Da für die Zwecke der phonologischen Analysen dieser Studie eine weite Transkription gewählt wurde, sind Störgeräusche bei der Analyse nicht ins Gewicht gefallen. Falls doch Überlagerungen stattfanden, wurde die jeweilige Produktion von der Analyse ausgenommen.

Die Aufnahmesitzungen fanden in der Regel einmal pro Woche für 30 Minuten und über einen Zeitraum von einem halben Jahr statt, wobei nach der 14. Aufnahmewoche keine Protowörter mehr produziert wurden und somit nur bis zu diesem Zeitpunkt ausgewertet wurde. Die Datenerhebung erfolgte mittels eines digitalen Aufnahmegeräts (Olympus DM-550 MP3-Rekorder), das während der Aufnahmen an einem Notizbrett befestigt war, so dass die Aufnahme-prozedur für das Kind nicht ersichtlich war und das Gerät auch nicht ablenkte oder irritierte.

Während der Aufnahmesitzungen vermied es der Untersucher, soweit dies möglich war, selbst Referenzen auf Objekte herzustellen oder Aufforderungen zum Sprechen zu geben, da dies die Produktion von Protowörtern eventuell eingeschränkt bis verhindert hätte.

4.2.2 Datenbearbeitung

Im Anschluss an jede Sitzung wurden die jeweiligen Äußerungen gemäß des Internationalen Phonetischen Alphabets transkribiert. Da in der vorliegenden Arbeit die systematisch

verwendeten Lautklassen und nicht die variierende Aussprache einzelner Laute im Vordergrund stehen, wurde eine weite Transkription als hinreichend erachtet. Alle in dieser Arbeit verwendeten phonetischen Transkriptionen wurden von einem zweiten Untersucher unabhängig transkribiert und auf ihre Reliabilität hin geprüft (Reliabilität: 89 %). Besonderheiten in der Zuordnung von Lauten zu Lautklassen müssen an dieser Stelle erwähnt werden, da sie für die statistische Analyse relevant sind.

- Der Glottisverschlusslaut [ʔ] wird als Plosiv gewertet (vgl. Hall 2000)
- [h] wird als stimmloser glottaler Frikativ gewertet (vgl. ebd.)
- Diphthonge werden als Kombinationen aus Vokal und Gleitlaut gewertet (Wurzel 1981; Vater 1992)
- Vokale werden nicht weiter differenziert (Clements 1990)

Die Äußerungen wurden schon während der Aufnahmen hinsichtlich ihres Anwendungskontextes (Angaben über Personen, Objekte und Handlungszusammenhang) klassifiziert. Hierzu wurden, wenn möglich, die von Halliday definierten semantischen Zuordnungen verwendet, wobei für diese die Spezifikationen aus Kapitel 3.2.5 gelten:

- 1) Instrumental: Ziel, ein Objekt zu bekommen
- 2) Regulatory: Ziel, eine Person zu einer Handlung aufzufordern
- 3) Interactional: Ziel, eine Interaktion mit einer Person herzustellen
- 4) Personal: Ziel, eigene innere Teilhabe an einer Situation kund zu tun.

Produktionen, wie zufällige, nicht erneut auftretende Lallabfolgen, direkte Imitationen oder reine Affektkundgebungen ohne segmentale Struktur wurden als nicht klassifizierbar verzeichnet.

Alle klassifizierbaren Äußerungen wurden sowohl phonologisch im Rahmen der 3D-Phonologie (qualitative Auswertung) als auch mittels statistischer Verfahren (quantitative Auswertung) analysiert. Durch diese sich ergänzenden Perspektiven auf die Daten sollte ein möglichst umfassendes Spektrum an Ergebnissen für die anschließende Interpretation gewährleistet werden.

Da in der quantitativen Analyse dieser Studie unter anderem untersucht werden soll, ob eine Korrelation zwischen der Protowortproduktion und der Produktion von Zielwörtern existiert, konnte hier nicht auf etablierte Messwerte zurückgegriffen werden, die eine uniforme, zielsprachliche Klassifikation von Wörtern innerhalb eines definierten Datensatzes zugrunde legen. Sowohl der Wert „number of different words (NDW)“, der alle unterschiedlichen Wörter angibt, der Wert „total number of words (TNW)“, der alle produzierten Wörter umfasst, als auch die „Type-Token-Ratio (TTR)“, bei welcher der Quotient aus Types und

Token gebildet wird, fallen dementsprechend in ihrer klassischen Anwendung weg. Aus diesem Grunde soll hier die Wort-Nichtwort-Ratio (W:NW-Ratio) als Maß des lexikalischen Fortschritts verwendet werden. Die Anzahl der zielsprachlichen Wörter wird hierbei durch die Anzahl der Nichtwörter dividiert (in der Studie sollten mindestens zwei Wörter und Nichtwörter in einer Sitzung gegeben sein), so dass mittels der Pearson Produkt-Moment-Korrelation ein möglicher Zusammenhang ermittelt werden kann.

Ausgenommen von dieser speziell zu begründenden Fortschrittsmessung kamen in der statistische Analyse die gängigen Verfahren für die gegebenen Datentypen zur Anwendung.

4.2.3 Spezifische Fragestellungen und Hypothesen

Es ergibt sich in dieser Arbeit eine Vielfalt von Untersuchungsfragen, für die spezifische Hypothesen gebildet werden müssen. Für einzelne Teilbereiche gibt es mehrere Teilfragestellungen. Ist dies der Fall, werden sie in Kapitel 4.3.2 in verkürzter Form angegeben, nämlich in der Reihenfolge

(Teilbereich (X)) + Nummer der Fragestellung (Y.) + Nummer der Hypothese (H_z). So z.B. mit (e) 2. H_0 . Im Folgenden werden die Fragestellungen und Hypothesen vollständig aufgelistet:

- (a) Fragestellung:
Gibt es einen Zusammenhang von Wort- und Nichtwortproduktion (Ziel- und Protowortproduktion)?
 H_0 = Es gibt keinen signifikanten Zusammenhang zwischen Wort- und Nichtwortproduktion.
 H_1 = Es gibt einen signifikanten Zusammenhang zwischen Wort- und Nichtwortproduktion ($\alpha=0,05$).
- (b) Fragestellung:
Werden Silbenstrukturtypen in Ziel- und Protowörtern unterschiedlich häufig produziert?
 H_0 = Silbenstrukturtypen werden in Ziel- und Protowörtern gleich häufig produziert.
 H_1 = Silbenstrukturtypen werden in Ziel- und Protowörtern bezogen auf die Auftretenshäufigkeit signifikant unterschiedlich produziert ($\alpha=0,05$).
- (c) Fragestellung:
Werden Reduplikationen in Ziel- und Protowörtern unterschiedlich häufig produziert?
 H_0 = Reduplikationen werden in Ziel- und Protowörtern gleich häufig produziert.
 H_1 = Reduplikationen werden in Ziel- und Protowörtern bezogen auf die Auftretenshäufigkeit signifikant unterschiedlich produziert ($\alpha=0,05$).
- (d) 1. Fragestellung:
Werden Demisilbentypen und damit weniger komplexe und komplexere Demisilben in der Produktion von Protowörtern initial unterschiedlich häufig produziert?
 H_0 = Demisilbentypen werden in Protowörtern initial gleich häufig produziert.

H_1 = Demisilbentypen werden in Protowörtern bezogen auf die Auftretenshäufigkeit initial signifikant unterschiedlich produziert ($\alpha=0,05$).

2. Fragestellung:

Werden Demisilbentypen und damit weniger komplexe und komplexere Demisilben in der Produktion von Protowörtern final unterschiedlich häufig produziert?

H_0 = Demisilbentypen werden in Protowörtern final gleich häufig produziert.

H_1 = Demisilbentypen werden in Protowörtern bezogen auf die Auftretenshäufigkeit final signifikant unterschiedlich produziert ($\alpha=0,05$).

(e) 1. Fragestellung:

Werden Demisilbentypen und damit weniger komplexe und komplexere Demisilben in der Produktion von Zielwörtern initial unterschiedlich häufig produziert?

H_0 = Demisilbentypen werden in Zielwörtern initial gleich häufig produziert.

H_1 = Demisilbentypen werden in Zielwörtern bezogen auf die Auftretenshäufigkeit initial signifikant unterschiedlich produziert ($\alpha=0,05$).

2. Fragestellung:

Werden Demisilbentypen und damit weniger komplexe und komplexere Demisilben in der Produktion von Zielwörtern final unterschiedlich häufig produziert?

H_0 = Demisilbentypen werden in Zielwörtern final gleich häufig produziert.

H_1 = Demisilbentypen werden in Zielwörtern bezogen auf die Auftretenshäufigkeit final signifikant unterschiedlich produziert ($\alpha=0,05$).

(f) 1. Fragestellung:

Ist die Produktion von initialen Demisilbentypen in Proto- und Zielwörtern unabhängig voneinander und besteht bei Unabhängigkeit der Produktion ein Zusammenhang?

H_0 = Die Produktion initialer Demisilben in Proto- und Zielwörtern ist nicht unabhängig voneinander und es existiert kein signifikanter Zusammenhang.

H_1 = Die Produktion initialer Demisilben ist unabhängig voneinander und es gibt einen signifikanten Zusammenhang ($\alpha=0,05$).

2. Fragestellung:

Ist die Produktion von finalen Demisilbentypen in Proto- und Zielwörtern unabhängig voneinander und besteht bei Unabhängigkeit der Produktion ein Zusammenhang?

H_0 = Die Produktion finaler Demisilben in Proto- und Zielwörtern ist nicht unabhängig voneinander und es existiert kein signifikanter Zusammenhang.

H_1 = Die Produktion finaler Demisilben ist unabhängig voneinander und es gibt einen signifikanten Zusammenhang ($\alpha=0,05$).

(g) 1. Fragestellung:

Ist die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Protowörtern gleich verteilt?

H_0 = Die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Protowörtern ist gleich verteilt.

H_1 = Die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Protowörtern ist nicht gleich verteilt. ($\alpha=0,05$).

2. Fragestellung:

Ist die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Zielwörtern gleich verteilt?

H_0 = Die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Zielwörtern ist gleich verteilt.

H_1 = Die Auftretenshäufigkeit der vier semantischen Klassifikationstypen in Zielwörtern ist nicht gleich verteilt. ($\alpha=0,05$).

3. Fragestellung:

Ist die Verwendung von Zeigegesten auf semantische Klassifikationstypen beschränkt?

H_0 = Die Auftretenshäufigkeit von Zeigegesten in der Produktion der vier semantischen Klassifikationstypen ist gleich verteilt.

H_1 = Die Auftretenshäufigkeit von Zeigegesten in der Produktion der vier semantischen Klassifikationstypen ist nicht gleich verteilt. ($\alpha=0,05$).

4. Fragestellung:

Ist die Verwendung von Zeigegesten auf spezifische Wortformen beschränkt?

H_0 = Die Auftretenshäufigkeit von Zeigegesten in der Produktion verschiedener Wortformen ist gleich verteilt.

H_1 = Die Auftretenshäufigkeit von Zeigegesten in der Produktion verschiedener Wortformen ist nicht gleich verteilt. ($\alpha=0,05$).

4.3 Ergebnisse

4.3.1 Qualitative Auswertung der Sprachdaten von Kind J.

In **Sitzung 1** produzierte J. so gut wie keine Zielwörter. Nach Maßgabe der semantischen Klassifikation konnte man die Formen, die er produzierte, in zwei Gruppen einteilen. Die eine Gruppe umfasste phonologische Formen, die unter PERSONAL fielen, die andere solche Formen, die als INTERAKTIONAL gelten konnten (vgl. Tab. 27).

PERSONAL	INTERAKTIONAL
ʔuj	de
ʔu ^h	da
ʔa	

Tabelle 28: Wörter von J. in Sitzung 1.

Schematisch kann man dementsprechend folgende Repräsentationen ableiten:

SEMANTIK [PERSONAL] →	PHONOLOGIE CVC	oder	CV
	Plosiv + Vokal + Gleitlaut Konsonantischer Artikulationsort: DORSAL		Plosiv + Vokal DORSAL
[INTERAKTIONAL] →	CV Plosiv + Kurzvokal Konsonantischer Artikulationsort: KORONAL		

Diese grundsätzlichen Zuordnungen blieben über zwei Aufnahmetermine bestehen, wobei es zu phonetischen Variationen in der Realisierung dieser Grundformen kam. So wurde z.B. [INTERAKTIONAL] in den Varianten [da], [de], [te], [ge], [di], [hi] oder [do] realisiert (wobei [ge] und [hi] sehr selten auftraten). Während die Formen unter PERSONAL keinen Zielwortbezug zeigen und entsprechend als Protowörter zu klassifizieren sind, können die Varianten unter INTERAKTIONAL als deiktisches Element im Sinne von *da* interpretiert werden, so dass hier eine Zielform, wenn auch noch phonologisch variabel, frühzeitig realisiert wird.

In **Sitzung 3** kam zur Realisierung des Merkmals [PERSONAL] eine weitere Form mit Diphthong hinzu ([ʔaw]), wobei der Gleitlaut [w] ebenfalls für den Artikulationsort [DORSAL] wie aber auch [LABIAL] spezifiziert ist (vgl. Hall 2000: 118).

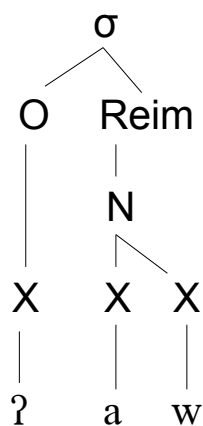


Abbildung 62: Silbenstruktur des Protoworts [ʔaw] in Sitzung 3.

Die Gleitlaute, [j] (aus [ʔuj]) und [w] (aus [ʔaw]) bekommen in **Sitzung 4** einen besonderen Status. J. bildet hier erste zweisilbige Formen mit dem Merkmal [PERSONAL], nämlich die ambisilbischen Protowörter [ʔaja] und [ʔawa].

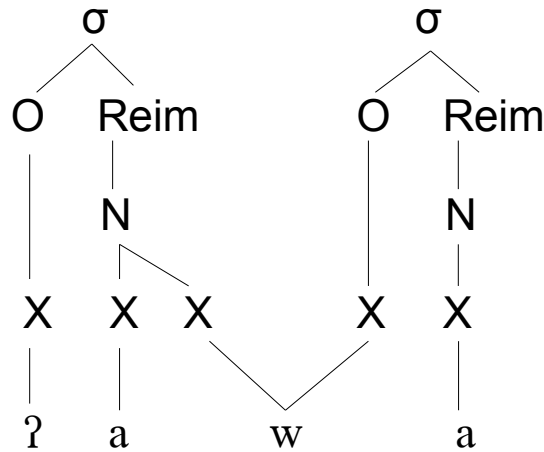


Abbildung 63: Ambisilbische Struktur des Protoworts [ʔawa].

Wie unschwer erkennbar ist, bilden [j] und [w] hier den Onset einer zweiten Silbe, wobei der Vokal redupliziert wurde (vgl. Abb. 63). Anders als es häufig in Studien beschrieben wird, liegt hier nicht der Fall vor, dass durch die Reduplikation ein Kodakonsonant erworben wird.⁷⁰ Es handelt sich schließlich um ein Phonem innerhalb der strukturellen Nukleusposition, das diese Position auch invariant beibehalten wird. Der Gleitlaut der ersten Silbe ist vielmehr das einzige konsonantische Element, das auf dieser Erwerbsstufe zur Bildung eines Silbengelenks verfügbar ist. Die zur gleichen Zeit gegebene Form Koronaler Plosiv + Kurzvokal (wie z.B. [de]) wird interessanterweise nicht redupliziert. Man muss annehmen, dass CVCV-Sequenzen auf dieser Stufe noch nicht prosodisch integriert sind, d.h., als disparate Einheiten für die Betonungszuweisung gelten können (wie es nach Fikkert's Modell später in Phase 2 der Trochäuszuweisung bei CVCV-Abfolgen geschieht). Diese Etablierung als prosodisch verwertbare Einheit respektive phonologisches Wort muss erst erfolgen. Wie kann man sich eine solche Prozedur der Strukturweiterung bzw. -etablierung anhand obiger Plosiv-Vokal-Gleitlaut-Strukturen vorstellen?

Wie in Abbildung 63 illustriert, wird durch die Bildung einer ambisilbischen Struktur eine neue Onset-Position erzeugt. Diesen Effekt hat Ambisilbifizierung in der Zielsprache über Wortgrenzen hinweg, unabhängig von der Betonung der jeweiligen Silben. Im Gegensatz dazu erzeugt wort-mediale Ambisilbifizierung eine neue Koda-Position und ist auf Onsets betonter Silben beschränkt (vgl. Ito & Mester 2009: 231). Während letztere Form der Ambisilbifizierung also innerhalb von schon etablierten respektive lexikalisierten

⁷⁰ So wird z.B. in Ross (1937) *bag* vor dem stabilen Erwerb zu [baga] redupliziert.

Worteinheiten geschieht, ist die über Wortgrenzen operierende Form der Ambisilbifizierung das Ergebnis der kontextsensitiven On-line-Silbifizierung (vgl. Kap. 1.5.2) und wird somit innerhalb des regulären Sprachproduktionsprozesses nach Maßgabe der Onset-Maximierung appliziert. Diese Form der Ambisilbifizierung kann somit als Mittel zur temporalen Realisierung eines Zweisilbers genutzt werden. Da auf dieser Erwerbsstufe keine zweisilbigen prosodischen Wörter in die Sprachproduktion eingehen können, ist das Höchstmaß an Information, die berechnet werden kann, zwei disparate Silben. Diese beide disparaten Informationselemente werden vermutlich im phonologischen Arbeitsgedächtnis in der gleichen Art und Weise zu einer Einheit verbunden, wie disparate phonologische Wörter in der Produktion erwachsener Sprecher. Im sprachtherapeutischen Prozess können genau solche Speicher- und (Re-)Analyseprozesse über die Onlineverarbeitung reaktualisiert und damit nutzbar gemacht werden. Bernhardt & Stemberger (2000: 145) illustrieren ein solch intervallbasiertes Vorgehen in der sprachtherapeutischen Behandlung:

„Gradually decrease the interval between the two syllables until C1VC2V is produced [...].
„Bee“ „Knee“ > „Bee, Knee“ > „Beany“
C1VC2V can be a starting point for a C1VC2 monosyllable, by eliminating the final rime “y.”
“Beany“ > “Bean, y“ > “Bean“ “y“ > “Bean“
CVCV > CVC, V > CVC V“ > CVC .“

Zwei schon gespeicherte Informationseinheiten können also über mehrmalige, gesteuerte Abrufprozesse zu einer neuen Informationseinheit zusammengefasst und umgeschrieben oder auch umgekehrt aus einer Informationseinheit zwei neue Informationseinheiten extrahiert werden. Unter einer solchen Perspektive bildet dies somit das phonologische Pendant zu Prozessen, die auch später noch im Erwerb in der Produktion von Sätzen beobachtet werden können. In Kapitel 1.4.1 wurde bereits auf das Phänomen der Bauplanüberfrachtung hingewiesen (vgl. Tracy 2000: 34). Die von Tracy so genannte *dazə*-Vorläuferkonstruktion erschließt als ganzheitlicher deiktischer Platzhalter neue strukturelle Positionen in der Produktion (ebd.: 33):

J. (2;4) [dazə] BAUernhof \ ... da IS das bauernhof\
J. (2;4) [daza] schava (=Scheuche) ... da IS-i vogelscheuch\

Auch Hohenberger (2002) beschreibt Satzproduktionen ihres Sohnes Tilman im Alter von 2;09 – 3;02 Jahren, in denen ein syntaktischer Bauplan überfrachtet wird, z.B. *der will aufs andere Blatt will der*, eine Verschränkung zweier Satzpläne (ebd. 214):

[IP1 der will [IP2 auf andere blatt IP1] will der IP2]

Hohenberger deutet solche Verschränkungen im Rahmen der Selbstorganisation eines dynamischen (Sprach-)Systems als Ausdruck einer liminalen Phase im Strukturbildungsprozess (ebd.: 217):

„I claim that these particular blends witness the last fraction of the liminal phase in which the determination of the various functions of IP/CP takes place. The blending is almost artistically exaggerated as if he wanted to tell himself and us: „Look, I have understood that I can leave an element in its base-generated position (i.e. in the VP) but I can also topicalize it to SpecIP/CP.“

Man kann solche Daten m.E. eher im Sinne der Strong Identity Hypothesis erklären. Worauf Hohenberger (ebd.: 213) selbst hinweist, ist, dass solche Oberflächenverschränkungen auch bei Wernicke-Aphasikern auftreten (Huber & Schlenck 1988) und bei diesen auf Dysfunktionen innerhalb des Sprachproduktionssystems zurückgeführt werden können. Oberflächenverschränkungen bestehen hier aus zwei Sätzen, die eine gemeinsame Konstituente teilen, die also Bestandteil der syntaktischen Struktur beider Sätze ist und mit der Wortordnung beider Sätze vereinbar ist, wie in „Ich bin [(am See) spazieren bin ich] am See“. Huber & Schlenck sehen solche Formen in einer Störung der Kontrollfunktion während der Satzplanung begründet; zwei Formen werden gleichzeitig linearisiert und da die zugrundeliegenden Strukturen gleiche Wortformadressen beinhalten, kommt es zum verzeichneten Verschränkungsphänomen, dem Wechsel von einer Struktur zur anderen im Linearisierungsprozess. Für die Basisstruktur NP (PP) V werden also zwei Oberflächenstrukturen NP AUX PP und PP AUX NP gleichzeitig eingespeist, die zugrundeliegende syntaktische Kenntnis selbst ist allerdings ungestört.

Wichtig erscheint mir, dass beide Strukturen Tilmans im Sinne ihrer Genese und der Verfügbarkeit struktureller Positionen nicht gleichrangig sind. Der Deklarativsatz ist unmarkierter als die Struktur, in welcher aus der Verbalphrase heraus topikalisiert wurde. Wie Hohenberger betont, ist die Abfolge der Sätze immer unmarkiert-markiert und beide Strukturen konnten zeitgleich unabhängig voneinander produziert werden. Nun beschreibt Hohenberger, dass Tilman vor dem zeitlichen Erscheinen der Oberflächenverschränkungen beide Formen simultan in einer Äußerung verwendete wie in *ah, da is meine kuh – meine kuh is da* (2;05, 30) (ebd.: 210).

Wir finden also die Reihenfolge: vereinzelte Produktion → simultane Produktion → Oberflächenverschränkungen. Letztere wird, wie Hohenberger betont, als ein überfrachteter, aber prosodisch monolithischer Phrasenstrukturbaum produziert⁷¹:

⁷¹ Die Strukturanalyse unterscheidet sich hier von Hohenberger (2002: 215). In ihrer Analyse fehlt die Spezifikator-Etikettierung unter IP und es ist keine Koindizierung der PP mit den beiden Spezifikatoren

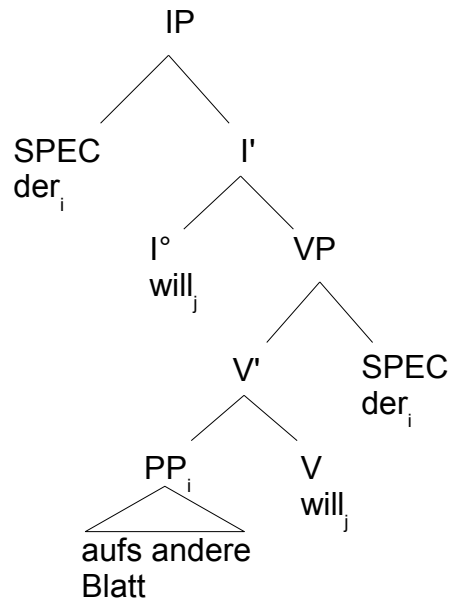


Abbildung 64: Struktur einer Oberflächenverschränkung nach Hohenberger (2002: 215).

Wichtig ist m.E., dass die Oberflächenverschränkungen nicht einfach nur der Ausdruck einer liminalen Phase sind, sondern eben die hauptsächlich relevante Voraussetzung für die nächsten Erwerbsschritte. Ab dem Alter von ca. drei Jahren, also kurz nach dem Alter, in dem Tilman Oberflächenverschränkungen produziert, beginnen Kinder mit der Produktion von komplexen Satzgefügen (was leider für Tilman nicht dokumentiert ist), ein Erwerbsschritt, der von Tracy (2000) als Meilenstein IV der syntaktischen Entwicklung bezeichnet wird. Was Tilman in der obigen Oberflächenverschränkung vorläufig in einer „falschen“ Struktur realisiert, beinhaltet ja nichts anderes als die Möglichkeit einer zweiten Subjektposition bzw. IP innerhalb einer syntaktischen Gesamtphrase, also genau das, was in Meilenstein IV strukturell realisiert werden soll. Die Annahme ist nun, dass Tilmans Sprachsystem durch die überfrachteten Oberflächenverschränkungen im Arbeitsgedächtnis eine solch komplexe Struktur als Lösungsentwurf vorerst simuliert, um später die angestrebte Zielstruktur produzieren zu können (die er passiv schon repräsentieren kann). Wenn man die Bezeichnung struktureller Positionen als Landeplätze beibehält, kann an dieser Stelle für solche überfrachteten Strukturen der Terminus *Simulator* eingeführt werden.⁷² Neue (syntaktische)

angegeben. Diese Verdopplung von Information und damit die Überfrachtung ist aber gerade für diese Strukturen wesentlich, weshalb sie hier angegeben wird.

⁷² Es muss an dieser Stelle betont werden, dass in Tracys Konzept der Bauplanüberfrachtung das Grundlegende der hier dargestellten Simulatorenkonzeption schon angelegt ist. So schreibt sie: „Im Deutschen simulieren Äußerungen wie [maxta], [dazitsta], [vo:ze] *ball* bereits Verbzweiteffekte des Erwachsenensystems.“ (Tracy

Landeplätze werden in solchen Simulatoren durch Nutzung bisheriger Ressourcen erstmalig realisiert. Die Operation, mithilfe der solche Strukturen im Arbeitsgedächtnis erzeugt werden, ist, wie bei allen anderen Strukturen, UNIFY. Jackendoff (2007b: 63) formuliert den genauen Prozess der Unifizierung wie folgt:

„Unification is sort of a Boolean union on feature structures: unifying A and B results in a composite that shares all common features of A and B and preserves all distinct features of A and B.“

Informationen der Form [V, +past] und [V, 3 sing] wird nach dieser Operation zu [V, +past, 3 sing] unifiziert (und nicht im Sinne von Merge zu [[V, +past] [V, 3 sing]] (ebd.)).⁷³ An dem Beispiel von Tilmans Oberflächenverschränkungen sieht man klar die Applikation von UNIFY. In Simulatoren werden demnach durch schon vorhandene Strukturen neue Optionen der Unifizierung getestet und durch Chunking im Kurzzeitgedächtnis die Ressourcennutzung des Arbeitsgedächtnisses optimiert. D.h., es muss erst einmal eine überfrachtete Struktur im Arbeitsgedächtnis erzeugt werden, um diesen Überschuss an Struktur durch Etikettierungen im Kurzzeitgedächtnis für eine ökonomische Strukturzeugung und letztlich Strukturweiterung nutzen zu können. Unter Zugrundelegung von TOTE-Einheiten, wie sie auch im MEKIV-Modell Verwendung finden (vgl. Kap. 1.6), kann so die sukzessive Erweiterung von sprachlicher Struktur im Rahmen einer klassischen (und kybernetischen) Architektur modelliert werden. Die Annahme kompensierender Rückkopplungsschleifen verdeutlicht den permanenten Abgleich von IST-Zustand der Produktion (Arbeits- und Kurzzeitgedächtnis) mit dem SOLL-Zustand rezeptiv bereits erworbener Strukturen (Langzeitgedächtnis). Während Störungen bei Wernicke-Aphasikern in dem Sinne strukturell rückläufig sind, als dass die Kontrollfunktionen von Performanzsystemen der Sprachplanung und -produktion ausfallen, findet im Spracherweb also ein progressiver Adjustierungsprozess derselben an vorhandene Kenntnis statt. Beide, Aphasiker wie das die Sprache erwerbende Kind, haben dennoch immer mehr strukturelle Kenntnis, als sie overt realisieren können.

Nach diesen etwas längeren Ausführungen zum Syntaxerwerb kann man also auch für die Formen mit ambisilbischen Nukleus annehmen, dass sie Simulatoren in einem solchen Sinne darstellen, als dass sie die reguläre Onset-Position einer zweiten Silbe innerhalb eines prosodischen Wortes zur Verfügung stellen, wobei auch eine „falsche“ Form produziert wird

2000; Hervorh. d. Verf.) Allerdings wird von ihr m.E. die Interaktion zwischen unifizierten Strukturen, Arbeitsgedächtnis, Sprachproduktion und -verstehen nicht explizit ausgearbeitet.

⁷³ Ähnlich wird es in Blanken, Dittmann, Wallesch (1988: 136-137) angenommen. Zu Verschmelzungen kann es nach ihnen nur kommen, wenn mindestens zwei Strukturen Verkettungsregeln mit gleichen Eingabeelementen beinhalten.

(eben ein ambisilbischer Nukleus). Dies geschieht ebenfalls nach dem Motto: „Eine Zielstruktur, die noch nicht realisiert werden kann, wird im Arbeitsgedächtnis strukturell durch die Unifizierung bereits produzierbarer Strukturen vorläufig simuliert.“ Auch hier ist es die grundlegende Voraussetzung dafür, dass neue, strukturell erweiterte Chunks gespeichert werden können. Wie in Abbildung 65 dargestellt, können so aus einem Simulator im Arbeitsgedächtnis drei neue Chunks im Kurzzeitgedächtnis als Basis möglicher Strukturen abgeleitet und gespeichert werden.

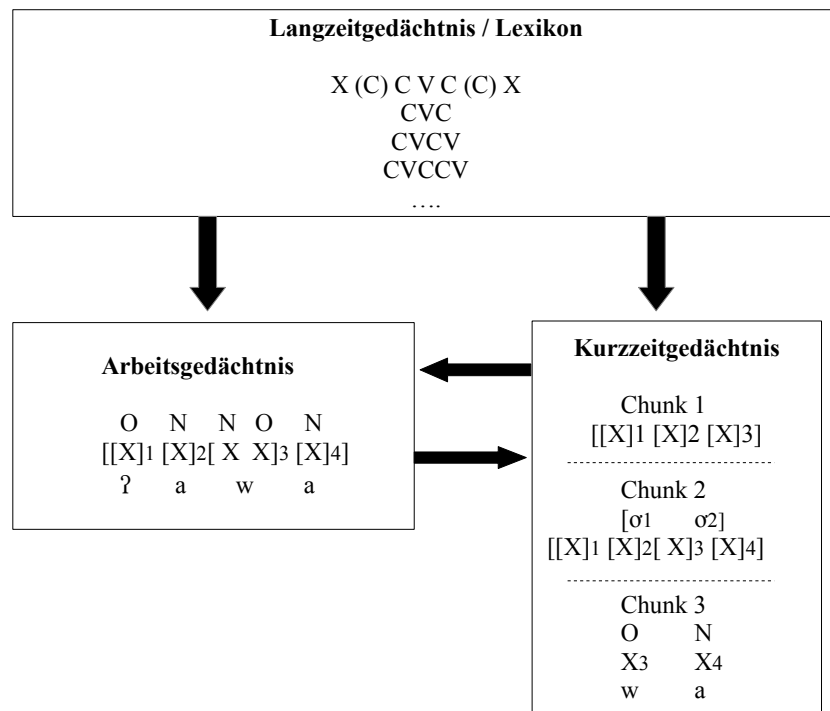


Abbildung 65: Speicherung von Chunks auf Basis einer Simulatorstruktur (X = Einheit auf X-Ebene, O = Onset, N = Nukleus).

Der Prozess der Reduplikation nimmt für die ambisilbischen Produktionen eine kritische Rolle ein. Was bedeutet es für die Reduplikationen der Sitzung 4 genau, wenn man von einer Überfrachtung des Bauplanes spricht? Im Grunde einfach nur, dass durch die Reduplikation einer Silbenkonstituente im Arbeitsspeicher eine Strukturweiterung stattfindet, die wie die Oberflächenverschränkungen Tilmans in einer unmarkiert-markiert-Abfolge realisiert wird. Die minimale Informationseinheit, die dafür benötigt wird, ist eindeutig ein vokalisches Element und damit notwendig eine silbenkernbildende Einheit. In Rekurs auf Framptons Distributed Phonology (vgl. Kap. 1.7.2) kann man die partielle Reduplikation eines Vokals modellieren:

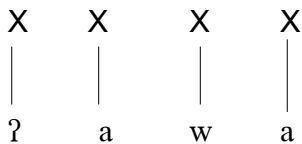
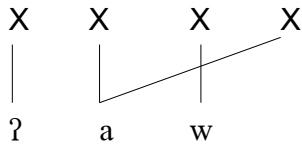
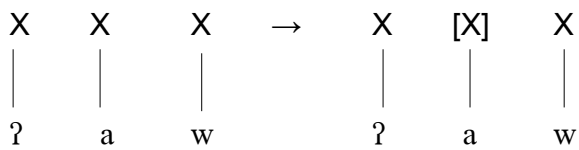


Abbildung 66: Partielle Reduplikation des Vokals in [ʔawa].

Unter Bezugnahme auf die Argumentation in Kapitel 1.4.2 kann man hinzufügen, dass hier schon zugrundeliegend, d.h. vor der eigentlichen Silbifizierung, der Sonoritätszyklus von Anstieg und Abfall gegeben ist. Dadurch, dass Diphthonge eben nicht als Vokal-Vokal-Einheiten verstanden werden, sondern als nukleuserne Kombination von Vokal und Gleitlaut, ist die in den Daten evidente ambisilbische Realisierung durch den Sprachplanungsapparat in ihrer Verarbeitungslogik nachvollziehbar.

Dieser hier beobachtete Freischaltungseffekt von Ressourcen durch chunkingbasierte Automatisierung, d.h. der Bildung von Abrufindizes im Kurzzeitgedächtnis, lässt sich in seiner Wirkung auch in den folgenden zwei Sitzungen nachweisen.

In **Sitzung 5** wird das zuvor freigeschaltete silbenstrukturelle Format CVCCV mit ambisilbischem Gleitlaut genutzt, um mit der PERSONAL etikettierten Form [ʔuja] eine erste Alternation der wortinternen vokalischen Elemente zu erzeugen. Mit [jaja] tritt, ebenfalls PERSONAL gebraucht, eine erste CVCV-Reduplikation auf. Zeitgleich produziert J. mit den Formen der vorherigen Protowortform [ʔu] und dem neu hinzugekommenen [wu] eine erste Variante eines Zielworts, entweder das Nomen *kuh* oder das Onomatopoetikum *muh*. D.h.,

dass in den Sitzungen 1-3 nur in der Silbenkoda von PERSONAL spezifizierten Wörtern und in Sitzung 4 in der Onset-Position ambisilbischer Strukturen realisierte Phonem [w] wird nun auch für INTERAKTIONAL markierte Wörter verwendet und tritt hier nun wortinitial in Onset-Position auf. Insofern wurden sowohl das ambisilbische Strukturformat als auch die positionale Verwertbarkeit von [w] durch die zuvor im Arbeitsgedächtnis realisierten Strukturbildungen im Kurzzeitgedächtnis etabliert. [w] ist bis zu diesem Zeitpunkt auch das einzige Phonem, das für das Merkmal LABIAL spezifiziert ist. Während in den bisherigen Sitzungen silbeninitial nur Konsonanten mit dem Artikulationsort DORSAL und KORONAL produziert wurden, kann in **Sitzung 6** nun erstmals der labiale Plosiv [b] in der Onset-Position beobachtet werden. Auch hier lässt sich vermuten, dass das Merkmal [LABIAL] auf dieser Position in Sitzung 5 eine Freischaltung erfahren hat. [b] wird nun sowohl in Formen des Bedeutungsmerkmals PERSONAL als auch INTERAKTIONAL in einer Zielwortvariante von *Buch* reduplikativ und mit Diphtong verwendet:

PERSONAL	INTERAKTIONAL
bo.ʔa	bubi/bøbø/bubu/buj/boj/bø= BUCH

Schon an dieser Stelle erkennt man ein wichtiges Prinzip. Während die Protowörter strikt PERSONAL markiert sind, bilden sie dennoch die strukturellen Vorläufer für die INTERAKTIONALEN und REGULATIVEN Zielwörter. Es scheint bei ihrer Produktion primär um die phonologische Form an sich, d.h., um die Elaboration und Etablierung von Struktur und neuer Merkmale und Merkmalsmengen zu gehen und nicht um ihre konversationelle Verwendung und Zieladäquatheit. Protowörter scheinen die virtuellen Testflüge von Zielformen darzustellen, da letztere in Form unmittelbarer Post-hoc-Produktionen die gleichen phonologischen Strukturformate, strukturellen Landeplätze wie auch Merkmalsmengen zu ihrer Realisierung nutzen. So wird z.B. in Sitzung 6 für die Formen [bubi/bøbø/bubu] die unmarkierte Reduplikationshypothese „Letztes Segment → erstes Segment“ im Sinne von Raimy (vgl. Kap. 2.2.4.1) verwendet, nachdem in Sitzung 5 mit [jaja] eine erste Form durchlaufen wurde.

Ein weiteres wichtiges Detail in Sitzung 6 ist, dass in der für PERSONAL spezifizierten Wortform [bo.ʔa] erstmalig eine zweisilbige nicht-reduplikative Form erzeugt wird, die

keinen ambisilbischen Gleitlaut enthält. Interessant ist für die Wortform [bo.ʔa] wie auch für die reduplikativen Formen mit Gleitlaut (wie [ʔa.wa]), dass der Glottisverschlusslaut als Grenzsymbol eine besondere Rolle einnimmt. Wiese (1996: 59-60) weist auf die Interdependenz der Glottisverschlusslautinsertion mit der Silbifizierung und der prosodischen Struktur insgesamt hin. Die Insertion eines nicht-phonemischen [ʔ] wird von Wiese als eine Default-Regel interpretiert, um dem Prinzip der Onset-Maximierung zu genügen. Insofern kann man [ʔ] in solchen Kontexten als Marker für Silben- und Wortgrenzen verstehen, umso mehr deshalb, weil die Insertionsregel aufgehoben wird, wenn die Silbifizierung unter bestimmten Bedingungen die Wortgrenzen überschreitet.⁷⁴ Gemessen an einer solchen Interpretation, nach der [ʔ] auch in Formen wie [ʔa.wa] erst während der Silbifizierung inseriert wird, gilt auch hier, dass die unmarkierte Reduplikationshypothese „Letztes Segment → erstes Segment“, die als morphologische Applikation der Silbifizierung vorangeht, wirksam ist. Für den weiteren Ausbau der Struktur ist der Glottisverschlusslaut in [bo.ʔa] eben als klarer Marker für zwei eindeutig disparate Silben zu deuten. Erstmals wird der Glottisverschlusslaut wortmedial inseriert und scheint damit die Produktion eines disparaten zweiten Silbenonsets zu initiieren. Dies kann unmittelbar in **Sitzung 7** beobachtet werden. Die für PERSONAL spezifizierte Wortform [ʔo.ba] wie auch die INTERKATIONAL gekennzeichnete Form [ʔa.jo] (Zielwort: Hallo) enthalten nun wortmedial einen Konsonanten im Onset, nämlich [b] und [j]. Auch hier zeigt sich also wieder, dass das Sprachsystem Ressourcen der Online-Verarbeitung nutzt, um sukzessive eine Annäherung an den SOLL-Zustand der Zielform zu erreichen.

In **Sitzung 8** produzierte J. nicht viele Wortformen. Auffallend war hier der häufige Gebrauch von [ja] analog zu dem deiktischen [da]. Der Gleitlaut [j] wurde zuvor hauptsächlich medial innerhalb eines Zweisilbers verwendet. Dies zuerst als ambisilbischer Konsonant in PERSONAL spezifizierten Protowörtern, dann im Onset einer Silbe von [ʔa.jo] (Zielwort: Hallo), die durch eine eindeutige Silbengrenze von der Nachbarsilbe getrennt ist. Nach dieser definitiven Etablierung einer Silbengrenze und der damit einhergehenden Etablierung eines Abrufindex im Kurzzeitgedächtnis scheint nun die Produktion als unabhängige Einzelsilbe möglich.

⁷⁴ „If syllabification crosses a word boundary (as it may in fast speech), creation of a consonantal onset bleeds ʔ-Insertion.“ (Wiese 1996: 60)

Nachdem im Onset sowohl ein labialer Plosiv [b] als auch ein labialer Frikativ [w] produziert werden konnten, tritt in **Sitzung 9** mit dem INTERAKTIONAL markierten Zielwort *mama* nun der erste Nasal mit Spezifizierung für das Merkmal [labial] auf. Ein eigenständiger Nasal wurde in Protowörtern nicht produziert. In der Produktion von [ʔaj] wurde wieder eine etablierte amsilbische Form [ʔa.jo] für eine Neubildung, diesmal zur Bildung eines Einsilbers durch Vokaltilgung gebraucht. Sie wird in Bezug auf rundlich geformte Objekte, wie *Bonbon* oder *Luftballon* verwendet. Die reduzierte Form [ʔa] wird in dieser Sitzung in einer Spielsituation mit Autos erstmalig gebraucht.

In den folgenden Sitzungen scheint eine Etablierung der bisherigen phonologischen Formen stattzufinden mit variierendem Gebrauch und sich die Anzahl der Protowortformen zu verringern. So wird schon in **Sitzung 10** nur noch das Protowort [ʔuj] produziert. Die Form [ʔaj] wird INSTRUMENTAL verwendet, wenn J. Gegenstände haben möchte, also eventuell eine frühe Form von *mein* oder *meins*. Sowohl für die Situation des Bücherlesens wie auch für Spielsituationen mit Tierfiguren werden Variationen mit labialem und stimmlosem Plosiv [b] gebraucht ([bow] *Buch lesen* vs. [bu] / [buj] *Mit Tieren spielen*). Dies gilt fast deckungsgleich für **Sitzung 11**. Erst in **Sitzung 12** wird mit der Form [naj] *nein* ein neuer, für koronal spezifizierter Nasal, nämlich [n] produziert, der aber vorher nicht in Protowörtern produziert wurde. Insgesamt wurde das Merkmal [nasal] also exklusiv in Zielwörtern produziert und scheint den definitiven Ablösezeitpunkt phonologischer Produktion von den notwendigen Vorläuferstrukturen zu bilden. Mit [ʔawaw] für *Hund* tritt eine CVC.CVC-Reduplikation auf, die auch in **Sitzung 13** für *Auto* verwendet wird. Die letzten beiden Sitzungen (13 und 14), in denen noch das einzige Protowort [ʔuj] verwendet wurde, sind durch keine wesentlichen strukturellen Innovationen gekennzeichnet. Vielmehr scheint das Prinzip der Zieladäquatheit zu dominieren und Annäherungen an diese nun bis zur abschließend analysierten **Sitzung 14** primär zu sein, wie in [bubu, bo, bobo] für *Bus/größeres Auto*; [baba, babaj, baj] für *Ball* oder [ʔawaw, ʔawa, ʔaw] für *Auto*.

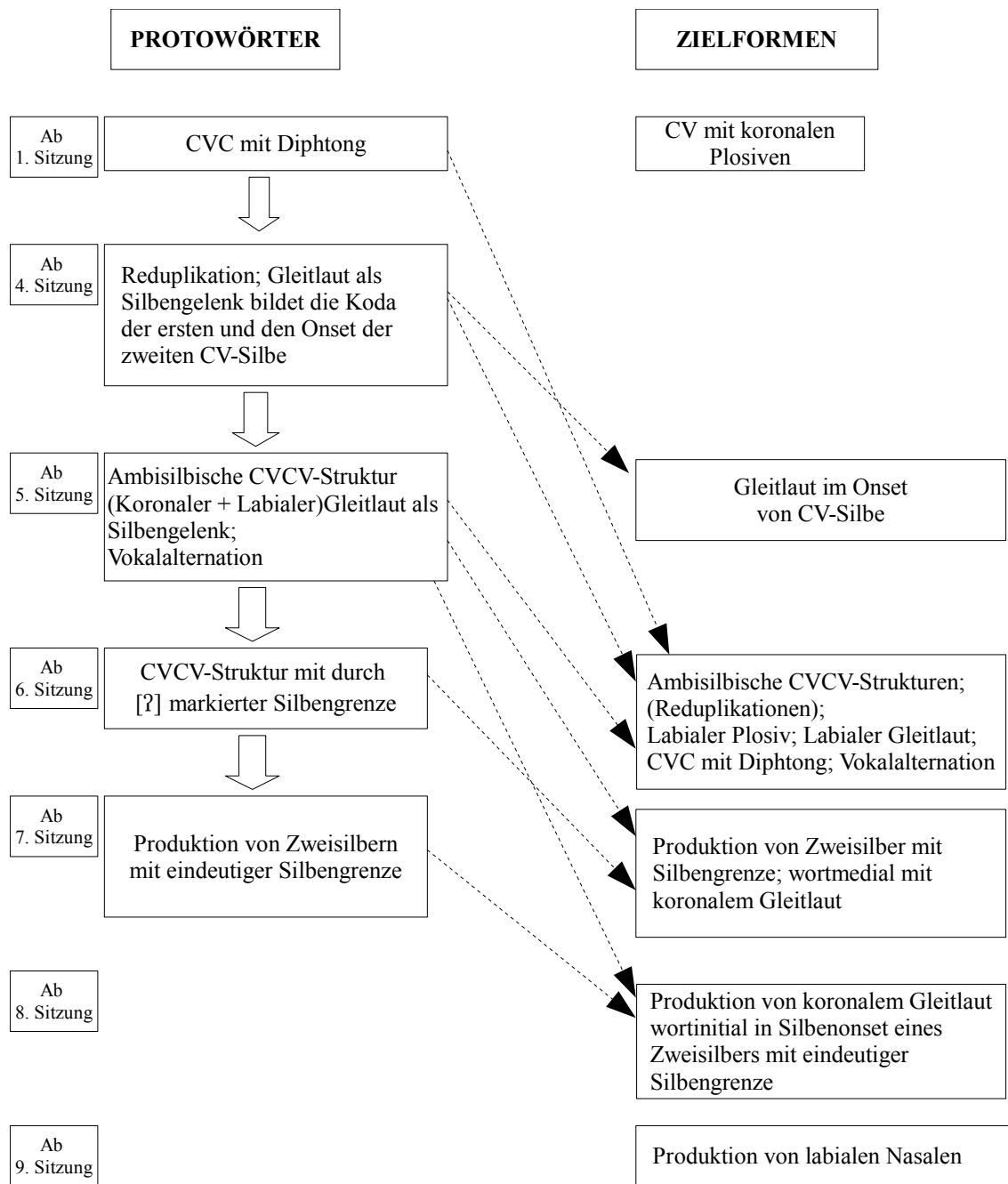


Abbildung 67: Freischaltungs- und Transferprozesse von Proto- zu Zielwörtern.

Es wurde bereits darauf hingewiesen, dass mit der Produktion des Merkmals [nasal] in Sitzung 9 die Vorläuferfunktion von Protowörtern endete. In Abbildung 67 sind noch einmal alle wesentlichen Transferprozesse zwischen Proto- und Zielwörtern bis zu diesem Zeitpunkt zusammengefasst.

Fokussiert man den Erwerb von phonologischen Merkmalskontrasten gesondert, so kann man also definitiv aus den Daten ableiten, dass ihr Erwerb schon auf Protowortebene beginnt.

Jakobsons angenommene Erwerbsreihenfolge vom Kontrast Konsonant-Vokal in labialen Plosiv-Vokal-Silben zum Kontrast zwischen labialen Plosiven und Nasalen (vgl. Jakobson 1941) kann hier nicht bestätigt werden. Vielmehr kann man mithilfe des Modells von Rice & Avery (1995; vgl. Kap. 2.1.1) eine etwas genauere Beschreibung der Erwerbsreihenfolge geben.

Merkmal	Laryngeal	Air Flow	SV	Place
Anfangsstruktur	(CG)?	Default (Stop)	-	Default (Coronal)
Neuer Kontrast	-	-	-	Ab Sitzung 6 Peripheral: Koronale vs. Labiale
Neuer Kontrast	-	-	Ab 9. Sitzung Default (Nasal): Bildung von [m] und [n] (12. Sitzung)	-

Tabelle 29: Kontrastentwicklung in J.s Produktionen.

Die ersten Konsonanten, die von J. gebildet wurden, waren stimmlose koronale Plosive. Diese und der Glottisverschlusslaut wurden in der Onsetposition von Silben produziert. Während die Defaultmerkmale [coronal] und [stop] gemäß der Modellierung von Rice & Avery (ebd.) früh produziert werden sollten, passt die Produktion von stimmhaften Plosiven nicht, da hier als Defaultmerkmal [spread glottis] gegeben sein sollte. Aber auch mit der sukzessiven Ausprägung der Merkmale [peripheral] und [nasal] ist der Erwerb von J. mit diesem Modell adäquat beschreibbar. Hervorheben muss man an dieser Stelle noch einmal, dass bis zur 9. Sitzung alle zielwortrelevanten Phoneme zuerst in Protowörtern auftraten. Dieser strukturelle Zusammenhang war in der qualitativen Analyse eindeutig evident. Im folgenden Kapitel soll untersucht werden, ob dieser Zusammenhang auch ein quantitatives Analogon findet.

4.3.2 Quantitative Auswertung der Sprachdaten von Kind J.

Über den Zeitraum von 14 Wochen, die den 14 Aufnahmesitzungen entsprechen, produzierte J. insgesamt 194 Tokens von Protowörtern und 250 Tokens von Zielwörtern. Wie in Abbildung 68 sichtbar ist, dominierte zu Anfang der Aufnahmen die Produktion von Protowörtern, während ab Sitzung 5 Zielwörter minimal häufiger produziert werden, die aber in den letzten Aufnahmesitzungen eindeutig das Gros der Gesamttokens bilden.

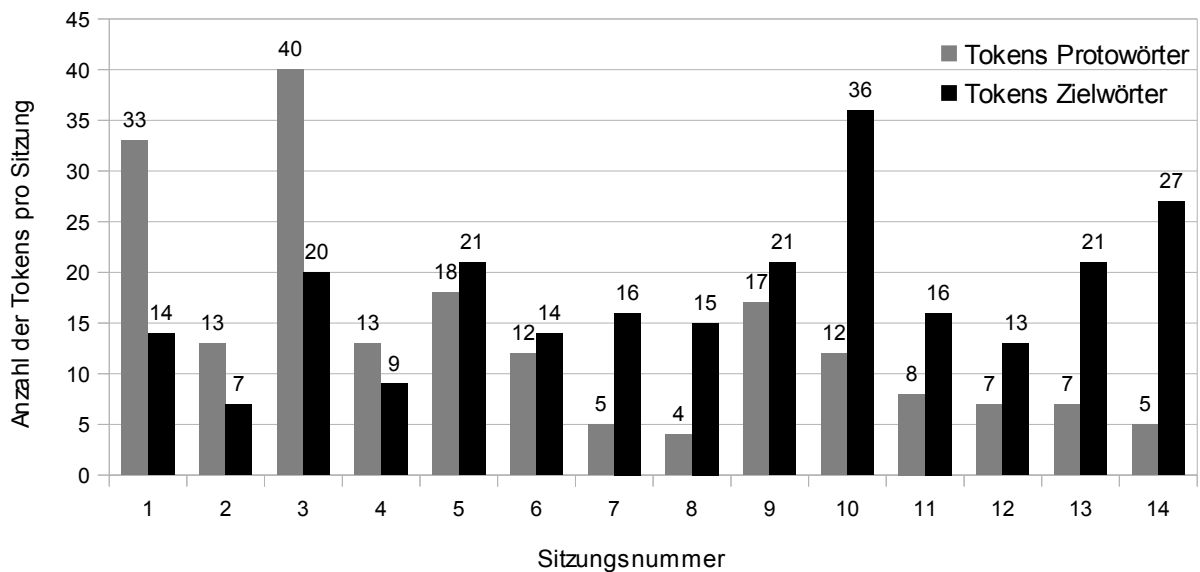


Abbildung 68: Tokens Protowörter und Zielwörter.

Diese rein häufigkeitsbasierte Darstellung kann mithilfe der W-NW-Ratio ergänzt werden, um festzustellen, ob es einen Zusammenhang gibt zwischen dem Lebensalter und dem Produktionsverhältnis von Proto- und Zielwörtern. Tatsächlich kann mittels der Pearson Produkt-Moment-Korrelation eine ausgeprägte Beziehung zwischen Lebensalter und W-NW-Ratio ($r=0.762$, $df=13$, $p<0,001$, $t=3,90$) attestiert werden (vgl. Abb. 69). Die Alternativhypothese (a) H_1 wurde entsprechend bestätigt (vgl. S. 218).

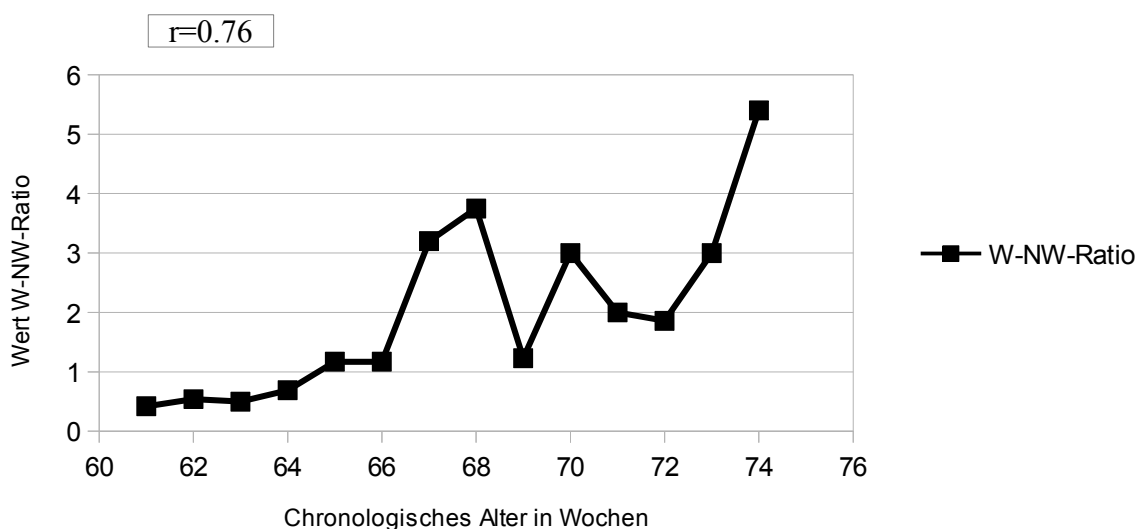


Abbildung 69: Entwicklung W-NW-Ratio.

Die einzigen Silbentypen, die sowohl in Proto- und Zielwörtern produziert wurden, waren CV- und CVC-Silben. Insgesamt wurden in Protowörtern 100 CV-Silben und 114 CVC-Silben, in Zielwörtern 258 CV-Silben und 47 CVC-Silben produziert. Abbildung 70 gibt einen Überblick über die Produktion der Silbentypen über den gesamten Aufnahmezeitraum.

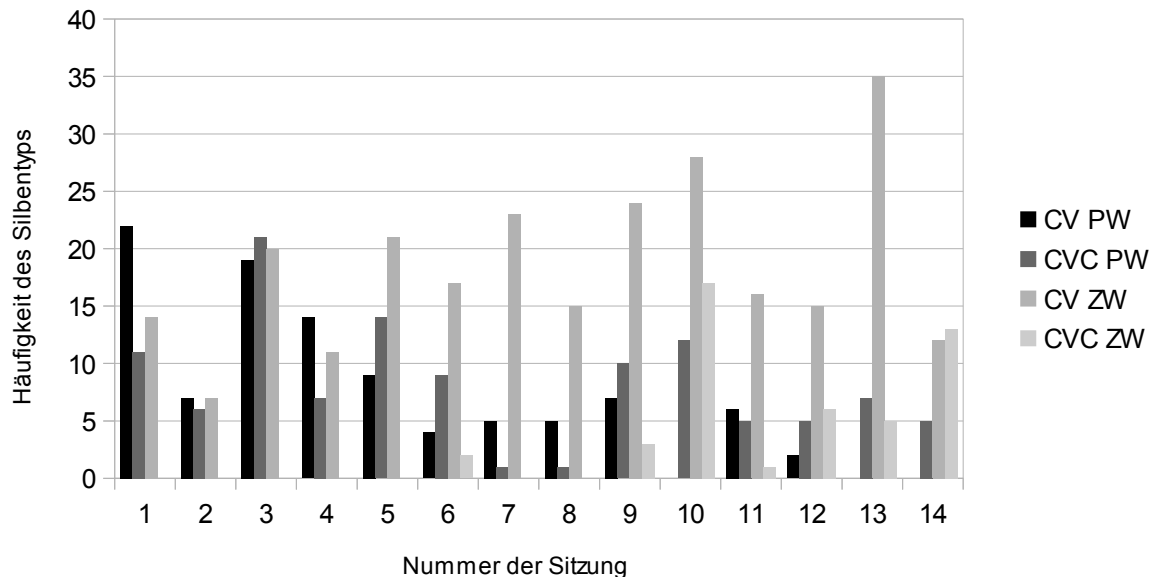


Abbildung 70: Produktion von CV- und CVC-Silben in Proto- und Zielwörtern.

Innerhalb der Protowortproduktion konnte keine signifikante Verteilung von CV- und CVC-Silben ($\chi^2=0,916$, $df=1$, $p<0,3386$) festgestellt werden. Innerhalb der Zielwortproduktion ist dagegen eine signifikante Verteilung von CV- und CVC-Silben zu beobachten ($\chi^2=145,970$, $df=1$, $p<0,0001$). Auf die Unabhängigkeit der Produktion der Silbentypen in Proto- und Zielwörtern getestet, zeigt sich, dass die Produktion hoch signifikant unterschiedlich ist und für den Zusammenhang eine mittlere Effektstärke besteht (Yates $\chi^2=82,49$, $df=1$, $p<0,0001$, Cramérs $V=0,4029$). Im Hinblick auf diese Ergebnisse kann die Alternativhypothese (b) H_1 als bestätigt angesehen werden.

Es wurden nur CV-Silben redupliziert, wobei in der Produktion von Protowörtern nur insgesamt 12, in Zielwörtern hingegen 52 CV-Reduplikationen festgestellt werden konnten. Während die ersten CV-Reduplikationen in Sitzung 4 und 5 in Protowörtern beobachtet werden konnten, entwickelt sich spätestens ab Sitzung 9 eine klare Dominanz von CV-Reduplikationen in Zielwörtern (vgl. Abb. 71). Mittels eines Chi-Quadrat-Anpassungstests kann ein hoch signifikantes ungleiches Vorkommen von Reduplikationen in Proto- und Zielwörtern

($\chi^2=25,000$, $df=1$, $p<0,001$) attestiert werden. Damit kann festgehalten werden, dass Alternativhypothese (c) H_1 verifiziert wurde.

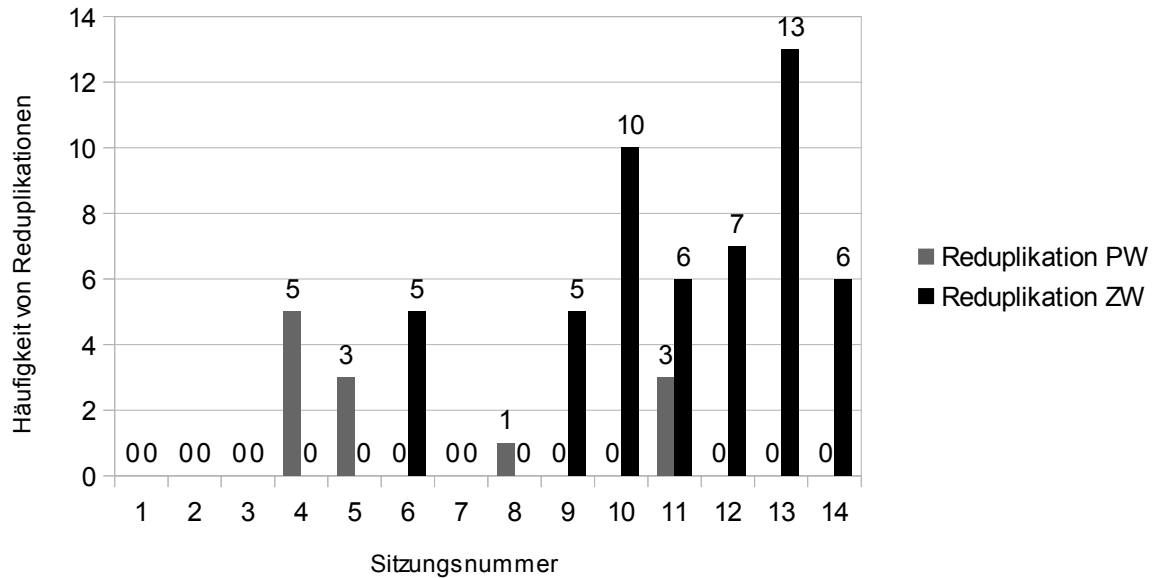


Abbildung 71: Produktion von Reduplikationen in Protowörtern und Zielwörtern.

Die Häufigkeitsverteilung der Produktion von initialen und finalen Demisilben insgesamt ist im Kreisdiagramm/Abbildung 72 dargestellt:

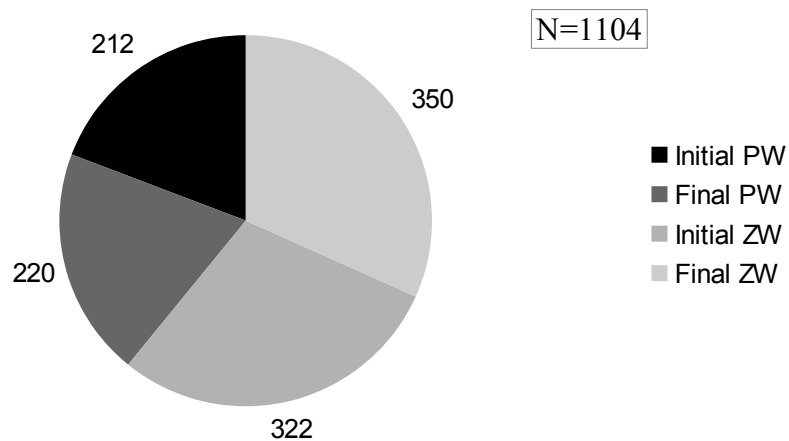


Abbildung 72: Absolute Häufigkeit initialer und finaler Demisilben in Proto- und Zielwörtern.

Die Alternativhypothese (d) 1. H_1 , die besagt, dass alle Demisilbentypen in Protowörtern initial nicht gleich häufig vorkommen, kann aufgrund eines Chi-Quadrat-Anpassungstests bestätigt werden ($\chi^2=699.37$, $df=4$, $p<0,001$), dessen Ergebnis hoch signifikant ist. Auf die reine Häufigkeit bezogen, zeigen die Daten, dass eine klare Dominanz von Plosiv-Vokal-Demisilben gegeben ist (vgl. Abb. 73; PV = Plosiv-Vokal, FV = Frikativ-Vokal, NV = Nasal-Vokal, LV = Liquid-Vokal, GV = Gleitlaut-Vokal).

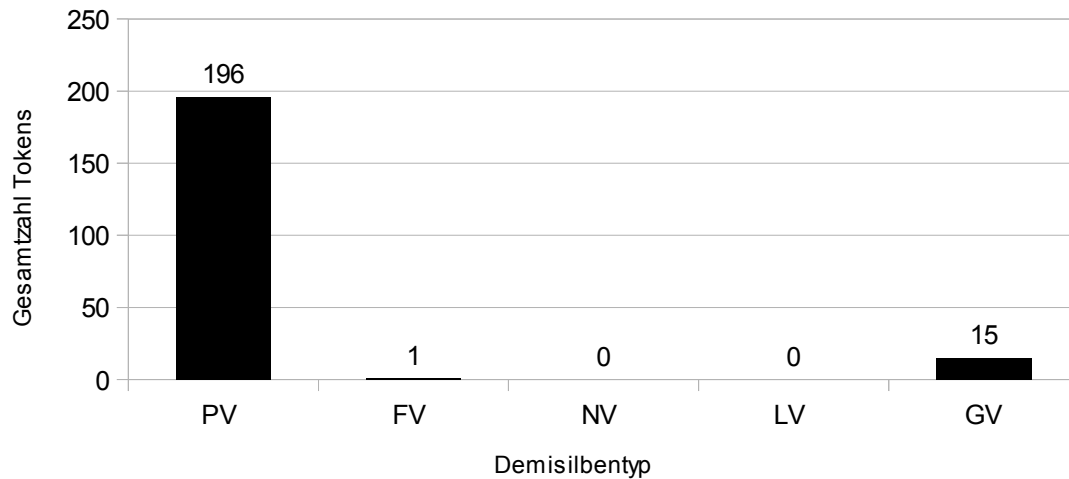


Abbildung 73: Häufigkeit initialer Demisilben in Protowörtern.

Für finale Demisilben in Protowörtern gilt analog zu den initialen Demisilben, dass die Demisilbentypen nicht gleich verteilt sind. Auch hier ist der Chi-Quadrat-Anpassungstests hoch signifikant ($\chi^2=442.127$, $df=5$, $p<0,0001$), wobei die Häufigkeiten zeigen, dass eine Dominanz von V- und VG-Demisilben vorliegt (vgl. Abb. 74; V = Vokal, VG = Vokal-Gleitlaut, VL = Vokal-Liquid, VN = Vokal-Nasal, VF = Vokal-Frikativ, VP = Vokal-Plosiv). Hiermit hat sich die Alternativhypothese (d) 2. H_1 als richtig erwiesen.

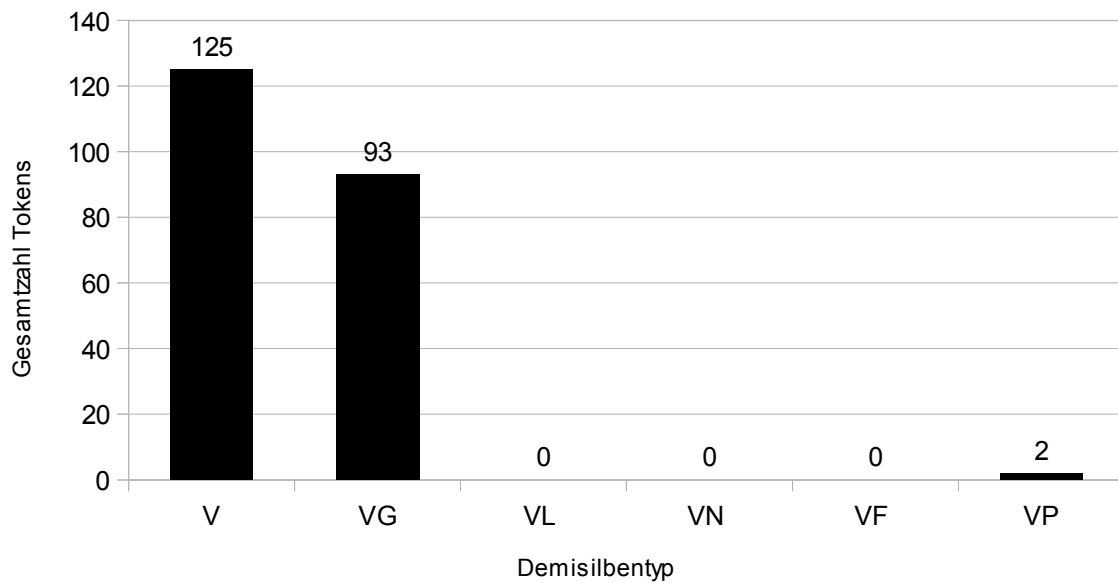


Abbildung 74: Häufigkeit finaler Demisilben in Protowörtern.

Führt man für initiale Demisilben in Zielwörtern einen Chi-Quadrat-Anpassungstests durch, so muss auch für diese verworfen werden, dass es eine gleichförmige Verteilung der Demisilbentypen gibt, somit Alternativhypothese (e) 1. H_1 stimmt. Auch hier findet sich ein hoch signifikantes Ergebnis ($\chi^2=531.882$, $df=4$, $p<0,0001$). Mit 227 Tokens von Plosiv-Vokal-Demisilben sind diese definitiv dominant, im Gegensatz zu dem Muster initialer Demisilbentypen in Protowörtern findet sich allerdings mit 46 Tokens von NV-Demisilben in Zielwörtern ein zusätzlich produzierter Demisilbentyp (vgl. Abb. 75).

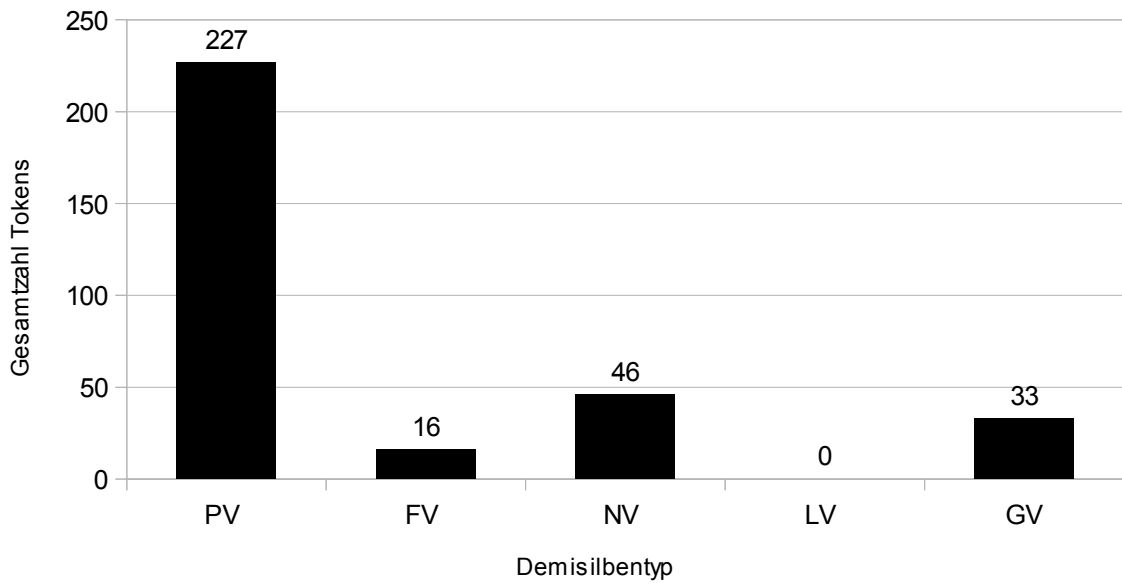


Abbildung 75: Häufigkeit initialer Demisilben in Zielwörtern.

Letztlich kann auch für die Produktion finaler Demisilben in Zielwörtern festgehalten werden, dass auch hier signifikant hoch keine Gleichverteilung von Demisilbentypen vorhanden ist ($\chi^2=886.103$, $df=5$, $p<0,0001$) und hiermit die Alternativhypothese (e) 2. H_1 durch die Daten eine Bestätigung gefunden hat. Vokal- und Vokal-Gleitlaut-Demisilben sind mit 250 bzw. 98 Tokens die klar am häufigsten auftretenden Typen in den Daten (vgl. Abb. 76).

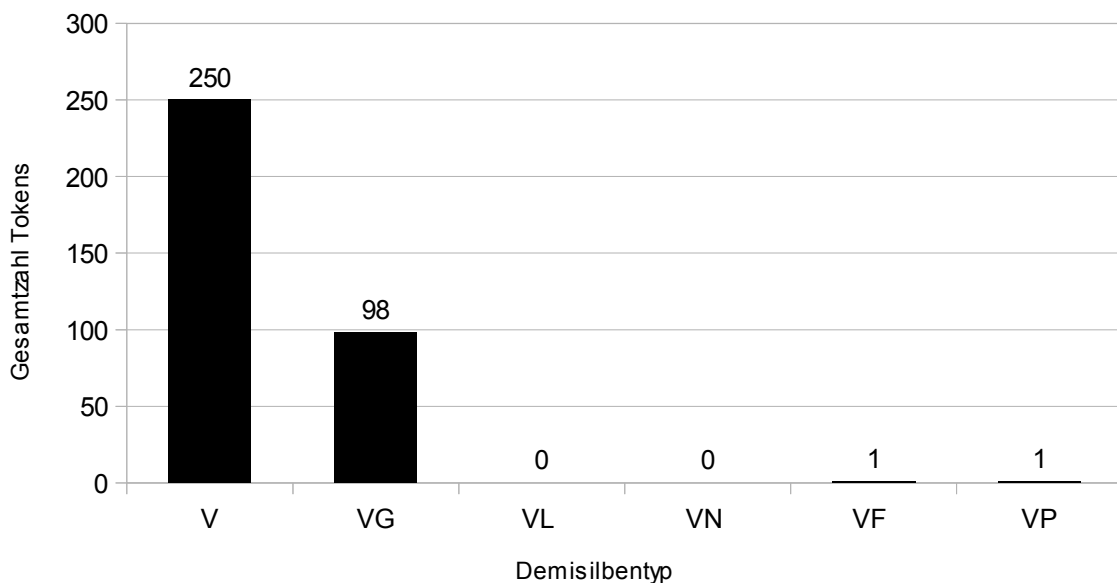


Abbildung 76: Häufigkeit finaler Demisilben in Zielwörtern.

Testet man auf die Unabhängigkeit der Produktion initialer Demisilbentypen von Proto- und Zielwörtern, so zeigt sich, dass die Produktion hoch signifikant unterschiedlich ist, ihr Zusammenhang besteht mit einer mittleren Effektstärke ($\chi^2=47,619$, $df=4$, $p<0,001$, Cramérs $V=0,30$). Damit ist die Alternativhypothese (f) 1. H_1 zum großen Teil verifiziert, da nur eine mittlere Effektstärke festgestellt wurde.

Für die Produktion von finalen Demisilben in Proto- und Zielwörtern lässt sich konstatieren, dass die Produktion nicht in gleichem Maße signifikant unterschieden und der Zusammenhang gering ist ($\chi^2=14,222$, $df=5$, $p<0,05$, Cramérs $V=0,16$). Hier kann Alternativhypothese (f) 2. H_1 nur eingeschränkt als bestätigt angesehen werden. Der Unterschied in der Produktion ist immer noch signifikant (bezogen auf das festgelegte Signifikanzniveau von $\alpha=0,05$), die geringe Effektstärke macht aber einen Zusammenhang unwahrscheinlich.

Auch wenn die Semantik von Proto- und Zielwörtern nicht zentraler Untersuchungsgegenstand dieser Arbeit ist, so lässt sich doch nach Maßgabe der methodischen Klassifikation eine Häufigkeitsdarstellung für die Produktion der semantischen Typen Personal, Interaktional, Instrumental und Regulativ angeben. Von insgesamt 444 Tokens sind 100% der Protowörter ($n=194$) mit dem semantischen Typ Personal kategorisiert. 92,4% der Zielwörter sind hingegen dem semantischen Typ Interaktional, 4,8% Instrumental und 2,8% Regulativ zugeordnet (vgl. Abb. 77).

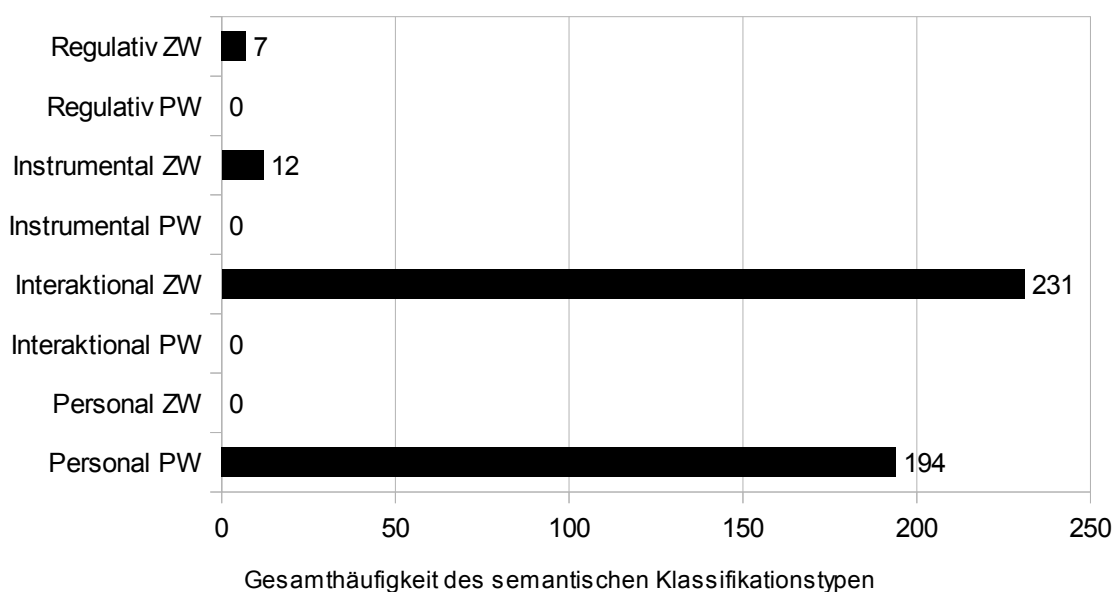


Abbildung 77: Produktion semantischer Klassifikationstypen in Proto- und Zielwörtern.

Damit ist die Auftretenshäufigkeit der semantischen Klassifikationstypen innerhalb von Protowörtern hoch signifikant nicht gleich verteilt ($\chi^2=582$, $df=3$, $p<0,0001$). Dies gilt ebenso für die selbige Auftretenshäufigkeit in Zielwörtern ($\chi^2=606,864$, $df=3$, $p<0,0001$). Damit ist sowohl die Alternativhypothese (g) 1. H_1 als auch die Alternativhypothese (g) 2. H_1 bestätigt. Das Auftreten von Zeigegesten beschränkte sich nur auf die Klasse der interaktionalen Zielwörter, wobei hier eine klare Prominenz in der Verwendung beobachtet werden konnte. Die Bestätigung der Alternativhypothese (g) 3. H_1 ist hier also eindeutig. 74,4% ($n=96$) der Zeigegesten wurden zusammen mit deiktischen Wortformen produziert, während 25,6% ($n=33$) der Zeigegesten auch mit anderen Zielwortformen auftraten. Allerdings konnte letztere Variante erst ab der siebten Aufnahmesitzung beobachtet werden (vgl. Abb. 78).

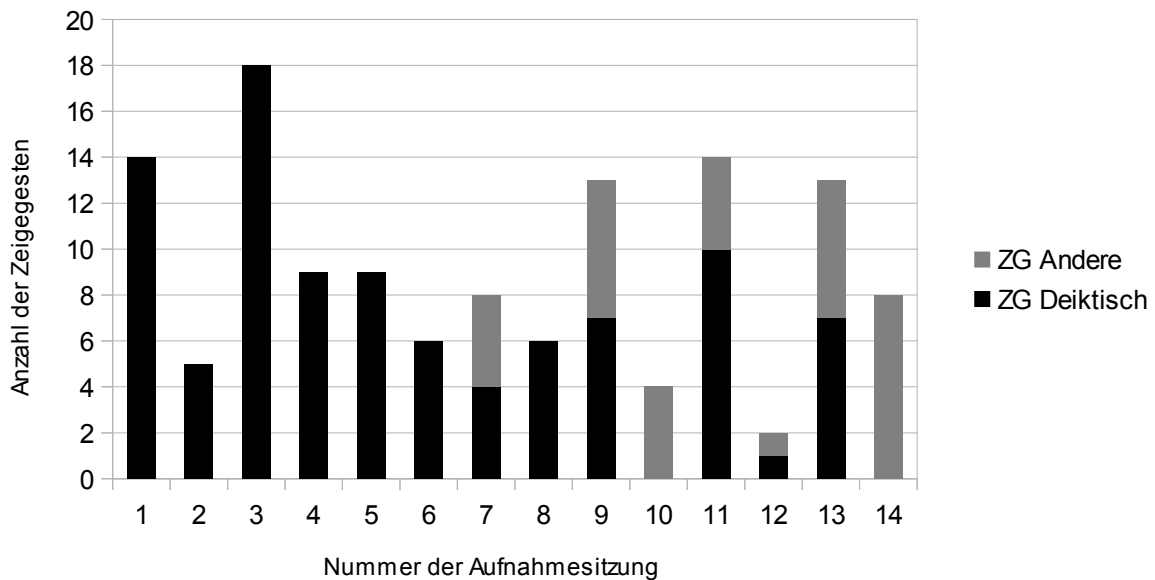


Abbildung 78: Verwendung von Zeigegesten in interaktionalen Zielwörtern.

Ein Chi-Quadrat-Anpassungstests bestätigt ein hoch signifikantes ungleiches Vorkommen von Zeigegesten in deiktischen Wortformen und anderen Wortformen ($\chi^2=30,767$, $df=1$, $p<0,0001$). Damit ist auch Alternativhypothese (g) 4. H_1 klar verifiziert.

4.4 Diskussion

Für die folgende Diskussion bietet sich an, von den Ergebnissen der quantitativen Untersuchung auszugehen und diese sukzessive mit den Ergebnissen der qualitativen zu verbinden.

In Kapitel 3.1.3 wurden verschiedene Untersuchungen vorgestellt, in denen die Protowortproduktion mit der Produktion von Zielwörtern negativ korrelieren (vgl. Dore et al. (1976); Robb et al. (1994)). Ein solcher Zusammenhang, d.h., dass durch die sukzessive Abnahme der Protowortproduktion die Produktion von Zielwörtern gleichfalls sukzessive zunimmt, konnte auch durch die Daten von J. bestätigt werden. Die interessante Frage hierbei ist natürlich, ob wir es im Sinne von Levelt (1998), Piske (2001) und Vihman (1996) mit ganzheitlichen phonetischen Formen zu tun haben oder ob Protowörter ab initio als genuin phonologisch strukturiert zu bewerten sind. Während die meisten Autoren Protowörter zum Großteil als CV-Reduplikationen klassifizieren, so dass ein fließender Übergang dieser Strukturen in die Zielwortproduktion angenommen wird, sprechen die Daten von J. gegen eine solche Interpretation. Während CV- und CVC-Silben in Protowörtern nicht signifikant unterschiedlich verteilt waren, konnte eine solch signifikante Verteilung innerhalb von Zielwörtern, mit einer deutlich höheren Produktion von CV-Silben (258 gegen 47 CVC-Silben), attestiert werden. Die quantitative Produktion der Silbentypen unterscheidet sich zwischen Proto- und Zielwörtern insgesamt signifikant, so dass nicht von einem einfachen Transfer der Protowortstrukturen in die Zielwortstrukturen gesprochen werden kann, vor allem auch deshalb nicht, weil markiertere CVC-Silben in Protowörtern sogar häufiger auftreten. Es ist dies m.E. ein erster deutlich Hinweis darauf, dass Protowörter innerhalb des Sprachsystems funktional anders bewertet werden als Zielwörter. Diese quantitativen Ergebnisse stützen insofern die qualitative Analyse, in der von Freischaltungseffekten durch chunkingbasierte Automatisierung, d.h., der Bildung von Abrufindizes im Kurzzeitgedächtnis ausgegangen wurde. Vor allem die häufigere Produktion von CVC-Strukturen in Protowörtern lässt sich unter dieser psycholinguistischen Perspektive dadurch motivieren, dass sie zur Bildung ambisilbischer Strukturen notwendig gegeben sein muss. Spezifische ambisilbische Strukturen in der Protowortproduktion wurden hier als Simulatoren benannt. In diesen wird mit den in der Produktion verfügbaren Repräsentationen zum Zeitpunkt z eine neue, aber überfrachtete Struktur y erzeugt, die letztlich als repräsentationaler Vorläufer für die Produktion von Zielwörtern gelten kann. Es war in den Daten ein ambisilbischer Nukleus, der

die reguläre Onset-Position einer zweiten Silbe innerhalb eines Zielwortes bereitstellte. Unter dieser Perspektive sitzt die Zielwortproduktion der Protowortproduktion zum Teil parasitär als Trittbrettfahrer auf, so dass auch die negative Korrelation durch die hier angenommene Vorläuferfunktion der Protowortproduktion eine Erklärung findet.

Auch Reduplikationen werden in der Protowortproduktion spezifisch und funktional verwendet. Es ist eben nicht der Fall, dass Protowörter hauptsächlich aus CV-Reduplikationen bestehen. Im Gegenteil wurde hier festgestellt, dass die Produktion von Reduplikationen in Proto- und Zielwörtern hinsichtlich der Vorkommenshäufigkeit signifikant ungleich ist. Ähnlich wie in Bezug auf den Realisierungszeitpunkt von CVC-Silben sind hier die Annahmen von Fikkert (1994) und Demuth et al. (1995) (vgl. Kapitel 2.2.1 und 2.2.2) nicht adäquat. Nach einem strikten 2D-Modell, aufbauend auf dem uniform wirksamen Konzept der prosodischen Lizenzierung mit Fuß- und Morarepräsentationen, dürfte kein Unterschied nach Worttyp (Zielwort vs. Protowort) in diesem frühen Erwerbsalter zu beobachten sein. Es konnte allerdings eine klare Dominanz von CV-Reduplikationen in Zielwörtern mit kurzzeitiger, ausschließlicher Produktion dieser Formen in Protowörtern beobachtet werden. Wie in Kapitel 2.2.4.1 schon ausgeführt wurde, unterstützen die hier vorliegenden Daten vielmehr die Annahme, dass Reduplikation als morphologischer Default-Prozess nur so lange systematisch verwendet wird, bis in der Produktion andere Formen etabliert sind. Im Übergang von Proto- zu Zielwörtern also so lange, bis spezifische Onsets von Zielwörtern etabliert sind, in der Produktion von Zielwörtern entsprechend bis die Insertion von Flexionsaffixen durch eine stabil etablierte CVC-Struktur möglich ist (vgl. Kap. 2.2.4.1). Auch die Natürliche Morphologie (vgl. Kap. 2.2.4), in der zwischen grammatischen, modularisierten und extragrammatischen, nicht-modularisierten Reduplikationen unterschieden und eine frühe prämorphologische, rein phonologische Reduplikation postuliert wird, kann die oben genannten Wortklassenunterschiede in der Verarbeitung nicht erfassen. Da hier ebenfalls eine uniform applizierte phonologische Reduplikation ohne modulare Interaktionen vorausgesetzt wird, kann die Natürliche Morphologie aus analogen Gründen wie 2D-Modelle der Phonologie die vorliegenden Daten nicht erklären.

Weisen alle bisher genannten Aspekte eindeutig darauf hin, dass Protowörter nicht einfach ganzheitliche phonetische Formen sind, so wird die phonologische Strukturiertheit von Protowörtern durch die Zugrundelegung von Clements Demisilbentheorie (vgl. Kap. 1.4.2) ersichtlich. Anhand der Verteilung kann klar konstatiert werden, dass in Protowörtern initial und final fast gänzlich nicht-komplexe Demisilben produziert werden, somit schon in einem

solch frühen Erwerbsalter eine systematische Produktion nach Maßgabe des Dispersion Principles stattfindet. Die finalen Demisilben der Zielwörter sind hinsichtlich der Komplexität von Demisilben in Protowörtern nicht unterschieden, initial zeigt sich allerdings ein signifikanter Unterschied. D.h., dass der Merkmalsausbau in Zielwörtern und damit die Zunahme der Demisilbenkomplexität im Onset beginnt, wobei Nasale die häufigste Produktionsvariante darstellen. Legt man die Sonoritätshierarchie von Clements (1990) zugrunde, sind Nasale auch die Lautklasse, deren Maß an Sonorität in einer kategorialen Klassifikation direkt über dem der Obstruenten liegt. Da allerdings auch Gleitlaute häufig initial produziert werden, kann hier keine eindeutige Übereinstimmung von Komplexitätsgrad und Erwerbsreihenfolge postuliert werden. Festgestellt werden kann aber, dass mit der Produktion von Protowörtern auch die Anwendung des Sonority Dispersion Principles stattfindet, somit in der nachfolgenden Zielwortproduktion nur eine phonologische Berechnung adaptiert und ausgebaut wird, die schon zuvor aktiv war.

Die Präferenz von nicht-komplexen initialen wie auch finalen Demisilbentypen zeigt auch, dass in der Erwerbslogik die Integration von Oberklassenmerkmalen primär ist und nicht die Spezifikationen für Einzelphoneme der Zielsprache. So müssen z.B. spezifizierende Artikulationsortmerkmale mediiert durch die Protowortproduktion erst eine Freischaltung für Silbenpositionen erfahren, wie z.B. das Merkmal [labial] in Sitzung 5 und 6 der qualitativen Analyse. Dies wäre insofern für die Produktion der Einzelphonem eine psycholinguistische Interpretation des Prozesses der De-Autosegmentalisierung wie in ihn Goldsmith (1976) und Spencer (1986) beschrieben haben, in dem Autosegmente zu lexikalisch gespeicherten, linear geordneten Segmentketten verschmolzen werden. Eine Konzentration auf eine einzelne Ebene, sei es nun Wort-, Silben- oder Merkmalsebene, kann diese komplexen Interaktionen im Erwerb auf keinen Fall wiedergeben. So muss z.B. die in Kapitel 3.2.2 genannte Anschauung von Velleman & Vihman (2002), nach der von ganzheitlichen Wortschablonen ausgegangen wird und jegliche früh zugänglichen Merkmale negiert werden, angesichts der hier dargelegten Ergebnisse zur Differenz von Proto- und Zielwörtern und ihrer spezifischen Berechnung als theoretisch schlicht simplifizierend und empirisch ungenügend qualifiziert werden. Sowohl solche Theorien, die von artikulatorischen Mustern oder phonetischen Schablonen ausgehen, als auch linguistisch orientierte Spracherwerbstheorien, in denen Phonologie in Nachfolge zum Strukturalismus mit der Realisierung zielsprachlicher Phoneme identifiziert wird, können die Systematik einer mehrdimensionalen Berechnung im frühen Spracherwerb aufgrund ihrer Prämissen hinsichtlich der phonologischen Verarbeitung nicht

erfassen. Beiden Forschungsströmungen kann vorgeworfen werden, dass ihr wissenschaftliches Analyseinstrumentarium bei weitem zu grobkörnig ist, als dass sie die Feinkörnigkeit phonologischer Subprozesse auch nur annähernd erfassen könnten.

Mit der Etablierung der Protowortproduktion als genuin formgetriebener Vorläufer der Zielwortproduktion wird auch die Annahme eines irgendwie gearteten kommunikativen oder intentionalen Zwecks von Protowörtern obsolet. Als wichtige Ergänzung zu den Ergebnissen der phonologischen Analyse können somit diejenigen der semantischen Analyse gelten.

Das eindeutige Ergebnis dieser Studie ist, dass Protowörter ausschließlich mit dem semantischen Merkmal PERSONAL produziert wurden und Zeigegesten niemals mit ihnen zusammen auftraten. Selbst wenn man die semantische Klassifikation in dieser Arbeit nicht annimmt, so bleibt als Ergebnis dennoch, dass Protowörter rein selbstbezüglich und ohne eine auf andere Personen gerichtete Intention produziert werden. So eindeutig die Datenanalyse ist, so eindeutig lässt es sich dieses Ergebnis m.E. unter Rückbezug auf die obigen Ausführungen und Schlussfolgerungen auch interpretieren: Die Produktion von Zeigegesten hat so gut wie keine Relevanz für die kognitive Berechnung sprachlicher Information, die in der Produktion von Protowörtern schon zuvor stattfindet, im Sprachverstehen zeitlich noch wesentlich früher. Mit sprachlicher Information ist hier ausdrücklich phonologische Information gemeint, die, wie in Kapitel 3.2.6 ausführlich und kritisch ausgeführt wurde, von Tomasello fast vollständig ignoriert wird.⁷⁵ Die Produktion von Zeigegesten hat allerdings auch bei J. dort eine Relevanz, wozu sie auch primär verwendet wird: zur Kommunikation. Und genau deshalb werden Zeigegesten auch mit solchen Wörtern verwendet, die auch primär zur Kommunikation verwendet werden, nämlich Zielwörtern, im Speziellen solche, die für das Merkmal INTERAKTIONAL spezifiziert sind. In für INTERAKTIONAL spezifizierten Wörtern sind es deiktische Varianten, die signifikant häufig mit Zeigegesten produziert werden. Aber angesichts der signifikanten Ergebnisse der fehlenden Zeigegesten in der Protowortproduktion lässt dieses gemeinsame Auftreten von Zeigegesten und Zielwörtern keine besonderen Schlussfolgerungen zu,⁷⁶ schon gar nicht solche, wie sie Tomasello, Carpenter & Liszkowski (2007: 720) vornehmen:

⁷⁵ In Tomasello (2003: 61-62) wird auf die Nähe zu den Annahmen von Vihman (1996) verwiesen und der Extraktionsprozess von Phonemen aus Wortschablonen als zielführend für den Erwerb insgesamt stark gemacht.

⁷⁶ Auch der Erwerb von Pronomen in der American Sign Language (ASL) spricht gegen die Annahmen Tomasellos (Petitto 1988). Hier werden Pronomen zeitgleich mit einfachen Zeigegesten produziert und das Pronomen „du“ (auf den Gesprächspartner deuten) anfänglich als „ich“ (auf sich selbst deuten) gebraucht. Pronomen sind also kein einfaches Deuten auf ein Objekt, sondern arbiträre Symbole, deren Referenz sich je nach Sprecher ändert und deren korrekte Verwendung erst nach einer gewissen Zeit gelingt.

„Pointing may thus represent a key transition, both phylogenetically and ontogenetically, from nonlinguistic to linguistic forms of human communication.“

Schon 1934 fasste Karl Bühler Ursprungstheorien in der Art Tomasellos als „Mythos vom deiktischen Ursprung der Sprache“ und kam zu einer klaren Einschätzung, die an dieser Stelle bekräftigt werden soll:

„Hindeuten ist Hindeuten und nie etwas mehr, ob ich es nun mit dem Finger oder zwiefach mit dem Finger und einem die Geste begleitenden Laute tue. Nein, der Fortschritt ist einzig und allein an die Bedingung geknüpft, daß der Laut etwas hinzubringt, etwas Neues an Leistung. Und wie immer man die Dinge auch drehen und wenden mag, so kann dieses Plus aus keiner anderen Quelle kommen als aus der Nennfunktion des Lautes. Auch eine stumme Gebärde kann das „Bedeutete“ charakterisieren, indem sie es nachbildet; der Laut symbolisiert es. In beiden Fällen ist der schlichte Hinweis auf ein da und da im Wahrnehmungsbereich zu Findendes durchaus zu trennen von der andersartigen Angabe, daß es ein so und so Beschaffenes sei.“ (Bühler 1934: 87)

Auch wenn bei Tomasello immer wieder betont wird, dass es die gemeinsame Aufmerksamkeit und mit ihr eine triadische Relation innerhalb dieses Aufmerksamkeitsrahmens ist, die die Anfänge menschlicher Kommunikation ausmachen, und nicht das isolierte Wahrnehmungsfeld einzelner kommunizierender Individuen, so ändert dies nichts am grundlegenden Faktum, dass der zielsprachlichen Produktion von phonologischen Strukturen die Produktion von Protowörtern vorangeht. Und diese sind, wie gezeigt wurde, letztlich nur zur Etablierung von Form innerhalb des sich ausbauenden Sprachsystems relevant und im eigentlichen Sinne einer internen kognitiven Informationsverarbeitung gegenüber sozialen und motivationalen Bezügen inert. Die Ergebnisse zur Protowortproduktion stützen insofern auch die in Kapitel 4.1.3 benannte Unabhängigkeit der (zielsprachlichen) phonologischen Entwicklung von demographischen Variablen (Dodd et al. (2003), Smit et al. (1990)).

Zusammen mit der in Kapitel 3.3 vorgestellten Fallanalyse des Falls P. weisen die hier vorliegenden Ergebnisse zur Sprachproduktion von Kind J. also vielmehr eindeutig darauf hin, dass einerseits Subkomponenten des Sprachsystems, hier die phonologische und die konzeptuelle Komponente, im frühesten Produktionsstadium unabhängig voneinander operieren und innerhalb der phonologischen Komponente selbst differente Bewertungen der jeweiligen informationellen Strukturen vorgenommen werden, d.h., die phonologischen Informationen von Protowörtern und Zielwörtern unterschiedlich gespeichert werden und sie in andere Verarbeitungsprozeduren eingehen. Unter letztere sind wiederum die spezifischen Simulatorfunktionen der Protowortproduktion gefasst. Im Rahmen des breiteren

kognitionswissenschaftlichen Diskurs', der in den Kapiteln 1.1, 1.2 und 1.3 besprochen wurde, kann man hieraus m.E. äußerst gewichtige Schlussfolgerungen ziehen.

Zum einen muss man konstatieren, dass schon in der frühen Sprachplanung eine modulare Verarbeitung stattfindet. Das Phonologie-Modul berechnet in der Produktion von Anfang an nach sprachsystematischen Prinzipien wie eben dem Sonority Dispersion Principle. Protowörter werden vor Zielwörtern produziert, wobei letztere als Trittbrettfahrer auf den Strukturen von ersteren aufbauen und sukzessive an Komplexität zunehmen, wie man an der Produktion von silbeninitialen Nasalen beobachten konnte. Da Protowörter zu einem Zeitpunkt produziert werden, zu dem passiv schon wesentliche Strukturen der Zielsprache repräsentiert sind, muss man mit Hale & Reiss (2008) die Unterscheidung zwischen Performanzfaktoren und Kompetenz annehmen und den Erwerb als sukzessive Integration von sprachlicher Information oder Verringerung der limitierenden Performanzfaktoren ansehen. Wesentliche Annahmen aus Kapitel 2.2.1 konnten in diesem Zusammenhang durch die Daten von J. bekräftigt werden, so vor allem, dass auch die frühe phonologische Produktion dem Prinzip der minimalen Struktur von Weissenborn (2000) folgt. Die Produktion von nicht-komplexen Demsilben kann hier als Analogon zu reduzierten morphosyntaktischen Strukturen in der Produktion gelten, so dass man Weissenborns Formulierungen (ebd.: 158; vgl. auch Kap. 2.2.1) für die frühe phonologische Produktion mit einer kleinen Modifikation übernehmen kann, nämlich,

„dass das Kind, solange es nicht die Anwendungsbedingungen einer zielsprachlichen Regel herausgefunden hat, vorzugsweise eine Äußerungsstruktur wählt, die unter Anwendung des zu diesem Zeitpunkt schon erworbenen zielsprachlichen grammatischen Wissens, den geringsten Aufwand an morphosyntaktischen [**oder phonologischen** – Anm. d. Verf.] Prozessen erfordert.“

Da Protowörter nach Maßgabe des Sonority Dispersion Principles aus nicht-komplexen Demsilben aufgebaut sind, demnach entgegen einer grundlegend artikulatorischen Einschränkung phonologisch strukturiert und dennoch ohne overten Bezug zu Zielwortrepräsentationen sind, wurden in dieser Arbeit als wesentliche Performanzlimitierungen die Operationen des Arbeits- wie auch des Kurzzeitgedächtnisses angenommen. Im Arbeitsgedächtnis werden sprachliche Informationen kombiniert und im Kurzzeitgedächtnis Hinweise für den schnelleren Abruf etabliert (vgl. Kap. 1.6). Diese Modellierung konnte durch die qualitative Analyse bekräftigt werden, in der verschiedene Freischaltungseffekte für Silbenpositionen und phonologische Merkmale in der Zielwortproduktion durch die vorherige Produktion von spezifischen Protowörtern

festgemacht werden konnten. Eine solche disparate Produktion mit differenter Speicherung und sukzessiver Elaboration kann nur unter Zugrundelegung einer klassischen Architektur mit der Trennung von Programm und (verschiedenen) Datenspeichern einer Erklärung finden. In Kapitel 1.2.3 wurde erläutert, dass in konnektionistischen Modellen keine Kombinatorik vorhanden ist und es dementsprechend keine Struktursensitivität gibt. Es ist in solchen Modellen entsprechend nicht möglich, die Produktion kombinationsbedingter, überfrachteter Strukturen, wie sie hier und von Tracy (2000) und Hohenberger (2002) beobachtet wurden, zu erklären. Dies gilt auch für die festgestellten Freischaltungseffekte, da hierfür die in der Struktur enthaltenen Silbenpositionen und phonologischen Merkmale symbolisch enkodiert sein müssten.

Mit der hier gegebenen Evidenz für eine modulare Verarbeitung im Rahmen einer klassischen Architektur nach Fodor & Pylyshyn (1988) kann entsprechend auch gegen jegliche epigenetischen oder emergenzbasierten Ansätze argumentiert werden. Die symbolischen Operationen und mehrdimensionalen Strukturinteraktionen nach Maßgabe des Prinzips der minimalen Struktur sind mit einer linguistischen Analyse kategorial klar erfassbar und modelltheoretisch rekonstruierbar. Eine neurobiologische „Fuzziness“, die auf dieser wissenschaftlichen Beschreibungsebene gegeben sein mag, spiegelt sich definitiv nicht auf der Ebene einer psycholinguistischen Analyse und Rekonstruktion wider. Das OIP von Poeppel & Embick (2005; vgl. Kap. 1.2.2) muss dementsprechend hier noch einmal als bestätigt angesehen werden.

Zu guter Letzt bleibt dennoch die Frage nach der neuronalen Implementation des Sonority Dispersion Principles. In Kapitel 3.4.1 wurde die Annahme gemacht, dass der Random Generator auch aktiv ist, wenn keine oder nur wenige Zielformen im Output-Lexikon gespeichert sind und wir innerhalb des Sprachproduktionssystems von Jargon-Aphasikern und Kindern im frühen Spracherwerb eine analoge Situation vorfinden, nämlich das Fehlen phonologischer Einträge im mentalen Output-Lexikon. Entsprechend wurde die Vorhersage gemacht, dass sich Protowörter ebenfalls aus Demisilben mit geringer Komplexität konstituieren, was hier auch bestätigt wurde.

Diese Vermutung, dass der Random Generator auch im Spracherwerb aktiv sein könnte, wurde auch schon von Buckingham (1990), ohne den Verweis auf Protowörter, gemacht:

„I believe that the generator should be considered to produce syllable-sized units rather than individual segments. Any subsequent substitutions or linear transpositions of the segments constituting the neologism will be accounted for by scan copier derailments. The stock of permissible syllables that the device produces would be precisely those of the speaker's

language, and consequently the phonotactic constraints would come ready made. Abstracted phonotactic information is most assuredly part of the normal cognitive system. This dissociated knowledge is likely set up during the course of language development by running inductions over the stock of syllable types that comprise the words of the language in order to arrive at the regularities that will be useful in certain circumstances later on. (ebd.: 219)

Diese Ausführungen müssen m.E. weiter spezifiziert werden. In Kapitel 1.5.1 wurde in Abgrenzung zu dem Slots-and-fillers-Modell von Shattuck-Hufnagel (1979, 1983) die Splitting-Hypothese entworfen. In dem Modell von Shattuck-Hufnagel werden die Silbenpositionen und Segemente unabhängig voneinander generiert. Im Gegensatz dazu wurde mit der Splitting-Hypothese ein Ansatz entworfen, in dem es vielmehr um die Integration von Merkmalen in spezifische X-Positionen der Silbe ging:

Splitting-Hypothese

Das Merkmal [silbisch], die Oberklassenmerkmale und die jeweils spezifischen Artikulationsort- und Artikulationsartmerkmale einer segmentalen Menge werden zeitlich konsekutiv in drei Stufen an definierte X-Positionen assoziiert.

Wenn Buckingham also davon ausgeht, dass der Random Generator Silben produziert, so muss hier betont werden, dass durch die festgestellte Wirksamkeit des Sonority Dispersion Principles in Protowörtern wie auch in den Produktionen von Jargon-Aphasikern, vielmehr ein minimales Set an silbischer wie auch segmentaler Information zeitgleich generiert wird. Die Funktion des Scan Copiers wird somit durch die Prozedur der Merkmalsintegration ersetzt. Bezieht sich die Splitting-Hypothese also genau auf diese Prozedur für die On-line-Produktion insgesamt, so kann weiter angenommen werden, dass im frühen Spracherwerb autosegmentale Information für Artikulationsart und -ort noch nicht stabil integriert ist und entsprechend lose assoziiert wird, was letztlich auch zu einer Variabilität in den overten Formen führt.

Da für Produktion von Proto- und Zielwörtern eine negative Korrelation festgestellt wurde, ist anzunehmen, dass vom Random Generator generierte Protowörter als vorläufige Substitute für fehlende Zielwörter vorläufig im Output-Lexikon abgespeichert werden und hier mit Zielwortrepräsentationen um Speicherplatz konkurrieren. Während der Zeit ihrer Speicherung sind sie hierbei aber auch modifizierbar. Für die erstmalige Abspeicherung wie auch für die Möglichkeit der sukzessiven Erweiterung von Repräsentationen muss man eine Input-Output-Schleife, in der der repräsentationale Output der Sprachproduktion den repräsentationalen Input des eigenen Sprachverstehens stellen kann, annehmen, so wie sie z.B. im Modell von Carruthers (2006) gegeben ist (vgl. Kap. 3.2.3).

Dass verschiedene Wortformtypen lexikalisch gespeichert werden, dürfte im Übrigen aus einer cross-linguistischen Perspektive nicht überraschen. Das japanische Lexikon ist z.B. stärker als andere Sprachen stratifiziert. Hier findet man vier morphologische Klassen (vgl. Itô & Mester 1995), Yamato (der native Wortschatz), Sino-Japanisch (Lehnwörter, die in der Zeit vom 7.-13. Jahrhundert n. Chr. aus dem Chinesischen übernommen wurden), Fremdwörter (Lehnwörter, die ab dem 16. Jahrhundert n. Chr. bis heute fast ausschließlich aus europäischen Sprachen übernommen wurden) und mimetische Formen. Letztere werden auch als laut-symbolische (engl. sound-symbolic) Elemente, Ideophone oder Onomatopoeia klassifiziert. Sie bilden im Japanischen mit über 2000 Wörtern das zweitgrößte Vorkommen in den Sprachen dieser Welt nach dem Koreanischen (Noma: 1998: 30, zit. nach Ivanova 2006). Im Gegensatz zum Englischen oder anderen europäischen Sprachen findet man also im Japanischen eine weitaus häufigere Verwendung von mimetischen Formen, die je nach Kulturraum unter anderen linguistischen Termini erfasst werden (vgl. Akita 2009: 10). Dies führte zu einer geringen Beachtung innerhalb der linguistischen Forschung:

„[...] English and European languages, on which a great deal of modern linguistic theories have customarily built on do not employ extensive sets of expressive words, with sporadic instances as *helter-skelter* or *bow-wow*. It seems undeniable that this latter fact has contributed to the scarcity of more scientific treatments about mimetics.“ (Tsujimura 2005: 371)

Unklar ist, in welcher Form die o.g. lexikalischen Strata separiert sind. Nach McCawley (1968) wäre z.B. eine Variante ein Sublexikon-Modell, in der alle Strata unabhängige phonologische Systeme beinhalten:

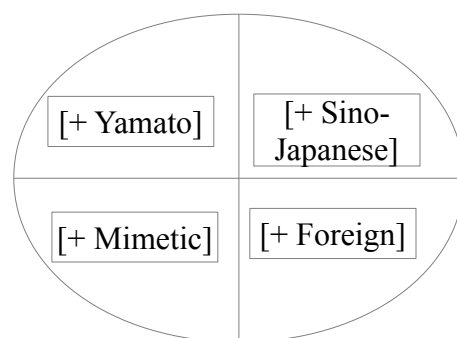


Abbildung 79: Sublexikon-Modell des Japanischen.

Man kann aus diesen cross-linguistischen Betrachtungen wichtige Hinweise für eine universell ausgerichtete Spracherwerbsmodellierung ziehen. Die Option eines stratifizierten

Lexikons muss universeller Bestandteil des Initialzustands des L1-Erwerbs zu sein und es erscheint deshalb durchaus plausibel, dass eine solche Option in der unterschiedlichen Speicherung von Proto- und Zielwörtern realisiert wird, aber relativ schnell verworfen wird. Die Frage danach aber, warum nicht für alle Kinder Protowortproduktionen dokumentiert sind, kann an dieser Stelle nicht abschließend beantwortet werden. Die vorliegende Untersuchung nutzte methodische Verfahrensweisen, die in Kapitel 4.1 genauer dargestellt wurden. Eine Spontansprachuntersuchung mit der durchgeführten Aufnahmefrequenz ist relativ zeitaufwändig. Auch bei J. wurden Protowörter nur in einem sehr kurzen Zeitraum produziert, so dass sie mit niedrig-frequenten Untersuchungszeitpunkten mit relativ kurzer Aufnahmedauer nicht zu erfassen sind. Da Protowörter als Datenklasse zudem individuell zu analysieren sind, können auf Vergleiche abzielende Gruppenstudien somit kaum zu relevanten Ergebnissen führen. Es mögen also diese methodischen Voraussetzungen zusammen mit der in dieser Arbeit ausführlich kritisierten Zielwortorientierung sein, die der bisherigen ausführlichen Untersuchung von Protowörtern im Wege standen.

5 Schlussbetrachtungen

In der vorliegenden Arbeit wurde die erste umfassende und linguistisch basierte Untersuchung von Protowörtern vorgestellt. Es gilt am Ende dieser Arbeit die gewonnenen Einsichten in eine prägnantere Definition dieses wichtigen Worttyps einmünden zu lassen. In Kapitel 3.1.3 wurde bereits auf die häufig zitierte Definition von Vihman (1996) verwiesen:

„We will use the term [...] to refer to relatively stable child forms with relatively consistent use which lack any clear connection with the form + meaning unit of a conventional adult model.“

Diese Definition wird sowohl der Funktion als auch der mehrdimensionalen Strukturiertheit von Protowörtern keineswegs gerecht. Nach Maßgabe der Untersuchung der Daten von Kind J. soll hier zuerst eine verkürzte Definition vorgestellt werden, die nur für den ungestörten, frühen Spracherwerb gilt, dennoch aber die wesentlichen Ergebnisse impliziert:

Protowortdefinition (Kurzfassung für den ungestörten frühen Spracherwerb)

Protowörter im frühen Spracherwerb sind über einen kurzen Zeitraum kontinuierlich produzierte Formen, deren wesentliche Charakteristikum ist, dass eine overte (phonetische) Ähnlichkeit zu Zielwörtern nicht gegeben ist. Sie spiegeln strukturell die konsekutive und

systematische Integration von phonologischen Merkmalen in Silbenpositionen wider und bestehen gemäß dem Sonority Dispersion Principle aus nicht-komplexen Demisilben. Bei Auftreten können sie durch den Ausbau phonologischer Strukturen spezifische Bootstrapping-Funktionen für Zielwörter ausüben. Protowörter werden im Verlauf des Spracherwerbs durch das Informationskorrelat [(Motorische Struktur) + Phonologische Struktur (\pm silbisch; [Oberklassenmerkmale]) + Bedeutung] konstituiert, wobei sie semantisch-pragmatisch nicht-intentional verwendet werden.

In Kapitel 3.4.2 wurde unter Zugrundelegung der Parallelen Architektur von Jackendoff (2002) eine grobe Schematik entwickelt, um den lexikalischen Eintrag von Protowörtern als Informationskorrelat verschiedener Domänen zu klassifizieren:

	Motorische Struktur	Phonologische Struktur	Morphosyntaktische Struktur	Bedeutung
PWIS-1	+	\pm silbisch	-	+
PWIS-2	+	(\pm silbisch) \pm sonorant \pm approximant \pm vocoid	-	+
PWIS-3	+	\pm silbisch \pm sonorant \pm approximant \pm vocoid	+	+

Auch wenn die Einteilung sehr grob ist, so lässt sich mit ihr doch der wesentliche Unterschied zwischen den lexikalischen Einträgen der Kinder P. (vgl. Kap. 3.3) und J. deutlich machen. Während P. in der Produktion Protowörter in reduzierten aber zieladäquaten morphosyntaktischen Strukturen mit zielsprachlicher Bedeutung ohne diagnostizierte motorische Einschränkungen realisiert, werden eben solche morphosyntaktischen Strukturen von J. noch nicht produziert und die Bedeutung kann nicht als zieladäquat bestimmt werden. Für eine grundlegende Klassifikation des Phänomens Protowort ist dieser Unterschied kein unwichtiges Faktum, da es noch einmal verdeutlicht, dass die wesentlichen Repräsentationen, die ein Protowort konstituieren, ausschließlich im Phonologie-Modul zu verorten sind. Werden diese Repräsentationen beibehalten, d.h., findet eine Plateaubildung innerhalb der phonologischen Komponente statt, ist es dennoch möglich, dass Schnittstellen zu anderen Modulen und ihren potentiell ziel- und altersadäquaten Repräsentationen etabliert werden. Das Gesamt dieser Schnittstellen und ihrer Repräsentationen bildet dementsprechend den lexikalischen Eintrag des jeweiligen Protoworts, so dass man klar die frühe, ungestörte und von einer durch eine (spezifische) Spracherwerbsstörung persistierende Protowortproduktion

unterscheiden kann und dies für diagnostische Zwecke auch dezidiert sollte. Im Übrigen zeigt ein Vergleich der Produktionen von P. und J. auch, dass motorisch-artikulatorische Einschränkungen kein relevanter Faktor sein können. Die Protowortproduktionen beider Kinder war gemäß dem Sonority Dispersion Principle genuin phonologisch strukturiert, allerdings wurden bei P. keine motorisch-artikulatorischen Beeinträchtigungen diagnostiziert. Die Plateaubildung von P. innerhalb der phonologischen Produktion bietet technisch gesprochen insofern ein Standbild des kognitiven Störungsherds in statu nascendi. Unter Einbeziehung dieser multi-modularen und klinisch relevanten Perspektiven, kann eine sehr umfassende Definition von Protowörtern gegeben werden, die weiterhin als eine wesentliche Zusammenfassung der Hauptergebnisse gelten kann:

Protowortdefinition (ausführliche Fassung, inklusive möglicher Plateaubildung)

Protowörter sind kontinuierlich produzierte Formen, deren wesentliches Charakteristikum ist, dass eine overt (phonetische) Ähnlichkeit zu Zielwörtern nicht gegeben ist.

Unter Zugrundelegung der Splitting-Hypothese spiegeln Protowörter die konsekutiv und systematische Integration von phonologischen Merkmalen in Silbenpositionen wider, wobei in der Protowortproduktion das Merkmal [±silbisch] wie auch die Oberklassenmerkmale integriert sind. Bei gegebenen Oberklassenmerkmalen erfolgt die Produktion regelhaft nach Maßgabe des Sonority Dispersion Principle, so dass nicht-komplexe Demisilben präferiert produziert werden. Artikulationsort- und Artikulationsartmerkmale werden hier zufällig und inklusive generiert, was die fehlende overt Ähnlichkeit zu Zielwörtern bedingt.

Für die Generierung der Demisilben ist ein Zufallsgenerator anzunehmen, der bei blockiertem Zugang zu Zielformen wieder aktiv werden kann.

Der beschriebene Intergrationsprozess erfolgt strikt modular und rein formbezogen, der Abgleich zu Zielwörtern spielt demgemäß keine Rolle. Allerdings kann in der Produktion von Zielwörtern auf generierte Strukturen der Protowortproduktion zurückgegriffen werden, so dass durch diese spezifische Bootstrapping-Funktionen erfüllt werden.

Protowörter werden im Verlauf des Spracherwerbs divers durch die Informationskorrelate

a) [(Motorische Struktur) + Phonologische Struktur ([±silbisch]) + Bedeutung] (PWIS-1)

a) [(Motorische Struktur) + Phonologische Struktur ([±silbisch; [Oberklassenmerkmale]]) + Bedeutung] (PWIS-2)

b) [(Motorische Struktur + Phonologische Struktur (vgl. a)) + Morphosyntax + Bedeutung] (PWIS-3),

konstituiert, wobei das jeweilige Informationskorrelat

a) semantisch-pragmatisch nicht-intentional verwendet wird und

b) nach Maßgabe des Entwicklungsstandes altersadäquat und voll symbolisch (d.h. typenbasiert) sein kann.

Die via regia zu dieser abschließenden Definition war letztlich die Zugrundelegung einer abstrakten „Galilean-Style Phonology“ (Hale 2007; vgl. Kap. 1.1) mit deren Hilfe erst die Reichhaltigkeit an strukturellen und prozessualen Nuancen der Protowortproduktion herausgearbeitet werden konnte. Es bleibt an dieser Stelle zu hoffen, dass in dieser

Untersuchung der Tiefenstrukturen der Protowortproduktion geleistet wurde, was Goethe in Bezug auf Galilei als das Wesentliche nennt:

„Alles kommt in der Wissenschaft auf das an, was man ein Apercu nennt, auf ein Gewährwerden dessen, was eigentlich den Erscheinungen zum Grunde liegt. Und ein solches Gewährwerden ist bis ins Unendliche fruchtbar.“ (Goethe 1948)

Abgesehen vom wissenschaftlichen Ertrag, ist es hier zum Abschluss ein ausdrückliches Movens, auf die ethische Dimension dieser Arbeit im Rahmen der Generativen Grammatik hinzuweisen. Konzepten wie Modularität, I-Sprachen-Linguistik, Angeborenheit von repräsentationalen Strukturen und Merkmalen ist eine solche Dimension immer schon immanent. Der paradigmatische Gegensatz zu Ansätzen, die die determinierende Formkraft sozialer Kontexte betonen, ist für die Generative Grammatik wesentlich, wie Chomsky eindringlich betont:

„Wäre der Mensch ein unendlich formbares, vollkommen plastisches Wesen, ohne angeborene geistige Strukturen, ohne die inhärenten Bedürfnisse eines kulturellen oder sozialen Charakters, wäre er in der Tat ein geeignetes Objekt für die „Verhaltensformung“ durch Staat, Unternehmen, Technokraten oder Zentralkomitees.“ (Chomsky 1987: 154)

Wie bereits in Dümig & Leuninger (2013) herausgestellt, ist es vor allem im therapeutischen Kontext die Etablierung der Generativen Grammatik und einer wirklichen phonologischen Expertise dringendst von Nöten. Als gravierendes Beispiel sei hier die gängige Diagnose einer sogenannten Inkonsequenten phonologischen Störung genannt. Eine Testung mit der PLAKKS durch den 25-Wörter-Test (vgl. Fox 2005) brächte bei jedem Kind mit einer Plateaubildung auf Protowortebene durch das verlangte Nachsprechen von Zielwörtern eine rein motorische Nachahmung hervor, die sich dementsprechend willkürlich und instabil gestalten würde. Eine zielwortorientierte Prozessanalyse erzeugt also, gerade weil sie nicht vom sich entwickelnden phonologischen System des Kindes bzw. der Kenntnis desselben ausgeht, einen „inkonsequenten“ Output. Die therapeutische Folgerung einer derart fatalen Diagnose ist, dass das Kind zehn Wörter immer wieder sagen bzw. nachsprechen muss, bis die Aussprache im Sinne der gewünschten Zielwörter konsequent ist. Wir haben hier in Form eines in der Logopädie weithin akzeptierten und populären Diagnostik- und Therapieverfahren m.E. unmittelbar vor Augen, welche schädigende Wirkung die Orientierung an einer E-Sprachen-Linguistik haben kann, bei der die Reichhaltigkeit, Feinkörnigkeit und Kreativität kognitiver Systeme komplett ignoriert wird.⁷⁷

⁷⁷ In Dümig & Leuninger (2013) sind Therapieprinzipien formuliert, die genau diesen Aspekten wie auch der Entwicklungslogik Form vor Bedeutung Rechnung tragen.

Entsprechend soll am Ende dieser Arbeit ein Plädoyer für die Revitalisierung des ursprünglichen Impetus und mit ihm der gesellschaftlichen Sprengkraft der Generativen Grammatik stehen, in der die Analyse kognitiver Systeme und ihrer Strukturen unmittelbar mit der der *conditio humana* und einer positiven Bestimmung derselben verbunden war:

„Wer ein bisschen Vertrauen in die Menschheit hat, hofft, dass es sich anders verhält, und versucht, die angeborenen Eigenschaften zu bestimmen, die den Rahmen für die geistige Entwicklung, die Entfaltung von moralischem Bewusstsein, kulturellen Leistungen und Teilhabe an einer freien Gemeinschaft bildet.“ (Chomsky 1987: 154)

6 Literaturverzeichnis

- Ahnert, L. (Hrsg.) (2014). *Theorien der Entwicklungspsychologie*. Heidelberg, Berlin: Springer VS.
- Akita, K. (2009). *A Grammar of Sound-Symbolic Words in Japanese: Theoretical Approaches to Iconic and Lexical Properties of Mimetics*. Ph.D. Thesis. Kobe: Kobe University.
- Ament, W. (1902). *Begriff und Begriffe der Kindersprache*. (Sammlungen von Abhandlungen aus dem Gebiete der pädagogischen Psychologie und Physiologie, V: 4.). Berlin: Reuther & Richard.
- Anderson, S. & Kiparsky, P. (Hrsg.) (1973). *A festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston.
- Archibald, J. (1995). The acquisition of stress. In J. Archibald (Hrsg.), *Phonological acquisition and phonological theory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 81-109.
- Archibald, J. (Hrsg.) (1995). *Phonological acquisition and phonological theory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Aronoff, M. & Oehrle, R. (Hrsg.) (1984). *Language Sound Structure*. Cambridge, MA: MIT Press.
- Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Hrsg.), *The psychology of learning and motivation: Advances in research and theory, Vol. II*. New York: Academic Press, 89-195.
- Auroux, S., Koerner, E. F. K., Niederehe, H.-J. & Versteegh, K. (Hrsg.) (2001). *Geschichte der Sprachwissenschaften. Ein internationales Handbuch zur Entwicklung der Sprachforschung von den Anfängen bis zur Gegenwart*, Teilband 2. Berlin: de Gruyter.
- Baayen, R. H., Piepenbrock, R. & van Rijn, H. (1993). *The CELEX Lexical Database (CDROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience* 4, 829–839.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. H. Bower (Hrsg.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 89-195.
- Baillargeon, R., Spelke, E. & Wasserman, S. (1985). Object permanence in 5-month old infants. *Cognition* 20, 191-208.
- Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology* 23, 655–664.
- Bakker, P. & Parkvall, M. (2005). Reduplication in pidgins and creoles. In B. Hurch (Hrsg.), *Studies on Reduplication*. Berlin: Mouton de Gruyter, 511-531.
- Baldwin, D. A., Baird, J. A., Saylor M. M. & Clark, M. A. (2001). Infants parse dynamic actions. *Child Development* 72, 708–717.
- Ball, M. & Kent, R. (Hrsg.) (1997). *The New Phonologies: Developments in Clinical Linguistics*. San Diego, CA: Singular Publishing Group.
- Baptista, B. O. & Silva Filho, J. L. A. (1997). The influence of markedness and syllable contact on the production of English final consonants by EFL learners. In J. Leather. & A. James (Hrsg.), *New sounds 97: Proceedings of the third*

- international symposium on the acquisition of second-language speech.* Klagenfurt: University of Klagenfurt, 336-350
- Barrett, M. D. (1981). The communicative functions of early child language. *Linguistics* 19, 273-305.
- Barrett, M. D. (1986). Early semantic representation and early word-usage. In S. A. Kuczaj & M. D. Barrett (Hrsg.), *The development of word meaning*. Berlin: Springer, 39-67.
- Bartke, S. (1998). *Experimentelle Studien zur Flexion und Wortbildung*. Tübingen: Niemeyer.
- Basser, L. S. (1962). Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. *Brain* 85, 427-460.
- Bates, E., Bretherton, I. & Snyder, L. (1988). *From first Words to Grammar. Individual differences and dissociable mechanism*. New York, Cambridge: University Press.
- Bates, E., Thal, D., Trauner, D., Fenson, J., Aram, D., Eisele, J. & Nass, R. (1997). From first words to grammar in children with focal brain injury. *Developmental Neuropsychology* 13(3), 275-334.
- Beckman, J., Walsh-Dickey, L. & Urbanczyk, S. (Hrsg.) (1995). *University of Massachusetts Occasional Papers in Linguistics: Papers in Optimality Theory*. Amherst: GLSA.
- Bell, A. & Hooper, J. B. (Hrsg.) (1978). *Syllables and segments*. New York: Elsevier
- Belletti, A. & Rizzi, L. (Hrsg.) (2002). *Noam Chomsky: On nature and language*. Cambridge: Cambridge University Press.
- Berent, I. (2013). *The Phonological Mind*. Cambridge: Cambridge University Press.
- Bernhardt, B. H., Gilbert, J. & Ingram, D. (Hrsg.) (1996). *Proceedings of the UBC international conference on phonological acquisition*. Somerville, MA: Cascadilla Press.
- Bernhardt, B. H. & Stemberger, J. P. (1998). *Handbook of phonological development: From the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.
- Bernhardt, B. H., & Stemberger, J. P. (2000). *Workbook in nonlinear phonology for clinical application*. Austin, Texas: PRO-ED.
- Berwick, R. (1986). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.
- Bickerton, D. (2009). *Adam's Tongue: How Humans made Language, How Language made Humans*. New York: Hill and Wang.
- Bierwisch, M. (1970). Semantics. In J. Lyons (Hrsg.), *New Horizons in Linguistics*. Harmondsworth: Penguin, 166-184.
- Bittner, D. (1998). Entfaltung grammatischer Relationen im NP-Erwerb: Referenz. *Folia Linguistica* XXXI (3-4), 255-283.
- Bittner, D. (1999). Erwerb des Konzepts der Quantifikation nominaler Referenten im Deutschen. In J. Meibauer & M. Rothweiler (Hrsg.), *Das Lexikon im Spracherwerb*. Tübingen, Basel: Francke, 51-74.
- Blanken, G., Dittmann, J. & Wallesch, C.-W. (Hrsg) (1988). *Sprachproduktionsmodelle. Neuro- und psycholinguistische Theorien der menschlichen Spracherzeugung*. Freiburg: Hochschulverlag.
- Blanken, G., Wallesch, C. W. & Papagno, C. (1990). Dissociations of language functions in aphasics with speech automatisms (recurring utterances). *Cortex* 26, 41-63.
- Blevins, J. (1993). The nature of constraints on the non-dominant hand in ASL. In G. Coulter (Hrsg.), *Current issues in ASL phonology*. New York: Academic Press, 43-62.
- Blevins, J. (1995). The Syllable in Phonological Theory. In J. A. Goldsmith (Hrsg.), *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell Publishers, 206-244.
- Bloom, L. (1973). *One word at a time*. The Hague: Mouton.
- Bloomfield, Leonard (1933). *Language*. New York: Holt, Rinehart and Winston.

- Blumstein, S. E., Alexander, M. P., Ryalls, J. H., Katz, W. & Dworetzky, B. (1987). On the nature of the foreign accent syndrome: A case study. *Brain and Language* 31, 215-244.
- Blumstein, S.E. & Kurowsky, K. (2006). The foreign accent syndrome. A perspective. *Journal of Neurolinguistics* 19, 346-355.
- Boomer, D. S. & Laver, J. D. M. (1968). Slips of the Tongue. *British Journal of Disorders of Communication* 3, 2-12.
- Bower, G. H. (Hrsg.) (1974). *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press.
- Bower, G. H. (Hrsg.) (1982). *The psychology of learning and motivation* (Vol. 16.). San Diego, CA: Academic Press.
- Broekhuis, H., Corver, N., Huybregts, R., Kleinherz, U. & Koster, J. (Hrsg.) (2005). *The organization of grammar*. Berlin: Mouton de Gruyter.
- Bromberger, S. & Halle, M. (1986). On the relationship of phonology and phonetics. In J. S. Perkell & D. H. Klatt (Hrsg.), *Invariance and variability in speech processes*. Hillsdale, N.J.: Erlbaum Associates, 510-519.
- Bromberger, S. & Halle, M. (1989). Why phonology is different? *Linguistic inquiry* 20(1), 51-70.
- Bromberger, S. & Halle, M. (1997). The contents of phonological signs: a comparison between their use in derivational theories and in optimality theories. In I. Roca (Hrsg.), *Derivations and Constraints in Phonology*. Oxford: OUP, 93-124.
- Bromberger, S. & Halle, M. (2000). The ontology of phonology (revised). In N. Burton-Roberts, P. Carr, & G. Docherty (Hrsg.), *Phonological knowledge: conceptual & empirical issues*. Oxford: OUP, 19-37.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Brugos, A., Micciulla, L. & Smith, C. E. (Hrsg.) (2004). *Proceedings of the Annual Boston University Conference on Language Development* 28.
- Bryan, K. (1989). Language prosody and the right hemisphere. *Aphasiology* 3, 285-299.
- Buckingham, H.W. (1990). Abstruse neologisms, retrieval deficits and the random generator. *Journal of Neurolinguistics* 5, 215-235.
- Bullowa, M. (Hrsg.) (1979). *Before speech: The beginning of human communication*. London: Cambridge University Press.
- Burton-Roberts, N., Carr, P. & Docherty, G. (Hrsg.) (2000). *Phonological knowledge: conceptual & empirical issues*. Oxford: OUP.
- Butterworth, B. (1979). Hesitation and the production of verbal paraphasias and neologisms in jargon aphasia. *Brain and Language* 18, 133-161.
- Butterworth, B. (Hrsg.) (1980). *Language Production. Volume I. Speech and Talk*. New York: Academic Press.
- Butterworth, B. (Hrsg.) (1983). *Language production. Volume II*. London: Academic Press.
- Bühler, K. (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer.
- Calabrese, A. (2012). Acoustic inputs and phonetic illusions: A linguist's perspective on the neuropsychological bases of speech perception. *Journal of Neurolinguistics* 25, 355-381.
- Cairns, C. (2002). Foot and syllable in Southern Paiute. Ms., City University of New York. <http://www.cunyphonologyforum.net/papers/20Sopai.pdf> (18 July 2012).
- Cairns, C. (2009). Phonological Representations and the Vaux-Wolfe Proposal. In E. Raimy & C. Cairns (Hrsg.) *Contemporary Views on Architecture and Representation in Phonology*, Cambridge, Massachusetts: MIT Press, 145-164.

- Caramazza, A. (1984). The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language* 21, 9-20.
- Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition* 5, 41–66.
- Caramazza, A. (1994). Parallels and divergences in the acquisition and dissolution of language. *Philosophical Transactions of the Royal Society of London, Series B*, Vol. 346, No. 1315, 121-127.
- Caramazza, A. & Zurif, E. (Hrsg.) (1978). *Language acquisition and language breakdown: Parallels and divergencies*. Baltimore: Johns Hopkins University Press.
- Caramazza, A. & Berndt, R. S. (1983). The Selective Impairment of Phonological Processing: A Case Study. *Brain and Language* 18, 128-174.
- Caramazza, A. & Badecker, W. (1989). Patient classification in neuropsychological research. *Brain and Cognition* 16, 256-295.
- Caramazza, A. & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate–inanimate distinction. *Journal of Cognitive Neuroscience* 10, 1-34.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.
- Carlson, L.A. & van der Zee, E. (Hrsg.) (2005). *Functional features in language and space: Insights from perception, categorization, and development*. New York: Oxford University Press.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competencies from 9 to 15 months of age. *Monographs of the Society of Research in Child Development* 63(4), 1-143.
- Carruthers, P. (2006). *The architecture of the mind: Massive modularity and the flexibility of thought*. New York: Clarendon Press/Oxford University Press.
- Carruthers, P., Laurence, S. & Stich, S. (Hrsg.) (2005). *The innate mind: Structure and content*. New York: Oxford University Press.
- Casasola, M. (2008). The development of infants' spatial categories. *Current Directions in Psychological Science*, 17 (1), 21-25.
- Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science* 26, 609-651.
- Chase, W G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Hrsg.), *The psychology of learning and motivation* (Vol. 16,). San Diego, CA: Academic Press, 1-58.
- Chomsky, N. (1975a): *The logical structure of linguistic theory*. New York: Plenum Press (Original 1955).
- Chomsky, N. (1975b). *Reflections on language*. New York: Pantheon Books.
- Chomsky, N. (1981a). *Regeln und Repräsentationen*. Frankfurt am Main: Suhrkamp.
- Chomsky, N. (1981b). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, N. (1987). *The Chomsky Reader*. New York: Pantheon Books.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1996). *Powers and Prospects: Reflections on Human Nature and the Social Order*. London: Pluto Press.
- Chomsky, N. (2002). An interview on minimalism. In A. Belletti & L. Rizzi (Hrsg.), *Noam Chomsky: On nature and language*. Cambridge: Cambridge University Press, 92-161.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
- Christman, S. S. (1992). Uncovering phonological regularity in neologisms: Contributions of sonority theory. *Clinical Linguistics and Phonetics* 6, 219–247.

- Clahsen, H. (Hrsg.) (1996). *Generative Perspectives on Language Acquisition*. Amsterdam: John Benjamins.
- Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Hrsg.), *Cognitive development and the acquisition of language*. New York: Academic Press, 65-110.
- Clements, G. N. & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the syllable*. Cambridge, Mass: MIT Press.
- Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2, 225-252.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckmann (Hrsg.), *Papers in laboratory phonology* 1. Cambridge: Cambridge Univ. Press, 283-333.
- Clements, G. N. & Hume, E. V. (1995). The internal organization of speech sounds. In John A. Goldsmith (Hrsg.), *The Handbook of phonological theory*. Cambridge, MA: Blackwell, 245-306.
- Clements, G. N. & Ridouane, R. (Hrsg.) (2011), *Where Do Phonological Features Come From? Cognitive, physical and developmental bases of distinctive speech categories*. Amsterdam : John Benjamins.
- Code, C. (1982). Neurolinguistic analysis of recurrent utterance in aphasia. *Cortex* 18, 141-152.
- Code, C. & Ball, M. J. (1994). Syllabification in aphasic recurring utterances: Contributions of sonority theory. *Journal of Neurolinguistics* 8, 257-265.
- Cohen, A. (1966). Errors of Speech and their Implication for Understanding the Strategy of Language Users. *Zeitschrift für Phonetik* 21, 177-181.
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Science* 3, 115-119.
- Coltheart, M., Sartori, G. & Job, R. (Hrsg.) (1987). *The cognitive neuropsychology of language*. Hillsdale, NJ: Erlbaum.
- Cooper, W. E. & Walker, E. C. T. (Hrsg.) (1979). *Sentence processing*. Hillsdale, N.J.: Erlbaum.
- Corina, D. & Sandler, W. (1993). On the nature of phonological structure in sign language. *Phonology* 10, 165-207.
- Coulter, G. (Hrsg.) (1993). *Current issues in ASL phonology*. New York: Academic Press.
- Csirmaz, A., Gualmini, A. & Nevins, A. (Hrsg.) (2004). *Plato's Problems: Papers on Language Acquisition*. Cambridge, MA: MITWPL.
- Cutler, A. (Hrsg.) (2005). *Twenty-first century psycholinguistics: four cornerstones*. Mahwah (NJ): Lawrence Erlbaum Associates.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.
- De Lacy, P. (Hrsg.) (2007). *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
- Dell, G. S. (1984). Representation of Serial Order in Speech: Evidence from the Repeated Phoneme Effect in Speech Errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10.2, 222-233.
- Dell, G. S. (1986). A Spreading Activation Theory of Retrieval in Sentence Production. *Psychological Review* 93, 283-321.
- Dell, G. S. (1988). The Retrieval of Phonological Forms in Production: Tests of Predictions from a Connectionist Model. *Journal of Memory and Language* 27, 124-142.
- Dell, G. S. & Reich, P. A. (1980). Toward a Unified Model of Slips of the Tongue. In V. Fromkin (Hrsg.), *Errors in Linguistic Performance*. New York: Academic Press, 273-286.

- Dell, G. S., & Sullivan, J. M. (2004). Speech errors and language production: Neuropsychological and connectionist perspectives. In B. H. Ross (Hrsg.), *The psychology of learning and motivation*. San Diego, CA: Elsevier, 63-108.
- Demuth, K. (1995). Markedness and the development of prosodic structure. *NELS* 25, 13-25.
- Demuth, K., & Fee, E. J. (1995). *Minimal prosodic words in early phonological development*. Ms.: Brown University and Dalhousie University.
- Diamond, A. (1985). The development of the ability to use recall to guide action, as indicated by infants performance on AB. *Child Development* 56, 868– 883.
- Dinnsen, D. (1997). Nonsegmental phonologies. In M. Ball & R. Kent (Hrsg.), *The New Phonologies: Developments in Clinical Linguistics*. San Diego, CA: Singular Publishing Group, 77- 126.
- Di Sciullo, A. M. (Hrsg.) (2003). *Asymmetry in Grammar, Volume 2*. Amsterdam: John Benjamins.
- Dittmann, J. (2002). *Der Spracherwerb des Kindes. Verlauf und Störungen*. München: Beck.
- Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics* 17, 617–643.
- Dore, J., Franklin, M. B., Miller, R. T. & Ramer, A. L. H. (1976). Transitional phenomena in early language acquisition. *Journal of Child Language* 3, 13-28.
- Dresher, B. E. & Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition* 34, 137-195.
- Dresher, B. E. (2003). Contrast and asymmetry in inventories. In A. M. Di Sciullo (Hrsg.), *Asymmetry in Grammar, Volume 2*. Amsterdam: John Benjamins, 239-257.
- Dressler, U. W. (Hrsg.) (1987). *Studies in Pre- & Protomorphology*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Dressler, W., Dziubalska-Kolaczyk, K., Gagarina, N. & Kilani-Schoch, M. (2005). Reduplication in child language. In B. Hurch (Hrsg.), *Studies on Reduplication*. Berlin: Mouton de Gruyter, 455-474.
- Dromi, E. (Hrsg.) (1993). *Language and cognition: A developmental perspective*. Norwood, NJ: Ablex.
- Dunlap, E. & Padgett, J. (Hrsg.) (1990). *University of Massachusetts Occasional Papers in Linguistics 14: Papers in Phonology*. Amherst, MA: GLSA
- Durand, J. (1990). *Generative and non-linear phonology*. London: Longman.
- Dümig, S. (2007). *Die Phonologie-Morphologie-Schnittstelle im L1-Erwerb: Eine klinische Einzelfallstudie*. Magisterarbeit, Universität Frankfurt. Unveröffentlicht.
- Dümig, S. (2009). Die Silbe und der Erwerb der Morphosyntax. *Sprache-Stimme-Gehör* 33(2), 100-101.
- Dümig, S. & Frank, A. (2008). The syllable and schwa in first language acquisition: Normal and impaired development. *Frankfurter Linguistische Forschung Sondernummer* 11, 65-90.
- Dümig, S. & Leuninger, H. (2013). *Phonologie der Laut- und Gebärdensprache: Linguistische Grundlagen, Erwerb, sprachtherapeutische Perspektiven*. Idstein: Schulz-Kirchner-Verlag.
- Dümig, S., Konradi, J. & Leuninger, H. (2016). The dispersion principle in aphasic language production. Effects of the sonority dispersion principle on error patterns in demisyllables of different phonological complexity. Mainz: Univ. https://publications.ub.uni-mainz.de/opus/frontdoor.php?source_opus=53720
- Echols, C.H. & Newport, E. L. (1992). The role of stress and position in determining first words. *Language Acquisition* 2, 189-220.

- Eckman, F. R., Iverson, G. K. (1994). Pronunciation difficulties in ESL: coda consonants in English interlanguage. In M. Yavaş (Hrsg.), *First and Second Language Phonology*. San Diego, CA: Singular Publishing Group, 251-266.
- Edelman, G. M. (1992). *Bright Air, Brilliant Fire - On the Matter of the Mind*. New York: Basic Books.
- Einstein, A. (1969). *Grundzüge der Relativitätstheorie*. Berlin: Akademie-Verlag.
- Eisenberg, P., Ramers, K. & Vater, H. (Hrsg.) (1992). *Silbenphonologie des Deutschen*. Tübingen: Gunter Narr Verlag.
- Eisenberg, P. (2000): *Grundriss der deutschen Grammatik*. Stuttgart, Weimar: Metzler.
- Ellis, A. W. (Hrsg.) (1985), *Progress in the Psychology of Language*. London: Lawrence Erlbaum.
- Ellis, A. W. (1987). Intimations of modularity, or, the modelarity of mind: doing cognitive neuropsychology without syndromes. In M. Coltheart, G. Sartori, & R. Job (Hrsg.), *The cognitive neuropsychology of language*. Hillsdale, NJ: Erlbaum, 397-408.
- Ellis, A. W. & Young, A. W. (1991). *Einführung in die kognitive Neuropsychologie*. Bern: Hans Huber.
- Elsen, H. (1991). *Erstspracherwerb: der Erwerb des deutschen Lautsystems*. Wiesbaden: Deutscher Universitäts-Verlag.
- Elsen, H. (1999). *Ansätze zu einer funktionalistisch-kognitiven Grammatik. Konsequenzen aus Regularitäten des Erstspracherwerbs*. Tübingen: Niemeyer.
- Engfer, H.-J. (1996). *Empirismus versus Rationalismus? Kritik eines philosophiegeschichtlichen Schemas*. Paderborn: Schöningh.
- Fee, E. J. (1996). Syllable structure and minimal words. In B. Bernhardt, J. Gilbert & D. Ingram, (Hrsg.), *Proceedings of the UBC international conference on phonological acquisition*. Somerville, MA: Cascadilla Press, 85-98.
- Fee, J. & Ingram, D. (1982). Reduplication as a strategy of phonological development. *Journal of Child Language* 9, 41-54.
- Ferguson, C. A., Peizer, D. B. & Weeks, T. A. (1973). Model-and-replica phonological grammar of a child's first words. *Lingua* 31(1), 35-65.
- Ferguson, C. A. (1978). Learning to pronounce: The earliest stages of phonological development in the child. In F. D. Minifie & L. L. Lloyd (Hrsg.), *Communicative and Cognitive Abilities – Early Behavioral Assessment*. Baltimore, MD: University Park Press, 273-297.
- Ferguson, C. A. (1983). Reduplication in child phonology. *Journal of Child Language* 10, 239-243.
- Ferguson, C. A., Menn, L. & Stoel-Gammon, C. (Hrsg.) (1992). *Phonological Development: Models, Research, Implications*. Timonium, Maryland: York Press.
- Féry, C. (1986). Metrische Phonologie und Wortakzent im Deutschen. *Studium Linguistik* 20, 16-43.
- Fikkert, P. (1994). *On the Acquisition of Prosodic Structure*. Ph.D. Dissertation, HIL dissertations 6, Leiden University. The Hague: Holland Academic Graphics.
- Fikkert, P. (2007). Acquiring phonology. In P. de Lacy (Hrsg.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 537-554.
- Fikkert, P., Penner, Z. & Wymann, K. (1998). Das Comeback der Prosodie. Neue Wege in der Diagnose und Therapie von phonologischen Störungen. *Logos Interdisziplinär* 6, 84-97.
- Fletcher, P. & Garman, M. (Hrsg.) (1979). *Language Acquisition: Studies in first language development*. Cambridge: Cambridge University Press.

- Fletcher, P. & MacWhinney, B. (Hrsg.) (1995). *The handbook of child language*. Oxford: Blackwell.
- Fodor, J. (1968). *Psychological Explanation: An Introduction to the Philosophy of Psychology*. New York: Random House.
- Fodor, J. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, Jerry, 1987. *Psychosemantics*. Cambridge, MA: MIT Press.
- Fodor, J. (1998). *In Critical Condition: Polemic Essays on Cognitive Science and the Philosophy of Mind*. Cambridge, MA: MIT Press.
- Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.
- Fodor, J., Garrett, M. F., Walker, E. C. T. & Parkes, C. H. (1980). Against definitions. *Cognition* 8, 263-367.
- Fodor, J. & Pylyshyn, Z. (1981) How direct is visual perception? Some reflections on Gibson's "Ecological Approach." *Cognition* 9, 139-196.
- Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. In S. Pinker & J. Mehler (Hrsg.), *Connections and symbols*. Cambridge, MA: MIT Press, 3-71.
- Fodor, J. D., Fodor, J., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry* 4, 515-531.
- Foster, S. H. (1990). *The Communicative Competence of Young Children: A Modular Approach*. London & New York: Longman.
- Fox, A. (2005): *Kindliche Aussprachestörungen. Phonologischer Erwerb – Differenzialdiagnostik – Therapie*. Idstein: Schulz-Kirchner Verlag.
- Fox, A. (2006). *TROG-D. Test zur Überprüfung des Grammatikverständnisses*. Idstein: Schulz-Kirchner Verlag.
- Frampton, John (2009). *Distributed Reduplication*. Linguistic Inquiry Monographs 52. Cambridge, MA: MIT Press.
- Frege, G. (1976) (Original: 1923). Gedankengefüge. In G. Patzig (Hrsg.), *Frege, Logische Untersuchungen*. Göttingen: Vandenhoeck & Rupprecht, 72-91.
- Freud S. (1897) Die infantile Cerebrallähmung. In H. Nothnagel (Hrsg.), *Spezielle Pathologie und Therapie*, Bd IX, Teil 111. Wien: Holder, 1-327.
- Friederici, A. & Menzel, R. (Hrsg.) (1999). *Learning: Rule abstraction and representation*. Berlin: de Gruyter.
- Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: From syllables to sentences. *Trends in Cognitive Science* 9, 481-488.
- Friederici, A. D. (2011). The Brain Basis of Language Processing: From Structure to Function. *Physiological Review* 91, 1357-1392.
- Fromkin, V. (1971). The Non-anomalous Nature of Anomalous Utterances. *Language* 47, 27-52.
- Fromkin, V. (Hrsg.) (1980). *Errors in Linguistic Performance*. New York: Academic Press.
- Fujimura, O. & Lovins, J. B. (1978). Syllables as concatenative phonetic units. In A. Bell & J. B. Hooper (Hrsg.), *Syllables and segments*. New York: Elsevier, 107-120.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience* 12, 207-222.
- Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Sathamnuwong, N., Tong, Y. & Xiaojian Li (2002). A crosslinguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience* 14(7), 1076-1087.

- García-Albea, J. E., Del Viso, S. & Igoa, J. M. (1989). Movement Errors and Levels of Processing in Sentence Production. *Journal of Psycholinguistic Research* 18.1, 145-161.
- Garrett, M. F. (1975). The analysis of sentence production. In B. Gorden (Hrsg.), *The psychology of learning and motivation. Advances in Research and Theory*. New York: Academic Press, 133-177.
- Garrett, M.F. (1980). Levels of Processing in Sentence Production. In B. Butterworth (Hrsg.), *Language Production. Volume I. Speech and Talk*. New York: Academic Press, 177-220.
- Gaskell, M. G. (Hrsg.) (2007). *The Oxford Handbook of Psycholinguistics*. New York: Oxford University Press.
- Gazzaniga, M. (Hrsg.) (2000). *The New Cognitive Neuroscience*, Second edition, Cambridge, MA: MIT Press.
- Gee, J. & Grosjean, F. (1983). Performance structures: a psycholinguistics and linguistic appraisal. *Cognitive Psychology* 15, 411-458.
- Gierut, J. A. & O'Connor, K. M. (2002). Precursors to onset clusters in acquisition. *Journal of Child Language* 29, 495-517.
- Goethe, Johann Wolfgang (1948). Gedenkausgabe der Werke, Briefe und Gespräche. Band 1–24 und Erg.-Bände 1–3, Band 16, Zürich, 418-419.
- Goldfield, B. A. & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17(1), 171-184.
- Goldsmith, J. (1976). *Autosegmental Phonology*. Doctoral dissertation, MIT.
- Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.
- Goldsmith, J. A. (Hrsg.) (1995). *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell Publishers.
- Gorden, B. (Hrsg.) (1975). *The psychology of learning and motivation. Advances in Research and Theory*. New York: Academic Press.
- Greenhill, A., Hughes, M., Littlefield, H. & Walsh, H. (Hrsg.) (1998). *Proceedings of the 22 Annual Boston Conference on Language Development*, Vol. 1. Somerville, Mass.: Cascadilla Press.
- Grégoire, A. (1947). *L'apprentissage du langage. La troisième année et les années suivantes*. Liège: Bibliothèque de la Faculté des Lettres.
- Grijzenhout, J. & S. Joppen (2002). The lack of onsets in German child phonology. In I. Lasser (Hrsg.), *The process of language acquisition*. Frankfurt/Berlin: Peter Lang Verlag.
- Grimm, H. (Hrsg.) (2000). *Sprachentwicklung. Enzyklopädie der Psychologie*, Serie III, Band 3. Göttingen: Hogrefe.
- Grimm, H. (2000). Entwicklungsdysphasie: Kinder mit spezifischer Sprachstörung. In H. Grimm (Hrsg.), *Sprachentwicklung. Enzyklopädie der Psychologie*, Serie III, Band 3. Göttingen: Hogrefe, 603-640.
- Grodzinsky, Y., Shapiro, L. & Swinney, D. (Hrsg.) (2000). *Language and the Brain: Representation and Processing*. San Diego: Academic Press.
- Grunwell, P. (1987). *Clinical Phonology*. London: Croom Helm.
- Guerssel, M. (1986). Glides in Berber and Syllabicity. *Linguistic Inquiry* 17, 1-12.
- Hagoort, P. (2005). On Broca, brain and binding: a new framework. *Trends in Cognitive Sciences* 9 (9), 416-423.
- Hagoort, P. (2008). Should Psychology Ignore the Language of Brain? *Psychological Science* 2(17), 96-100.
- Hagoort, P. & Levelt, W. J. M. (2009). The Speaking Brain. *Science* 326, 372-373.
- Hale, M. (2007). *Historical Linguistics: Theory and Method*. Oxford: Blackwell.

- Hale, M. & Reiss, C. (2003). The Subset Principle in phonology: why the tabula can't be rasa. *Journal of Linguistics* 39, 219–244.
- Hale, M. & Reiss, C. (2007). The Subset Principle in phonology: why the tabula can't be rasa. <http://linguistics.concordia.ca/reiss/subby.pdf>, 1-19.
- Hale, M. & C. Reiss. 2008. *The phonological enterprise*. Oxford: Blackwell.
- Hall, T. (2000). *Phonologie*. Berlin: de Gruyter.
- Halle, M. (1990). Respecting metrical structure *Natural Language and Linguistic Theory*, 149-176.
- Halle, M. & Vergnaud, J.-R. (1980). Three Dimensional Phonology, *Journal of Linguistic Research* 1, 83-105.
- Halle, M. & Idsardi, W. (1995). General Properties of Stress and Metrical Structure." In J. A. Goldsmith (Hrsg.), *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell Publishers, 403-443.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold.
- Halliday, M. A. K. (1979). One child's protolanguage. In M. Bullowa (Hrsg.), *Before speech: The beginning of human communication*. London: Cambridge University Press, 171-190.
- Halpern, A. L. (Hrsg.) (1990). *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*. Stanford: Stanford Linguistics Association/CSLI.
- Hannahs, S. J. & Young-Scholten, M. (Hrsg.) (1997). *Focus on phonological acquisition*. Amsterdam: Benjamins.
- Happ, D. & Vorköper, M.-O. (2006). *Deutsche Gebärdensprache: Ein Lehr- und Arbeitsbuch*. Frankfurt am Main: Fachhochschulverlag.
- Harris, J., Watson, J. & Bates, S. (1999). Prosody and melody in vowel disorder. *Journal of Linguistics* 35, 489-525.
- Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Hauser, M. D. & Fowler, C. A. (1992). Fundamental frequency declination is not unique to human speech: Evidence from nonhuman primates. *Journal of the Acoustical Society of America* 91, 363-369.
- Hauser, M. D., Chomsky, N. & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve?. *Science* 298, 1569-1579.
- Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.
- Hayes, B. (1990). Diphthongisation and Coindexing. *Phonology* 7, 31-72.
- Hayes, B., Kirchner, R. & Steriade, D. (Hrsg.) (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.
- Hayes, B. & Steriade, D. (2004). Introduction: the phonetic bases of phonological Markedness. In B. Hayes, R. Kirchner & D. Steriade (Hrsg.), *Phonetically based phonology*. Cambridge: Cambridge University Press, 1–33.
- Heidolph, K. E., Flämig, W. & Motsch, W. (Hrsg.) (1984). *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag.
- Henning, A., Daum, M. & Aschersleben, G. (2009). Frühkindliche Handlungswahrnehmung und Theory of Mind: Vom Verständnis zielgerichteter Handlungen zum

- Verständnis mentalistisch gesteuerter Handlungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* Jg. 41, 233-242.
- Hennon, E., Hirsh-Pasek, K. & Golinkoff, R. M. (2000). Die besondere Reise vom Fötus zum spracherwerbenden Kind. In H. Grimm (Hrsg.), *Sprachentwicklung. Enzyklopädie der Psychologie*, Serie III, Band 3. Göttingen: Hogrefe, 41-103.
- Hintikka, J., Suppes, P. & Moravcsik, J. M. E. (Hrsg.) (1973). *Approaches to natural language*. Dordrecht: Reidel.
- Hirst, D. (1993). Detaching intonational phrases from syntactic structure. *Linguistic Inquiry* 24, 781-788.
- Hohenberger, A. (2002). *Functional Categories in Language Acquisition*. Tübingen: Niemeyer.
- Höhle, B. (2002). *Der Einstieg in die Grammatik: Die Rolle der Phonologie-Syntax-Schnittstelle für Sprachverarbeitung und Spracherwerb*. Habilitationsschrift. Freie Universität Berlin.
- Höhle, B. & Weissenborn, J. (1998). Sensitivity to closed-class-elements in preverbal children. In A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (Hrsg.), *Proceedings of the 22 Annual Boston Conference on Language Development*, Vol. 1. Somerville, Mass.: Cascadilla Press, 348-359.
- Höhle, B. & Weissenborn, J. (1999). Discovering grammar. Prosodic and morphosyntactic aspects of rule formation in first language acquisition. In A. Friederici & R. Menzel (Hrsg.), *Learning: Rule abstraction and representation*. Berlin: de Gruyter, 37-69.
- Höhle, B., Bijeljic-Babic, R. & Herold, B., Weissenborn, J. & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Cognitive Psychology* 32, 262-274.
- Hook, S. (Hrsg.) (1960). *Dimensions of Mind: A Symposium*. New York: New York University Press.
- Huber, W. & Schlenck, K.-J. (1988). Satzverschränkungen bei Wernicke-Aphasie. In: G. Blanken, J. Dittmann, C.-W. Wallesch (Hrsg.) (1988): *Sprachproduktionsmodelle. Neuro- und psycholinguistische Theorien der menschlichen Spracherzeugung*. Freiburg: Hochschulverlag.
- Huck, G. & Ojeda, A. (Hrsg.) (1987). *Syntax and Semantics, Vol. 20, Discontinuous Geometry*. Orlando, FL: Academic Press.
- Hulst, Harry van der. (2000). Modularity and Modality in Phonology. In N. Burton-Roberts, P. Carr & G. Docherty (Hrsg.), *Phonological Knowledge. Conceptual and Empirical Issues*. Oxford: OUP, 207-243.
- Hulst, H. van der. (2003). Cognitive Phonology. *Germania at alia. A linguistic webschrift for Hans den Besten on the occasion of his 55th birthday*, 1-24.
- Hulst, H. van der. (2005). Why phonology is the same. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinherz, & J. Koster (Hrsg.), *The organization of grammar*. Berlin: Mouton de Gruyter, 252-261.
- Humboldt, Wilhelm von (1998). *Über die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechts*. Paderborn: Schöningh (Original 1836).
- Hurch, B. (Hrsg.) (2005). *Studies on Reduplication*. Berlin: Mouton de Gruyter.
- Hussy, W. (1993). *Denken und Problemlösen*. Stuttgart: Kohlhammer.
- Huttenlocher, P. R. (1994). Synaptogenesis, synapse elimination, and neural plasticity in human cerebral cortex. In C. Nelson (Hrsg.), *Threats to Optimal Development: Integrating Biological, Psychological, and Social Risk Factors*, Vol. 27. Hillsdale, NJ: Lawrence Erlbaum Associates, 35-54.

- Huttenlocher, P. R. & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *J. Comp. Neurol.* 387, 167-178.
- Hyams, N. & Sigurjónsdóttir, S. (1990). The development of 'long distance anaphora,': a cross-linguistic study with special reference to Icelandic. *Language Acquisition* 1, 57-93.
- Idsardi, W. (1992). *The Computation of Prosody*, Doctoral dissertation: MIT.
- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101-144.
- Ingram, D. (1974). Phonological rules in young children. *Journal of Child Language* 1, 49-64.
- Ingram, David. (1978). The role of the syllable in phonological development. In A. Bell & J. B. Hooper (Hrsg.), *Syllables and Segments*. New York: Elsevier, 143-156.
- Itô, J. (1986). *Syllable Theory in Prosodic Phonology*. Ph.D. Dissertation. Amherst: University of Massachusetts.
- Itô, J. & Mester, A. (1995). The core-periphery structure of the lexicon and constraints on reranking. In J. Beckman, L. Walsh-Dickey & S. Urbanczyk (Hrsg.), *University of Massachusetts Occasional Papers in Linguistics: Papers in Optimality Theory*. Amherst: GLSA, 181-209.
- Itô, J. & Mester, A. (2009). The onset of the prosodic word. In S. Parker (Hrsg.), *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox, 227-260.
- Ivanova, G. (2006). Sound symbolic approach to Japanese mimetic words. *Toronto Working Papers in Linguistics* 26, 103-114.
- Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge, Mass.: MIT Press.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, Mass.: MIT Press.
- Jackendoff, R. (2000). Fodorian Modularity and Representational Modularity. In Y. Grodzinsky, L. Shapiro & D. Swinney (Hrsg.), *Language and the Brain: Representation and Processing*. San Diego: Academic Press, 4-30.
- Jackendoff, R. (2002). *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, R. (2003). Précis of Foundations of Language: Brain, Meaning, Grammar, Evolution. *Behavioral and Brain Sciences* 26, 651-707.
- Jackendoff, R. (2007a). A Parallel Architecture perspective on language processing. *Brain Research* 1146, 2-22.
- Jackendoff, R. (2007b). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.
- Jackendoff, R. (2014). *Genesis of a theory of language: From thematic roles (source) to the Parallel Architecture (goal) (Sort of an intellectual memoir)*. <http://ase.tufts.edu/cogstud/jackendoff/papers/GenesisofPA.pdf>
- Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.
- James, A. & Leather, J. (Hrsg.) (1987). *Sound patterns in second language acquisition*. Providence, RI: Foris Publications.
- Jespersen, O. (1913). *Elementarbuch der Phoentik*. Leipzig/Berlin: Teubner.
- Jones, G. (2012). Why chunking should be considered as an explanation for developmental change before short-term memory capacity and processing speed. *Frontiers in Psychology* 3, 167.
- Jovanovic, B., Király, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2007). The role of effect for infants' perception of action goals. *Psychologia* 50, 273-290.
- Julesz, B. (1975). Experiments in the Visual Perception of Texture. *Scientific American* 232, 34-43.

- Jusczyk, P. W., Luce, P. A. & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language* 33, 630-645.
- Jusczyk, P. W., Houston, D. M. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39, 159-207.
- Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99, 122-149.
- Kager, R., Pater, J. & Zonneveld, W. (Hrsg.) (2004). *Constraints in phonological acquisition*. Cambridge, England: Cambridge University Press.
- Karmiloff-Smith, A. (1992). *Beyond modularity. A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Kauschke, C. (2000). *Der Erwerb des frühkindlichen Lexikons. Eine empirische Studie zur Entwicklung des Wortschatzes im Deutschen*. Tübingen: Narr.
- Katz, J. (1964). Mentalism in Linguistics. *Language* 40, 124-137.
- Katz, J. (1996). The unfinished Chomskyan revolution. *Mind and Language* 11, 270-294.
- Katz, L., Shatz, C., Weliky, M. & Crowley, J. (2000). Activity and the development of the visual cortex: new perspectives. In M. Gazzaniga (Hrsg.), *The New Cognitive Neuroscience*, Second edition, Cambridge, MA: MIT Press.
- Kehoe, M. & Lleó, C. (2003). The acquisition of nuclei: A longitudinal analysis of phonological vowel length in three German-speaking children. *Journal of Child Language* 30, 527-556.
- Kehoe, M. & Stoel-Gammon, C. (2001). Development of syllable structure in English-speaking children with particular reference to rhymes. *Journal of Child Language* 28, 393-432.
- Keller, E. & Gopnik, M. (Hrsg.) (1987). *Motor and sensory processes of language*. Hillsdale N.J.: Erlbaum.
- Kessel, F. (Hrsg.) (1988). *The development of language and language researchers*. Hillsdale, NJ: Erlbaum.
- Kiese-Himmel, C. (2005). *Aktiver Wortschatztest für 3- bis 5-jährige Kinder (AWST-R)*. Göttingen: Göttingen Verlag.
- Kingston, J. & Beckmann, M. (Hrsg.) (1990). *Papers in laboratory phonology 1*. Cambridge: Cambridge Univ. Press.
- Kiparsky, P. (1973). "Elsewhere" in phonology. In S. Anderson & P. Kiparsky (Hrsg.), *A festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston, 93-106.
- Király, I., Jovanovic, B., Prinz, W., Aschersleben, G. & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition* 12, 752-769.
- Kitagawa, Y. (1987). Redoing reduplication: a preliminary sketch. In G. Huck and A. Ojeda (Hrsg.), *Syntax and Semantics, Vol. 20, Discontinuous Geometry*. Orlando, FL: Academic Press, 71-106.
- Knobloch, C. (2001). Psychologische Ansätze bei der Erforschung des frühkindlichen Spracherwerbs. In S. Auroux, E. F. K. Koerner, H.-J. Niederehe & K. Versteegh (Hrsg.), *Geschichte der Sprachwissenschaften. Ein internationales Handbuch zur Entwicklung der Sprachforschung von den Anfängen bis zur Gegenwart*, Teilband 2. Berlin: de Gruyter, 1705-1718.
- Kuczaj, S. A. & Barrett, M. D. (Hrsg.) (1986). *The development of word meaning*. Berlin: Springer.
- Lasser, I. (Hrsg.) (2002). *The process of language acquisition*. Frankfurt/Berlin: Peter Lang Verlag.
- Laubstein, A. S. (1987). Syllable Structure: The Speech Error Evidence. *Canadian Journal of Linguistics* 32.4, 339-363.

- Leather, J. & James, A. (Hrsg.) (1997). *New sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech*. Klagenfurt: Universität Klagenfurt.
- Leopold. W. (1947). *Speech development of a bilingual child: a linguist's report: Vol. 2. Sound-learning in the first two years*. Evanston IL: Northwestern University Press.
- Leopold. W. (1949). *Speech development of a bilingual child: a linguist's report: Vol. 3. Grammar and general problems in the first two years*. Evanston IL: Northwestern University Press.
- Leuninger, H. (1993). *Reden ist Schweigen, Silber ist Gold. Gesammelte Versprecher* (3. ed.). Zürich: Ammann.
- Leuninger, H. (2005). Erfolgreiche und weniger erfolgreiche Wege vom Gedanken zur Äußerung: Das wirft einige Fragezeichen auf. http://www1.uni-frankfurt.de/fb/fb10/KogLi/Lehrstuehle/ehem__Lehrstuhl_Leuninger/Download/Gedanken zur.ppt
- Leuninger, H., Dümig, S., Kops, F., Koch, C., Grimm, T., Becker, J., Paul, M. & Schwarze, R. (2013). Phonologische Komplexität: Demisilben in Versprechern, Paraphrasen und Neologismen. *Zeitschrift für angewandte Linguistik* 59 (1), 1-19.
- Levelt, C. C. (1994). *The acquisition of place*. Holland Institute of Generative Linguistics Publications.
- Levelt, C., Schiller, N. & Levelt, W. J. M. (2000). The acquisition of syllable types. *Language Acquisition* 8, 237-264.
- Levelt, C. & van de Vijver, R. (2004). The acquisition of syllable types in cross-linguistic and developmental grammars. In R. Kager, J. Pater, & W. Zonneveld (Hrsg.), *Constraints in phonological acquisition*. Cambridge, England: Cambridge University Press, 204-218.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1991). Die konnektionistische Mode. *Sprache & Kognition* 10 (2), 61-72.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition* 42, 1-22.
- Levelt, W. J. M. (1998). The genetic perspective in psycholinguistics. Or where do spoken words come from? *Journal of Psycholinguistic Research* 27, 167-180.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1-75.
- Levin, J. (1985). *A Metrical Theory of Syllabicity*, Doctoral dissertation. Cambridge, MA; MIT Press.
- Levitte, A. G. & Healy, A. F. (1985). The Roles of Phoneme Frequency, Similarity, and Availability in the Experimental Elicitation of Speech Errors. *Journal of Memory and Language* 24, 717-733.
- Liberman, A., Cooper, F., Shankweiler, D. & Studdert-Kennedy, M. (1967). The Perception of The Speech Code. *Psychological Review* 74, 431-461.
- Liberman, M. (1975). The intonational system of English. PhD thesis, MIT. <http://dspace.mit.edu/handle/1721.1/27376#files-area>
- Liu, L., Peng, D., Ding, G., Jin, Z., Zhang, L., Li, K., et al. (2006). Dissociation in the neural basis underlying Chinese tone and vowel production. *Neuroimage* 29, 515-523.
- Lleó, C. (1990). Homonymy and reduplication: an extended availability of two strategies in phonological acquisition. *Journal of Child Language* 17, 267-278.
- Lleó, C. & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language* 23, 31-56.

- Lleó, C. & Prinz, M. (1997). Syllable structure parameters and the acquisition of affricates. In S. J. Hannahs & M. Young-Scholten (Hrsg.), *Focus on phonological acquisition*. Amsterdam: Benjamins, 143-164.
- Lohuis-Weber, H. & Zonneveld, W. (1996). Phonological acquisition and Dutch word prosody. *Language Acquisition* 5, 245-283.
- Löblich, B. (2000). *Semantikerwerb: ein Beitrag zu einer empiristisch-naturalistischen Bedeutungstheorie*. Tübingen: Niemeyer.
- Lyons, J. (Hrsg.) (1970). *New Horizons in Linguistics*. Harmondsworth: Penguin.
- MacKay, D. G. (1970). Spoonerisms: The Structure of Errors in the Serial Order of Speech. *Neuropsychologia* 8, 323-350.
- Macken, M. A., & Ferguson, C. A. (1983). Cognitive aspects of phonological development: Model, evidence and issues. In K. Nelson (Hrsg.), *Children's language*. Vol. 4. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Macken, M. A. (1992). Where's Phonology? In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Hrsg.), *Phonological Development: Models, Research, Implications*. Timonium, Maryland: York Press, 249-269.
- Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
- MacNeilage, P. F. (Hrsg.) (1983). *The production of speech*. New York: Springer.
- Markman, E. (1993). Ways in which children constrain word meanings. In E. Dromi (Hrsg.), *Language and cognition: A developmental perspective*. Norwood, NJ: Ablex, 61-87.
- Marr, D. (1982). *Vision*. Cambridge, MA: MIT Press.
- Marslen-Wilson, W. (1973). *Speech Shadowing and Speech Perception*. Ph.D.thesis, Massachusetts: M.I.T.
- McCarthy, J. (1988). Feature geometry and dependency: a review. *Phonetica* 45, 84-108.
- McCarthy, J. & Prince, A. (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8, 209-283.
- McCawley, J. D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.
- McMahon, A. (2007). Why phonology is plural. In M. C. Pennington (Hrsg.), *Phonology in context*. Basingstoke: Palgrave Macmillan, 159-185.
- McNally, L. (1990). Multiplanar reduplication: Evidence from Sesotho. In A. L. Halpern (Hrsg.), *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*. Stanford: Stanford Linguistics Association/CSLI, 331-346.
- Meibauer, J. & Rothweiler, M. (Hrsg.) (1999). *Das Lexikon im Spracherwerb*. Tübingen, Basel: Francke.
- Meisel, J. (1995). Parameters in acquisition. In P. Fletcher & B. MacWhinney (Hrsg.), *The handbook of child language*. Oxford: Blackwell, 10-35.
- Meinschaefer, J. (1998): *Silbe und Sonorität in Sprache und Gehirn*. Universität Bochum: <http://www-brs.ub.ruhr-uni-bochum.de/net/html/HSS/Diss/MeinschaeferJudith/diss.pdf>.
- Menn, L. (1976). *Pattern, Control and Contrast in Beginning Speech: A case study of word form and word function*. Unpublished Ph.D. Thesis, University of Illinois (Veröffentlicht (1978). Bloomington: Indiana University Linguistics Club.)
- Menn, L. (1983). Development of articulatory, phonetic and phonological capabilities. In B. Butterworth (Hrsg.), *Language production. Volume II*. London: Academic Press, 3-50.
- Menn, L. & M. M. Vihman (2011). Features in child phonology : Inherent, emergent, or artefacts of analysis?. In G. N. Clements & R. Ridouane (Hrsg.), *Where Do*

- Phonological Features Come From? Cognitive, physical and developmental bases of distinctive speech categories.* Amsterdam : John Benjamins, 259–302.
- Menn, L. (2013). Development of articulatory, phonetic, and phonological capabilities. In M. M. Vihman & T. Keren-Portnoy (Hrsg.), *The Emergence of Phonology. Whole-word Approaches and Cross-linguistic Evidence.* Cambridge: Cambridge University Press, 3-50.
- Menyuk, P. & Menn, L. (1979). Early strategies for the perception and production of words and sounds. In P. Fletcher & M. Garman (Hrsg.), *Language Acquisition: Studies in first language development.* Cambridge: Cambridge University Press, 49-70.
- Meyer, A.S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17, 1146-1160.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63, 81-97.
- Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Holt, Rinehart, & Winston.
- Minifie, F. D. & Lloyd, L. L. (Hrsg.) (1978). *Communicative and Cognitive Abilities – Early Behavioral Assessment.* Baltimore, MD: University Park Press.
- Moen, I. (2000). Foreign accent syndrome: A review of contemporary explanations. *Aphasiology* 14, 5-15.
- Mohammed, M.A. (2001). *Modern Swahili Grammar.* Nairobi: East African Education Press.
- Moore, T. E. (Hrsg.) (1973). *Cognitive development and the acquisition of language.* New York: Academic Press.
- Moskowitz, A. I. (1973). The acquisition of phonology and syntax: a preliminary study. In J. Hintikka, P. Suppes & J. M. E. Moravcsik (Hrsg.), *Approaches to natural language.* Dordrecht: Reidel, 48-84.
- Moss, H. E., Tyler, L. K., & Taylor, K. I. (2007). Conceptual structure. In M. G. Gaskell (Hrsg.), *The Oxford Handbook of Psycholinguistics.* New York: Oxford University Press, 217-234.
- Motsch, H.-J. (1999). *ESGRAF-Testmanual: Evozierte Sprachdiagnose grammatischer Fähigkeiten.* München: E. Reinhardt.
- Neef, M. (2004). Segments with inherently falling sonority. *Studia Linguistica* 58(3), 252-268.
- Nelson, C. (Hrsg.) (1994). *Threats to Optimal Development: Integrating Biological, Psychological, and Social Risk Factors,* Vol. 27. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* 38.
- Nelson, K. (Hrsg.) (1983). *Children's language,* Vol. 4. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Nespor, M. & Vogel, I. (1986). *Prosodic phonology.* Dordrecht: Foris.
- Neville, H. J. & Bavelier, D. (2000). Specificity and plasticity in neurocognitive development in humans. In M. Gazzaniga (Hrsg.), *The New Cognitive Neurosciences.* Cambridge, MA: MIT Press, 83-98.
- Nevins, A. (2004). What UG Can and Can't do to Help the Reduplication Learner. In A. Csirmaz, A. Gualmini & A. Nevins (Hrsg.), *Plato's Problems: Papers on Language Acquisition.* Cambridge, MA: MITWPL, 113-126.
- Noma, H. (1998). Languages richest in onomatopoeic words. *Language Monthly* 27(5), 30-34.

- Nooteboom, S. G. (1969). The Tongue Slips into Patterns. In A. Sciarone, A. van Essen & A. van Raad (Hrsg.), *Leiden Studies in Linguistics and Phonetics*. The Hague: Mouton, 144-156.
- Nothnagel, H. (Hrsg.) (1897). *Spezielle Pathologie und Therapie*, Bd IX, Teil 111. Wien: Holder.
- Obler, L. & Menn, L. (Hrsg.) (1982). *Exceptional Language and Linguistics*. New York: Academic Press.
- Ohala, D. K. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders* 32, 397-422.
- Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komishian, J. F. Kavanagh & C. A. Ferguson (Hrsg.), *Phonology, Vol.1: Production*. New York: Academic Press, 93-112.
- Onn, Farid. (1976). *Aspects of Malay Phonology and Morphology: A Generative Approach*. Doctoral dissertation. Urbana: University of Illinois.
- Ott, S., van de Vijver, R. & Höhle, B. (2006). The effect of phonotactic constraints in German-speaking children with delayed phonological acquisition: Evidence from Production of word-initial consonant clusters. *International Journal of Speech-Language Pathology* 4, 323-334.
- Parker, S. (Hrsg.) (2009). *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox.
- Patzig, G. (Hrsg.) (1976). *Frege, Logische Untersuchungen*. Göttingen: Vandenhoeck & Rupprecht.
- Pauen, S. & Vonderlin, E. (2007). Entwicklungsdiagnostik in den ersten drei Lebensjahren. *Research Notes* 21. Berlin: Deutsches Institut für Wirtschaftsforschung.
- Penner, Z. (2002). Plädoyer für eine präventive Frühintervention bei Kindern mit Spracherwerbstörungen. In W. Suchodoletz (Hrsg.), *Therapie von Sprachentwicklungsstörungen. Anspruch und Realität*. Stuttgart: Kohlhammer, 106-142.
- Penner, Z. & Weissenborn, J. (1996). Strong continuity, parameter setting and the trigger hierarchy. In H. Clahsen (Hrsg.), *Generative Perspectives on Language Acquisition*. Amsterdam: John Benjamins, 161-200.
- Penner, Z., Tracy, R. & Wymann, K. (1999). Die Rolle des Fokuspartikel AUCH im frühen kindlichen Lexikon. In J. Meibauer & M. Rothweiler (Hrsg.) (1999), *Das Lexikon im Spracherwerb*. Tübingen, Basel: Francke, 88-105.
- Penner, Z., Fischer, A. & Krügel, Ch. (2006). *Von der Silbe zum Wort*. Bildungsverlag EINS. Troisdorf.
- Pennington, M. C. (Hrsg.) (2007). *Phonology in context*. Basingstoke: Palgrave Macmillan.
- Perkell, J. S. & Klatt, D. H. (Hrsg.) (1986). *Invariance and variability in speech processes*. Hillsdale, N.J.: Erlbaum Associates.
- Petitto, L. (1988). "Language" in the prelinguistic child. In F. Kessel (Hrsg.), *The development of language and language researchers*. Hillsdale, NJ: Erlbaum, 187-222.
- Peuser, G. (2000). *Sprachstörungen. Einführung in die Patholinguistik*. München: Fink.
- Pinker, S. & Mehler, J. (Hrsg.) (1988). *Connections and symbols*. Cambridge, MA: MIT Press.
- Pinker, Steven. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Pinker, S. & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition* 95, 201-236.

- Piske, T. (2001). *Artikulatorische Muster im frühen Laut- und Lexikonerwerb*. Tübingen: Gunter Narr.
- Poeck, K., Kerschensteiner, M., Stachowiak, F.-J. & Huber, W. (1974). Die amnestische Aphasie. Klinisches Bild und Überlegungen zur neurolinguistischen Struktur. *Journal of Neurology* 207, 1-17.
- Poeppel, D & Embick, D. (2005). Defining the relation between linguistics and neuroscience. In A. Cutler (Hrsg.), *Twenty-first century psycholinguistics: four cornerstones*. Mahwah (NJ): Lawrence Erlbaum Associates, 103-118.
- Pollock, F. (1878). An infant's progress in language. *Mind* 3, 392-401.
- Pomnitz, P. & Siegmüller, J. (2013). Ergänzende Daten zur Variabilität und Kontinuität im frühkindlichen Lexikon: Eine Tagebuchstudie. *Forschung Sprache* 1, 51-66.
- Poulisse, N. (1999). *Slips of the Tongue. Speech Errors in First and Second Language Production*. Amsterdam: John Benjamins.
- Preyer, W. (1882). *Die Seele des Kindes*. Leipzig: Grieben.
- Prince A. & Smolensky P. (1997). Optimality: from neural networks to universal grammar. *Science* 275, 1604–1610.
- Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics* 25, 423-459.
- Putnam, H. (1960). Minds and Machines. In S. Hook (Hrsg.), *Dimensions of Mind: A Symposium*. New York: New York University Press, 148-179.
- Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.
- Quinn, P. C. (2005). Developmental constraints on the representation of spatial relation information: Evidence from preverbal infants. In L.A. Carlson & E. van der Zee (Hrsg.), *Functional features in language and space: Insights from perception, categorization, and development*. New York: Oxford University Press, 293-310.
- Raffler-Engel W. (1965). Del bilinguismo infantile. *Archivio Glottologico Italiano* 50, 175-180.
- Raimy, E. (2000). *The Phonology and Morphology of reduplication*. No. 52 in *Studies in Generative Grammar*. Berlin: Mouton de Gruyter.
- Raimy, E. (2009). A Case of Appendicitis. In E. Raimy, & C. Cairns (Hrsg.), *Contemporary Views on Architecture and Representation in Phonology*, Cambridge, Massachusetts: MIT Press, 177-188.
- Raimy, E. (2009). Deriving Reduplicative Templates in a Modular Fashion. In E. Raimy, & C. Cairns (Hrsg.), *Contemporary Views on Architecture and Representation in Phonology*, Cambridge, Massachusetts: MIT Press, 383-404.
- Raimy, E. & Cairns, C. (Hrsg.) (2009) *Contemporary Views on Architecture and Representation in Phonology*, Cambridge, Massachusetts: MIT Press.
- Ramers, K.-H. (2001). *Einführung in die Phonologie*. München: Wilhelm Fink.
- Ramus, F., Nespors, M., Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.
- Rice, K. & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In J. Archibald (Hrsg.), *Phonological acquisition and phonological theory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 23-42.
- Robb, M. P., Bauer, H. R., Tyler, A. A. (1994). A quantitative analysis of the single-word stage. *First Language*, 14, 37-48.
- Robinson, J. (2009). Review of Charles O. Nussbaum, *The Musical Representation: Meaning, Ontology, and Emotion*. *Notre Dame Philosophical Reviews*: <http://ndpr.nd.edu/news/23934-the-musical-representation-meaning-ontology-and-emotion/>

- Roca, I. (Hrsg.) (1997). *Derivations and Constraints in Phonology*. Oxford: OUP.
- Roeper, T. & Williams, E. (Hrsg.) (1987), *Parameter Setting*. Dordrecht: Reidel.
- Romani, C. & Calabrese, A. (1998). Syllabic constraints on the phonological errors of an aphasic patient. *Brain and Language* 64, 83-121.
- Ross, A. (1937). An example of vowel-harmony in a young child. *Modern Language Notes*, 508-509.
- Ross, B. H. (Hrsg.) (2004). *The psychology of learning and motivation*. San Diego, CA: Elsevier.
- Rothweiler, M. & Meibauer J. (1999). Das Lexikon im Spracherwerb – Ein Überblick. In J. Meibauer & M. Rothweiler (Hrsg.) (1999), *Das Lexikon im Spracherwerb*. Tübingen. Basel: Francke, 9-31.
- Rothweiler, M. (2001). *Wortschatz und Störungen des lexikalischen Erwerbs bei spezifisch sprachentwicklungsgestörten Kindern*. Heidelberg: Winter.
- Rumelhart, D.E., McLelland, J.L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.
- Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science* 274, 1926–1928.
- Sagey, E. C. (1986). *The representation of features and relations in non-linear phonology*. PhD dissertation: MIT.
- Salidis, J. & Johnson, J. S. (1997). The production of minimal words: A longitudinal case study of phonological development. *Language Acquisition* 6, 1-36.
- Sandler, W. (1993). A sonority cycle in American sign language. *Phonology* 10, 243-279.
- Sarimski, K. (1985). *Sprachentwicklungsskalen*. München: Röttger.
- Sauer, N. J. (2015). *Das Tip-of-the-Tongue-Phänomen. Zur Rolle der Silbe beim Auflösen von Wortfindungsstörungen*. Promotionsschrift: Goethe-Universität Frankfurt.
- Schade, U. (1999). *Konnektionistische Sprachproduktion*. Wiesbaden: Deutscher Universitätsverlag.
- Schiller, N. O., Jansma, B. M., Peters, J. & Levelt, W. J. M. (2006): Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes* 21, 112–140.
- Schröder, C. & Höhle, B. (2011). Prosodische Wahrnehmung im frühen Spracherwerb. *Sprache-Stimme-Gehör* 35, 91–98.
- Schwartz, R. G., Leonard, L. B., Wilcox, M. J. & Folger, M. K. (1980). Again and again: Reduplication in child phonology. *Journal of Child Language* 7, 75-87.
- Schulte, G. (2001). *Neuromythen*. Frankfurt: Zweitausendeins.
- Sciarone, A., van Essen, A. & van Raad, A. (Hrsg.) (1969). *Leiden Studies in Linguistics and Phonetics*. The Hague: Mouton.
- Selkirk, E. (1984). *Phonology and Syntax*, Cambridge, MA: MIT Press.
- Selkirk, E. (1984). On the Major Class Features and Syllable Theory. In M. Aronoff & R.Oehrle (Hrsg.), *Language Sound Structure*. Cambridge, MA: MIT Press, 107-113.
- Selkirk, E. (1990). A Two-Root Theory of Length. In E. Dunlap & J. Padgett (Hrsg.), *University of Massachusetts Occasional Papers in Linguistics 14: Papers in Phonology*. Amherst, MA: GLSA, 123-171.
- Shallice, T. (1979). Case study approach in neuropsychological research. *Journal of Clinical Neuropsychology* 1, 183–211.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Hrsg.), *Sentence processing*. Hillsdale, N.J.: Erlbaum, 295-342.

- Shattuck-Hufnagel, S. (1982). Three Kinds of Speech Error Evidence for the Role of Grammatical Elements in Processing. In L. Obler & L. Menn (Hrsg.), *Exceptional Language and Linguistics*. New York: Academic Press, 133-142.
- Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Hrsg.), *The production of speech*. New York: Springer, 109-136.
- Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik (Hrsg.), *Motor and sensory processes of language*. Hillsdale N.J.: Erlbaum, 17-51.
- Shattuck-Hufnagel, S. & Klatt, D. H. (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.
- Shusterman, A. & Spelke, E. S. (2005). Language and the development of spatial reasoning. In P. Carruthers, S. Laurence, & S. Stich (Hrsg.), *The innate mind: Structure and content*. New York: Oxford University Press, 89-108.
- Sidtis, J. J. & Van Lancker Sidtis, D. (2003). A Neurobehavioral Approach to Dysprosody. *Seminars in Speech and Language* 24(2), 93-105.
- Sievers, E. (1901). *Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf & Härtel.
- Siok W. T., Jin, Z., Fletcher, P., Tan, L. H. (2003). Distinct brain regions associated with syllable and phoneme. *Human Brain Mapping* 18, 201-207.
- Skarabela, B., Fish, S. & Do, A. H.-J. (Hrsg.) (2002). *Proceedings of the 26th Annual Boston University Conference on Language Development*, Vol. 2. Somerville, MA: Cascadilla Press,
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appelon.
- Smit, A., Hand, L., Freilinger, J., Bernthal, J. & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders* 55, 779-798.
- Smith, B. L. (1990). Elicitation of slips of the tongue from young children: a new method and preliminary observations. *Applied Psycholinguistics* 11, 131-144.
- Smith, M. D. & Locke, J. (Hrsg.) (1988). *The emergent lexicon: The child's development of a linguistic vocabulary*. New York: Academic Press.
- Smith, N. V. (1973). *The acquisition of phonology: A case study*. New York: Cambridge University Press.
- Smith, N. V. (2010). *Acquiring phonology: a cross-generational case-study*. Cambridge: Cambridge University Press.
- Smolensky, P. (1999): Grammar-based Connectionist Approaches to Language. *Cognitive Science* 23 (4), 589-613.
- Snow, D. (2001). Imitation of intonation contours by children with normal and disordered language development. *Clinical Linguistics and Phonetics* 15, 567-584.
- Soderstrom, M. (2008). Early perception – late comprehension of grammar? The case of verbal -s: A response to de Villiers & Johnson (2007). *Journal of Child Language* 35, 671-676.
- Soderstrom, M., Wexler, K. & Jusczyk, P. W. (2002). English-learning toddlers' sensitivity to agreement morphology in receptive grammar. In B. Skarabela, S. Fish & A. H.-J. Do (Hrsg.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, Vol. 2. Somerville, MA: Cascadilla Press, 643-652.
- Soderstrom, M., White, K. S., Conwell, E. & Morgan, J. L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy* 12, 1-29.

- Sodian, B. (2014). Entwicklung begrifflichen Wissens: Kernwissenstheorien. In L. Ahnert (Hrsg.), *Theorien der Entwicklungspsychologie*. Heidelberg, Berlin: Springer VS, 122-147.
- Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular *-s*. *Journal of Speech, Language and Hearing Research* 52, 623-642.
- Spelke, E. (1994). Initial knowledge: six suggestions. *Cognition* 50, 431-445.
- Spelke, E. & Kinzler, K. D. (2007). Core knowledge. *Developmental Science* 10 (1), 89-96.
- Spence, K. W. & Spence, J. T. (Hrsg.) (1968). *The psychology of learning and motivation: Advances in research and theory, Vol. II*. New York: Academic Press.
- Spring, D. R. & Dale, P. S. (1977). Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research* 20, 224-232.
- Stachowiak, F. J. (1979). *Zur semantischen Struktur des subjektiven Lexikons*. München: Fink.
- Stemberger, J. P. (1985). An interactive Activation Model of Language Production. In A. W. Ellis (Hrsg.), *Progress in the Psychology of Language*. London: Lawrence Erlbaum, 143-186.
- Stenneken, P., Bastiaanse, R., Huber, W. & Jacobs, A. M. (2005). Syllable structure and sonority in language inventory and aphasic neologisms. *Brain and Language* 95, 280-292.
- Stern, C. & Stern, W. (1928). *Die Kindersprache. Eine psychologische und sprachtheoretische Untersuchung*. (4. Auflage, Nachdruck 1987) Darmstadt: WBG.
- Stevens, Kenneth (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America* 111, 1872-1891.
- Stites, J., Demuth, K. & Kirk, C. (2004): Markedness vs. Frequency Effects in Coda Acquisition In A. Brugos, L. Micciulla, & C. E. Smith (Hrsg.), *Proceedings of the Annual Boston University Conference on Language Development* 28, 565-576.
- Stoel-Gammon, C. & Cooper, J. A. (1984). Patterns of early lexical and phonological development. *Journal of Child Language* 11, 247-271.
- Stoel-Gammon, C. & Dunn, C. (1985). *Normal and disordered phonology in children*. Austin, TX: PRO-ED.
- Stumpf, Carl. (1901). Eigenartige sprachliche Entwicklung eines Kindes. *Zeitschrift für pädagogische Psychologie und Pathologie* 3:6, 419-447.
- Suchodoletz, W. (Hrsg.) (2002). *Therapie von Sprachentwicklungsstörungen. Anspruch und Realität*. Stuttgart: Kohlhammer.
- Sussman, H. (1984). A Neuronal Model of Syllable Representation. *Brain and Language* 22, 167-177.
- Szagun, G. (2006). *Sprachentwicklung beim Kind. Ein Lehrbuch*. Weinheim und Basel: Beltz.
- Taelman, H., & Gillis, S. (Im Druck). Do Dutch-speaking children prefer to produce trochees? A reevaluation of the trochaic template hypothesis of Dutch truncations. *First Language*.
- Thoenen, H. (1995). Neurotrophins and neuronal plasticity. *Science* 279, 593-598.
- Tomasello, M. (2002). *Die kulturelle Entwicklung des menschlichen Denkens*. Frankfurt a.M.: Suhrkamp.
- Tomasello, M. (2003). *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development* 78, 705-722.

- Tracy, R. (1995). *Child languages in contact: The simultaneous acquisition of two languages (English/German) in early childhood*. Habilitationsschrift: Universität Tübingen.
- Tracy, R. (2000). Sprache und Sprachentwicklung: Was wird erworben? In H. Grimm (Hrsg.) (2000), *Sprachentwicklung. Enzyklopädie der Psychologie*, Serie III, Band 3. Göttingen: Hogrefe, 3-35.
- Trubetzkoy, N. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht. (6. Auflage (1977))
- Truckenbrodt, H., (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.
- Trotzke, A., Bader, M. & Frazier, L. (2013). Third factors and the performance interface in language design. *Biolinguistics* 7, 1-34.
- Tropf, H. (1987). Sonority as a variability factor in second language phonology. In A. James & J. Leather (Hrsg.), *Sound patterns in second language acquisition*. Providence, RI: Foris Publications, 173-191.
- Tsakosta (2004). Multiple parallel grammars in the acquisition of stress in Greek L1: <http://www.lotpublications.nl/index3.html>
- Tsujimura, N. (2005). Mimetic verbs and innovative verbs in the acquisition of Japanese. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 31(1), 371-382.
- Tyler, L. K. & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science* 5, 244-252.
- Uhmann, S. (1991). *Fokusphonologie: Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie* (Linguistische Arbeiten Nummer 252). Tübingen: Niemeyer.
- Uttal, W. (1967). Evoked Brain Potentials: Signs or Codes? *Perspectives in Biology and Medicine* 10, 627-639.
- Van den Broecke, M. P. R. & Goldstein, R. (1980). Consonantal Features in Speech Errors. In Fromkin, V. (Hrsg.), *Errors in Linguistic Performance*. New York: Academic Press, 47-65.
- Van Oostendorp, M. (2000). *Phonological projection: a theory of feature content and prosodic structure*. *Studies in Generative Grammar* 47. Berlin: Mouton de Gruyter.
- Van Oostendorp, M. & van de Weijer, J. (Hrsg.) (2005). *The internal organization of phonological segments*. *Studies in Generative Grammar* 77. Berlin: Mouton de Gruyter.
- Vater, H. (1992). Zum Silben-Nukleus im Deutschen. In P. Eisenberg, K. Ramers & H. Vater (Hrsg.), *Silbenphonologie des Deutschen*. Tübingen: Gunter Narr Verlag, 100-133.
- Velleman, S. & Vihman, M. M. (2002). Whole-word phonology and templates: trap, bootstrap, or some of each? *Language, Speech and Hearing Services in Schools* 33, 9-23.
- Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Hrsg.), *Phonological Development: Models, Research, Implications*. Timonium, Maryland: York Press., 393-422.
- Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Oxford: Blackwell.
- Vihman, M. M. & Keren-Portnoy, T. (Hrsg.) (2013). *The Emergence of Phonology. Whole-word Approaches and Cross-linguistic Evidence*. Cambridge: Cambridge University Press.
- Vihman, M. M. (2014). *Phonological development. The first two years*. Malden, MA: Wiley Blackwell.

- Vihman, M. M., Macken, M. A., Miller, R., Simmons, H. & Miller, J. (1985). From babbling to speech: a re-assessment of the continuity issue. *Language* 61, 397-445.
- Vihman, M. M. & McCune, L. (1994). When is a word a word? *Journal of Child Language* 21, 517-542.
- Vihman, M. M., & Miller, R. (1988). Words and babble at the threshold of language acquisition. In M. D. Smith & J. Locke (Hrsg.), *The emergent lexicon: The child's development of a linguistic vocabulary*. New York: Academic Press, 151-184.
- Vihman, M. M. & Croft, W. (2007). Phonological development: toward a 'radical' templatic phonology. *Linguistics* 45, 683-725.
- Wagner, L. & Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science*, 4(2), 177-184.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392-393.
- Webb, S., Monk, C. & Nelson, C. (2001). Mechanisms of postnatal neurobiological development: implications for human development. *Developmental Neuropsychology* 19, 147-171.
- Weir, R. H. (1962). *Language in the crib*. The Hague: Mouton.
- Weissenborn, J. (1996). Kinder steigen mit schlafwandlerischer Sicherheit in das System der Muttersprache ein (Interview). *Frankfurter Linguistische Forschung* 19, 89-102.
- Weissenborn, Jürgen (2000). Der Erwerb von Morphologie und Syntax. in H. Grimm (Hrsg.) (2000), *Sprachentwicklung. Enzyklopädie der Psychologie*, Serie III, Band 3. Göttingen: Hogrefe, 141-169.
- Wells, R. (1951). Predicting Slips of the Tongue. *The Yale Scientific Magazine* 263, 9-30.
- Wells, W. & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech, Language and Hearing Research* 46, 5-20.
- Wexler, K. & Manzini, M. R. (1987). Parameters and Learnability in Binding Theory. In T. Roeper & E. Williams (Hrsg.), *Parameter Setting*. Dordrecht: Reidel, 41-76.
- Wiese, R. (1988). *Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen*. Tübingen: Niemeyer.
- Wiese, R. (1996). *The Phonology of German. The Phonology of the World's Languages*. Oxford: Clarendon Press.
- Wijnen, E. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language* 31, 734-755.
- Winitz, H. & Irwin, O. C. (1958). Syllabic and phonetic structure of infants' early words, *Journal of Speech and Hearing Research* 1:3, 250-256.
- Wojcik, P. & Smoczynska, M. (1997) Acquisition of Lithuanian Verb Morphology. In U. W. Dressler (Hrsg.), *Studies in Pre- & Protomorphology*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 83-100.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1-34.
- Woodward, A. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development* 22, 145-160.
- Woodward, A. & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science* 11, 73-77.
- Wundt, W. (1900). *Die Sprache*, erster Halbband. Stuttgart: Kröner. (4. Aufl., 1921)
- Wurzel, W. U. (1984). Phonologie: Segmentale Struktur. In K. E. Heidolph, W. Flämig, & W. Motsch (Hrsg.), *Grundzüge einer deutschen Grammatik*. Berlin: Akademie-Verlag, 898-990.
- Yang, C. (2002). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

- Yang, C. (2004). Universal grammar, statistics, or both? *Trends in Cognitive Sciences* 8 (10), 451-456.
- Yavaş, M. (Hrsg.) (1994). *First and Second Language Phonology*. San Diego, CA: Singular Publishing Group, 251-266.
- Yavaş, M. (1998). *Phonology: Development and disorders*. San Diego, CA: Singular Publishing Group.
- Yavaş, M. (2003). The role of sonority in developing phonologies. *Journal of Multilingual Communication Disorders* 1: 79-98.
- Yeni-Komishian, G., Kavanagh, J. F. & Ferguson, C. A. (Hrsg.) (1980). *Phonology, Vol.1: Production*. New York: Academic Press.
- Yu, A. C. L. (2003). *The morphology and phonology of infixation*. Doctoral dissertation: University of California, Berkeley.
- Zimmermann, E. T. & Sternefeld, W. (2013). *Introduction to Semantics: An Essential Guide to the Composition of Meaning*. Berlin/Boston: Mouton de Gruyter.

Anhang A

1. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔui	0	Personal	11	0	0	0	0	0	0	0	0	0	0	0	0
ʔuh	0	Personal	19	0	0	0	0	0	0	0	0	0	0	0	0
dɛ	da	Interaktional	1	0	0	0	0	0	0	0	0	0	0	0	1
da	da	Interaktional	13	0	0	0	0	0	0	0	0	0	0	0	13
ʔa	0	Personal	3	0	0	0	0	0	0	0	0	0	0	0	0

2. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔuj	0	Personal	6	0	0	0	0	0	6	0	0	0	0	0	0
ge	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
dɛ	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
da ta do	da	Interaktional	2	0	0	0	0	2	0	0	0	0	0	0	2
di	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
ʔo	0	Personal	5	0	0	0	0	5	0	0	0	0	0	0	0
hi	da	Interaktional	0	1	0	0	0	1	0	0	0	0	0	0	1
ʔu	0	Personal	2	0	0	0	0	2	0	0	0	0	0	0	0

3. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔu	0	Personal	16	0	0	0	0	16	0	0	0	0	0	0	0
ʔuj	0	Personal	21	0	0	0	0	0	21	0	0	0	0	0	0
da / dɛ / dʉ / di / tɛ	da	Interaktional	18	0	0	0	0	18	0	0	0	0	0	0	18
ʔa	0	Personal	2	0	0	0	0	2	0	0	0	0	0	0	0
ʔaw	0	Personal	1	0	0	0	0	0	1	0	0	0	0	0	0

4. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
dɛ	da	Interaktional	2	0	0	0	0	2	0	0	0	0	0	0	2
ʔuj	0	Personal	7	0	0	0	0	0	7	0	0	0	0	0	0
ʔaho	Hallo	Interaktional	1	1	0	0	0	2	0	0	0	0	0	0	0
ʔa ja	0	Personal	1	0	0	0	1	1	1	0	0	0	0	1	0
ʔa.wa	0	Personal	4	0	0	0	4	8	0	0	0	0	0	4	0
dɪ	da	Interaktional	2	0	0	0	0	2	0	0	0	0	0	0	2
da	da	Interaktional	3	0	0	0	0	3	0	0	0	0	0	0	3
dø	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
det	da	Interaktional	1	0	0	0	0	0	0	0	0	0	0	0	1
ʔu	0	Personal	1	0	0	0	0	1	0	0	0	0	0	0	0

5. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔuj	0	Personal	14	0	0	0	0	0	14	0	0	0	0	0	0
ʔu	0	Personal	1	0	0	0	0	1	0	0	0	0	0	0	0
dε	da	Interaktional	5	0	0	0	0	5	0	0	0	0	0	0	5
ʔuja	0	Personal	2	0	0	0	2	4	0	0	0	0	0	0	0
ja.ja	0	Personal	0	0	0	0	4	0	0	0	0	0	0	1	0
da	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
ʔa.ja	0	Personal	2	0	0	0	2	4	0	0	0	0	0	2	0
dɪ	da	Interaktional	3	0	0	0	0	3	0	0	0	0	0	0	3
ʔu / wu	Kuh/muh	Interaktional	7	0	0	0	5	12	0	0	0	0	0	0	0

6. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔuj	0	Personal	9	0	0	0	0	0	9	0	0	0	0	0	0
ʔu	0	Personal	1	0	0	0	0	1	0	0	0	0	0	0	0
dε	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
bo.ʔa	0	Personal	2	0	0	0	0	2	0	0	0	0	0	0	0
ja	0	Personal	0	0	0	0	1	1	0	0	0	0	0	0	0
da	da	Interaktional	2	0	0	0	0	2	0	0	0	0	0	0	2
dɪ	da	Interaktional	2	0	0	0	0	2	0	0	0	0	0	0	2
da	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
bubi/ʔəbəʔ/bubu/ʔuj / ʔoj /	Buch	Interaktional	9	0	0	0	0	10	2	0	0	0	0	3	0
wuwu	Hund	Interaktional	0	4	0	0	0	4	0	0	0	0	0	2	0

7. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔuj	0	Personal	1	0	0	0	0	0	1	0	0	0	0	0	0
ʔu	0	Personal	3	0	0	0	0	3	0	0	0	0	0	0	0
bu	Bus	Interaktional	7	0	0	0	0	7	0	0	0	0	0	0	4
ʔoba	0	Personal	2	0	0	0	0	2	0	0	0	0	0	0	0
da, di	da	Interaktional	4	0	0	0	0	4	0	0	0	0	0	0	4
ʔa.jo	Hallo	Interaktional	6	0	0	0	6	12	0	0	0	0	0	0	0

8. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔo	0	Personal	1	0	0	0	0	1	0	0	0	0	0	0	0
da	da	Interaktional	6	0	0	0	0	6	0	0	0	0	0	0	6
ja	zustimmung?	Interaktional	0	0	0	0	9	9	0	0	0	0	0	0	0
ʔuj	0	Personal	1	0	0	0	0	0	1	0	0	0	0	0	0
ʔa.ʔa	0	Personal	2	0	0	0	0	2	0	0	0	0	0	1	0
ʔe.ja	0	Personal	1	0	0	0	1	2	0	0	0	0	0	0	0

9. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
mama	mama	Interaktional	0	0	10	0	0	10	0	0	0	0	0	5	0
da	da	Interaktional	7	0	0	0	0	7	0	0	0	0	0	0	7
ʔuj	0	Personal	10	0	0	0	0	0	10	0	0	0	0	0	0
ʔu	0	Personal	4	0	0	0	0	4	0	0	0	0	0	0	0
ʔo	0	Personal	3	0	0	0	0	3	0	0	0	0	0	0	0
ʔaj	Bonbon/Luftballon/ Eifer	Interaktional	3	0	0	0	0	0	3	0	0	0	0	0	1
ʔa	Auto	Interaktional	5	0	0	0	0	5	0	0	0	0	0	0	5
ʔajo	Hallo	Interaktional	1	0	0	0	1	2	0	0	0	0	0	0	0

10. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
mama	Erzieherin	Interaktional	0	0	16	0	0	16	0	0	0	0	0	8	4
ʔaja	Hallo	Interaktional	2	0	0	0	2	4	0	0	0	0	0	2	0
ʔija	Nina	Interaktional	1	0	0	0	1	2	0	0	0	0	0	0	0
ʔaj	Mein	Instrumental	10	0	0	0	0	0	10	0	0	0	0	0	0
bow	Buch lesen	Regulativ	5	0	0	0	0	0	5	0	0	0	0	0	0
da	Gegenstand haben woll	Instrumental	2	0	0	0	0	2	0	0	0	0	0	0	0
ʔuj	0	Personal	12	0	0	0	0	0	12	0	0	0	0	0	0
bu:(6x), buj (2x)	Kuh/Hund/Tier spiel	Interaktional	8	0	0	0	0	6	2	0	0	0	0	0	0

11. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
de, dada (2x), dr	da	Interaktional	10	0	0	0	0	10	0	0	0	0	0	2	10
bu.bu, buj	Hund	Interaktional	3	0	0	0	0	2	1	0	0	0	0	1	3
ʔuj	0	Personal	5	0	0	0	0	0	5	0	0	0	0	0	0
bo, bubu (3x)	Buch	Interaktional	7	0	0	0	0	7	0	0	0	0	0	3	1
ʔaja	0	Personal	3	0	0	0	3	6	0	0	0	0	0	3	0

12. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
ʔawvaw	Hund	Interaktional	1	0	0	0	1	0	2	0	0	0	0	0	1
ʔuj	0	Personal	5	0	0	0	0	0	5	0	0	0	0	0	0
ʔo	0	Personal	2	0	0	0	0	2	0	0	0	0	0	0	0
mama	mama	Interaktional	0	0	14	0	0	14	0	0	0	0	0	7	0
da	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1
naj	nein	Interaktional	0	0	4	0	0	0	4	0	0	0	0	0	0

13. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
bubu	auto/bus	Interaktional	8	0	0	0	0	8	0	0	0	0	0	4	4
bo, bobo	auto/bus	Interaktional	6	0	0	0	0	6	0	0	0	0	0	2	0
ʔuj	0	Personal	7	0	0	0	0	0	7	0	0	0	0	0	0
baba, babaj	ball	Interaktional	4	0	0	0	0	3	1	0	0	0	0	2	2
ʔawaw	auto	Interaktional	2	2	0	0	0	0	4	0	0	0	0	2	0
dɛ	das/dies	Interaktional	6	0	0	0	0	6	0	0	0	0	0	0	6
ʔeja	Aufmerksamkeit	Regulativ	2	0	0	2	4	0	0	0	0	0	0	0	0
da	da	Interaktional	1	0	0	0	0	1	0	0	0	0	0	0	1

14. Sitzung

Phonetisch	Zielwort	Akt	PV	FV	NV	LV	GV	V	VG	VL	VN	VF	VP	Reduplikation	Zeigegesten
mama	mama	Interaktional	0	0	2	0	0	2	0	0	0	0	0	1	0
ʔawa	Auto	Interaktional	6	6	0	0	0	3	0	0	0	0	0	3	3
ʔaj, ʔaja	ei, streicheln	Interaktional	6	0	0	0	0	4	6	0	0	0	0	2	0
ʔaw	Auto	Interaktional	10	0	0	0	0	0	10	0	0	0	0	0	3
ʔuj	0	Personal	5	0	0	0	0	0	5	0	0	0	0	0	0
baj	Ball	Interaktional	6	0	0	0	0	0	6	0	0	0	0	0	2
jaj	zustimmung?	Interaktional	0	0	0	0	4	0	4	0	0	0	0	0	0

Lebenslauf

geboren: 02.07.1977 in Offenbach am Main
Familienstand: verheiratet, keine Kinder

Berufsausbildung und Studium

1998 – 2001	Ausbildung zum staatlich anerkannten Erzieher an der Fachschule für Sozialpädagogik Bertha-Jourdan-Schule Frankfurt am Main (Abschluss: 1,1)
2001 – 2007	Magisterstudium Germanistik an der Johann-Wolfgang-Goethe-Universität Frankfurt am Main (Hauptfächer: Sprachwissenschaft/ Didaktik; Nebenfächer: Psychologie und Psychologie; Abschluss: 1,2)

Berufserfahrung

im pädagogischen Bereich

1996 – 1997	Zivildienst in einer Krabbelstube des Sozialpädagogischen Vereins für familienergänzende Erziehung in Frankfurt am Main
2000 – 2001	Anerkennungsjahr als staatlich anerkannter Erzieher im Bergkindergarten Frankfurt am Main
11/2001 – 12/2002	Studentischer Mitarbeiter bei der Ambulanten Familienhilfe der Lebenshilfe für Menschen mit geistiger Behinderung Frankfurt e.V.
3/2003 – 6/2004	Sozialpädagogischer Lernhelfer für den Verein für sozialpädagogische Modelle in Frankfurt am Main

im universitären Betrieb, Forschung und Lehre

10/2003 – 4/2009	studentische und wissenschaftliche Hilfskraft am Institut für Linguistik der Johann-Wolfgang-Goethe-Universität
4/2009 – 4/2014	wissenschaftlicher Mitarbeiter am Institut für Linguistik der Johann-Wolfgang-Goethe-Universität
2009 – 2012	Lehraufträge und externe Gutachtertätigkeiten für die Fresenius Hochschule Idstein
7/2013 – heute	Dozent an der Hochschule Fresenius

Praktika

8/2005 – 9/2005	Praktikum im Clementine Kinderhospital (Logopädie) FFM
9/2006	Praktikum im Pädiatrischen Dysphagiezentrum der Kinderkliniken Prinzessin Margaret-Hospital in Darmstadt

Schulbildung

1983 – 1987	Grunelius-Grundschule Frankfurt am Main
1987 – 1996	Freiherr-v.-Stein-Gymnasium Frankfurt am Main (Abitur)

Persönliches Engagement

2/2005 – 4/2006	Studentische Vertretung im Direktorium des Instituts für Linguistik
6/2010 – 12/2012	Mittelbauvertreter am Institut für Linguistik

Auszeichnungen

2007	Posterpreis der Gesellschaft für Aphasieforschung und -behandlung
2011	Neue Impulse-Preis für die beste Nachwuchsarbeit der Zeitschrift <i>Sprache-Stimme-Gehör</i>

Publikationsliste

- Dümig, S. (2016). Lebendiges Wort? Schopenhauers und Goethes Anschauungen von Sprache im Vergleich. In: Schubbe D. & Fauth, S. R.: *Schopenhauer und Goethe. Biographische und Philosophische Perspektiven*. Hamburg: Felix Meiner-Verlag. <https://meiner.de/schopenhauer-und-goethe.html>
- Dümig, S., Konradi, J. & Leuninger, H. (2016). Sonority in aphasic language production: effects of the sonority dispersion principle on error patterns in demisyllables of varying phonological complexity. https://publications.ub.uni-mainz.de/opus/frontdoor.php?source_opus=53720
- Dümig, S. & Leuninger, H. (2013). *Phonologie der Laut- und Gebärdensprache: Linguistische Grundlagen, Erwerb, sprachtherapeutische Perspektiven*. Idstein: Schulz-Kirchner-Verlag.
- Bader, M. & Dümig, S. (2013): Dissociating grammaticality and word-order choice: A case study on object pronouns in German. In: Wielfaert, T., Heylen, K. and Dirk Speelman (eds.). *Proceedings of the 5th Conference on Quantitative Investigations in Theoretical Linguistics*. 11-13.
- Dümig, S. & Leuninger, H. (2013). Auf ein Wort. *Praxis Sprache* 58/3, 144.
- Bader, M., Dümig, S. & Kops, R. (2012): Linearization Processes during Language Production: A Case Study of Object Pronouns in German. In: *Schmid, U. et al. (eds.). Proceedings of KogWis 2012. 11th Biannual Conference of the German Cognitive Science Society*. 62.
- Dümig, S. (2009). Die Silbe und der Erwerb der Morphosyntax. *Sprache-Stimme-Gehör* 33(2), 100-101.
- Dümig, S. & Frank, A. (2008). The syllable and schwa in first language acquisition: Normal and impaired development. *Frankfurter Linguistische Forschung Sondernummer* 11, 65-90.
- Dümig, S. (2007). *Die Phonologie-Morphologie-Schnittstelle im L1-Erwerb: Eine klinische Einzelfallstudie*. Magisterarbeit, Universität Frankfurt. Unveröffentlicht.
- Dümig, S. (2005): Erster Dolmetschertag Gebärdensprache in Hessen. In: *Das Zeichen* 19: 71, 464-465.